

**The Moral Justification of Retributive
Punishment by Reference to the Notion of Balance**

By Stephen Paul Brown

Submitted for the degree of PhD, March 1998

University of Sheffield

Department of Law

Accepted 1998

Thanks

Sincere thanks to the following for their help: Pete and Linda Birtles; Pete Thompson; John Hopkins; Henrik Olsen; Mark Smith; my Aunt Rose; Mick Cavadino, who co-supervised this thesis for the first half of its duration and in many ways set me on the right path; Lilian Bloodworth.

I cannot thank enough my mother, my father and Nicole for their supererogatory help and love throughout the last four years.

Finally, many, many thanks to Deryck Beyleveld for his guidance, wisdom and patience over the course of my studies.

Note on the Text

In this thesis I have employed the fiction that all agents are male. This is for ease of writing.

The Thesis

There is a school of retributive punishment theory that seeks to justify punishment by reference to the notion of balance. Retributive theories typically refer to the past crime as being the single event justifying intervention. The specific retributive theories considered claim that punishment is justified because it restores a rightful equilibrium upset by that crime. I contend that this school of 'balance theories' can be traced within enlightenment philosophy through the works of Immanuel Kant and Georg Hegel to the present day. It is my further contention that none of the theories considered is successful in justifying punishment. Both Kant and Hegel rely on formal notions of balance and fail to prove that punishment is in fact an appropriate mechanism for the restoration of equilibrium. The twentieth century writers propose a substantive notion of balance in terms of rights/freedoms wrongly appropriated by the criminal (and hence denied to all other members of society). They claim that punishment is justified because of its ability to re-appropriate wrongly taken rights/freedoms, hence restoring the rightful equilibrium. I contend that these theorists fail to justify punishment because they fail to identify the exact locus of imbalance caused by crime. I propose that Alan Gewirth's rights theory is sufficient to ground categorical human rights and, hence, categorically binding laws. I conduct an analysis of the imbalance caused by crime in terms of Gewirth's theory and propose, what I believe to be, an original aspect of right disturbed by the criminal's action: dispositional interest in the value of right. It is my thesis that this is restored by punishing the criminal and that, given the importance of rights, punishment is imperative. Punishment is therefore justified by reference to the idea of restoring an equilibrium upset by crime. This is a retributive justification of punishment.

INDEX

1	<u>Chapter One: A Framework for Considering the Justification of Retributive Punishment</u>
3	1 The Definition of Punishment
5	1.1 Offence/non-offender
5	1.2 Non-offence/offender
6	1.3 Non-offender/non-offence
6	1.4 Offender/offence
9	2 The Moral Foundation
10	2.1 Rational beings
11	2.2 The necessary conditions of agency
13	2.3 Self-regarding implications of the necessary conditions of agency
13	2.3.1 The generic features of agency as necessary goods
14	2.3.2 The generic features of agency as rights
16	2.3.2.1 A contingent argument for the pre-eminence of freedom and well-being as rights
17	2.4 Other-regarding implications: the supreme moral principle
20	3 The Definition of Law
21	3.1 Legal idealism outlined
24	3.2 Law under the PGC - formal definition
26	3.3 Law under the PGC - substantively outlined
27	3.3.1 Suggested outline for specific rights under the PGC
27	3.3.1.1 The right not to be physically harmed
28	3.3.1.2 The right not to be physically coerced
28	3.3.1.3 The right to be rescued
30	3.3.1.4 The right to property as a feature of freedom and both nonsubtractive and additive well-being
32	3.3.1.5 The right not to be physically endangered
33	3.3.1.6 Additive harms

34	3.4 Crime
35	3.5 The legitimate State
36	4 Criminality
39	5 Summary
41	<u>Chapter Two: Kant on Punishment</u>
42	1 Does Kant have a theory of Punishment?
50	2 Kant as a Retributivist
52	2.1 The equilibrium to be maintained by punishment: Kant's notion of right as an external lawgiving protecting freedom
55	2.1.1 Deterrence as a constraint on the will
57	2.1.2 Prevention as a constraint on the will
58	2.1.3 Self-defence as a constraint on the will
58	2.1.4 Enforcement of the law as a constraint on the will
60	2.2 How does punishment restore the equilibrium: punishment as a hindrance to a hindrance of freedom
65	2.3 Punishment as right
66	3 The Shipwreck Example
73	4 Conclusions on Kant's Theory of Punishment
75	<u>Chapter Three: Hegel on Punishment</u>
76	1 Punishment as the Annuling of a Wrong
83	2 Punishment as the Criminal's Right
90	2.1 Qualification
90	3 Modern Interpretations of Hegel
102	4 Reconstruction of a Hegelian Justification of Punishment
102	4.1 Punishment as a restoration of equilibrium on a metaphysical level
104	4.2 The need for the State to do <i>something</i> in the face of crime
107	4.3 Punishment as the attempt to uphold right

108 4.4 Hegel's failure to demonstrate how punishment restores right

109 **5 Conclusions**

111 **Chapter Four: Benefits and Burdens Theory**

112 **1 The Equilibrium which Punishment Maintains**

112 1.1 An equilibrium of the benefits and burdens of social living

116 1.1.1 Rights and the benefits of having rights

118 1.2 An equilibrium spread across the members of society

121 1.3 Criticisms

121 1.3.1 The notion of equilibrium is metaphorical

123 1.3.2 Balance theory has a distorted view of crime

125 1.3.3 Balance theory does not account for the victim

130 1.3.4 The Marxist critique

136 **2 Why the Equilibrium of Benefits and Burdens is Held to be Important**

136 2.1 The equilibrium as a feature of distributive justice

139 2.2 Intuition

143 2.3 Egalitarianism

146 2.4 The Social Contract

152 2.5 Criticisms

152 2.5.1 Intuitive anti-retributivism

153 2.5.2 The incompatibility between determinism and retribution

154 2.5.3 Desert is a necessary but not a sufficient condition of punishment

156 2.5.4 Punishment fails to respect the criminal as an end in himself

159 **3 Punishment's Ability to Restore Equilibrium**

159 3.1 Criticisms

159 3.1.1 Balance theory allows some response but not Punishment

162 3.1.2 Unfairness can't be measured

166	4 Conclusions
168	<u>Chapter Five: The Justification of Retributive Punishment Within a Gewirthian Framework.</u>
170	1 Gewirth on Punishment
173	1.1 Punishment's ability to offset an occurrent inequality of noninterference between agents
183	2 A Gewirthian Justification of Punishment
183	2.1 An equilibrium of dispositional interest in the value of right
184	2.1.1 The notion of "value"
186	2.1.2 Crime as an attack on the dispositional interest in the value of right
190	2.1.3 In what "units" is equilibrium to be measured?
192	2.1.4 The move from the particular victim to Potential Future Victims (PFVs)
196	2.2 How does punishment function to maintain the equilibrium?
198	2.2.1 Strict derivation of the State's duty to punish
204	2.2.2 The requirement of proportionality between punishment and offence
205	2.2.3 Is this theory of punishment deterrent in justification?
206	3 Conclusions
211	References

Chapter One: A Framework for Considering the Justification of Retributive Punishment

This thesis is part of a tradition within moral philosophy. The tradition is one traced through Immanuel Kant and Georg Hegel to modern day thinkers including Jeffrie Murphy, Herbert Morris and John Finnis. The common strand to be found in the works of these writers is that punishment can be conceived in terms of, and justified by reference to, the notion of balance. There is a state of affairs, held to be right, which is disturbed by crime. Punishment is the restoration of the right state of affairs and is justified as such. The specifics of the theory vary from thinker to thinker but the tradition is based around this notion of balance and the restoration of balance. In my thesis I will trace the tradition from Kant to the modern day before setting out my own position within that tradition. I will contend that at most stages there is a degree of progress, at many stages there is also a degree of regress but that at every stage there are significant questions left unanswered.

Kant, commonly held to be the paradigmatic retributivist, set up the basic premises for the theory. He talked in terms of punishment operating as a hindrance to a hindrance of freedom. For Kant crime is definitionally wrong because it limits the freedom of an agent against the will of that agent. Thus punishment is held to be right because it limits the criminal limitation of freedom. One crucial question, which I believe Kant leaves unanswered, is exactly how punishment operates as a hindrance to a hindrance of freedom. As commonly conceived it is too late, after the event, to hinder the hindrance of freedom represented by crime. Another problem in Kant and one which I will contend ultimately renders his justification of punishment unsuccessful, is the confounding of freedom with right. We have to wait until the later work of Hegel for a justification of punishment framed entirely in terms of the notion of right. This I hold to be a step forward in the tradition. However, Hegel's work is hampered by his insistence that punishment is the criminal's right. In other words, not only is punishment right but it is *the* right of the criminal. This tends

to render the justification of punishment unnecessary. The central problem with punishment is that it is something done to the criminal against his will. Furthermore, Hegel, like Kant, fails to fully answer the question of how punishment acts as a restorative (in his case for right, in Kant's for freedom). The twentieth century thinkers take the tradition forward by concentrating on a wrong done to a wider constituency than the immediate victim of the crime. This allows us to conceive of an ongoing wrong that must be addressed if the pre-crime situation is to be restored. This is a step further than either Kant or Hegel but I will contend that because of the notion of right which the modern thinkers rely on they are not successful in offering a justification of punishment. Finally, I will offer a Gewirthian analysis of the rights interests breached by crime and suggest that these interests are restored in the act of punishment.

The purpose of this introductory chapter is to draw the framework within which the rest of my thesis will be laid out. There are principally four issues which I wish to discuss. None of them is the focus of this current work but each is necessary in order for that work to make sense. First, I will discuss the definition of "punishment". This is a problem which has troubled many punishment theories, even preventing some from getting off the ground. I will not offer a conclusive definition of punishment but rather I will state the precise field of study of my thesis. I am not concerned with what punishment is: I am concerned with how, if at all, the practice of deliberately inflicting harm on someone because they have committed a criminal act is to be justified. Second, I will outline a derivation of human rights. This provides me with a substantive moral philosophy on which to base talk of justification. The derivation of human rights presented is that of the philosopher Alan Gewirth. Third I will discuss what bearing this derivation has on the law, outlining the debate between legal idealism and legal realism and offering tentative arguments in favour of the former school of thought. The definition of law offered in turn has a bearing on the definition of crime and the legitimate role of the state. Fourth, I will deal with the notion of free-will necessary in order to hold someone responsible for his actions. Although this subject has filled longer and better works than my

thesis I will suggest that given the derivation of human rights offered by Alan Gewirth, agents are bound to view themselves as being the authors of their own actions. This at least allows us to make sense of the claim that someone is guilty of an offence whatever the solution to the wider debate as between free-will and determinism.

1 The Definition of Punishment

An initial problem facing those who seek to justify punishment is the exact ambit of the object of inquiry. This point was famously made by H.L.A Hart in his essay *Prolegomenon to the Principles of Punishment*. The danger which Hart identified in this work is that of the “definitional stop” (1968, 5-6). This occurs when we start with a definition of punishment and work towards a conclusion which states that certain practices outside of the definition are unjustified. Punishment, for example, may be defined as the harsh treatment of an offender for an offence. This practice is then given a justification and then the claim made that punishment of the innocent is unjustified because, in fact, it does not constitute punishment at all. For example, A.M. Quinton in his essay *On Punishment*, states that retributivism provides us with an answer to the question “when (logically) *can* we punish?”, while utilitarianism answers the question “when (morally) *may* we punish?” (1969, 55-56). Retributivism defines punishment as *inter alia* being of an offender for an offence, while utilitarianism tells us that we may harshly treat such an offender when it is for the greater good of the people. This allows Quinton to conclude that punishment of the innocent is logically impossible unless we twist the word “punishment” beyond its natural meaning (1969, 62; cf Ewing 1972, 138). The problem with such a line of thinking is that while punishment of the innocent may be logically impossible (given Quinton’s definition of that term), harsh treatment of the innocent is certainly not. The definitional stop is actually a classic case of *non sequitur*: even if punishment is defined as the harsh treatment of an offender for an offence, this does not preclude the harsh treatment of the innocent. All it precludes is that this second case should be *called* punishment but there are many instances of state intervention other than punishment which are justified. Thus the fact that

something is not called punishment does not mean that, therefore, it is unjustified. After Rawls, we might call such a practice “telishment”, and consider what it is, if anything, which would justify telishing someone (Rawls 1969, 113).

My thesis is concerned with retributive punishment. Already then the object of inquiry is refined somewhat. It is not so important to me that what I am considering is called punishment to the exclusion of all other practices, or that “retributive punishment” is held to be a tautology on the grounds that all punishment is retributive. For convenience I may write as if this is the case but the issue which actually concerns me is whether there is something intrinsic to the case of an offender who has committed an offence which legitimates state intervention. This is a narrower starting point than, for examples, Honderich and Hart who consider secondary uses of the word punishment, or Flew who includes punishment intrinsic to a system such as the family, school or army (Hart 1968, 5; Honderich 1969, 5; Flew 1969, 87). Which is not to say that state intervention, punishment, call it what you will, is never justified in such instances, but merely to leave this for another inquiry.

Given the above my starting point is the following table:

	<i>Offender</i>	<i>Non offender</i>
<i>Offence</i>	Present inquiry	Mental Illness
<i>Non offence</i>	Dangerousness	Taxation Compulsory Education etc.

This table suggests four discrete areas of state intervention. It is to be noted that it is an exhaustive table although the category of non-offender/non-offence is somewhat of a catch-all containing many disparate grounds of state intervention. The examples given at present (e.g. mental illness, dangerousness etc.) are not supposed to be *the* morally legitimate cases of state intervention under a given heading; at this stage it would be entirely question begging to fill in the blanks

authoritatively. Rather the examples are proposals as to the sorts of things that we might expect to find under each heading.

1.1 Offence/non-offender

This is where a law has clearly been broken in that the *actus reus* has been committed but the necessary *mens rea* is missing: the person who has committed the offence can not be held responsible for it. This includes not only instances of mental illness but also, for example, automata and young offenders. Furthermore, under certain moral codes, this category would include scapegoating. A strict utilitarian must claim that it is legitimate to punish a person for an offence that he has not committed where this serves the interests of utility. Such an argument could be used (by a utilitarian) for keeping the victims of miscarriages of justice in jail if their release would severely dent public faith in the justice system.

1.2 Non-offence/offender

I have suggested dangerousness as a prospective entry in this category although the whole issue of dangerousness is extremely contentious. The possibility that I have in mind is where the state intervenes to treat, train or detain someone who is a proven offender but who has not committed an offence so as to justify this particular intervention under the heading offence/offender. The grounds for the detention are that the individual poses a threat to the public safety over and above that which is normally tolerable. Of course, this is problematic. Even accepting that we can predict dangerousness many would argue that the case therefore becomes one of a non-offender because a person whose behaviour is pre-determined so as to be predictable can not be held responsible for his offence. Ultimately this is a problem for psychiatry rather than philosophy; it may even be that there is no case of legitimate intervention for an offender who has not committed an offence. But this does not remove the logical possibility that there could be such a case and thus the need to consider the particular category. A further possible example of state intervention under this heading would be scapegoating as where a known paedophile is arrested when a horrific

crime has been committed in order to assuage public blood-lust. It may well be that such an intervention is not *morally* justifiable. The point at this stage is that it would represent a case of state intervention against an offender who has not actually committed the offence in question and the question is then whether this sort of intervention *is* ever morally justifiable.

1.3 Non-offender/non-offence

As has already been noted this is a catch-all category. It includes state intervention in cases of taxation, licensing laws, compulsory education, regulatory laws, quarantine etc. Hard-line anarchists and Texans may denounce any such interventions but there is the logical possibility that in regulating citizens' relationships the state will be morally justified in intervening in some non-culpable activities. There is nothing intrinsically morally wrong in driving on the left but the state may be justified in prohibiting all subjects from doing so in order to efficiently and safely regulate road traffic.

1.4 Offender/offence

The object of my inquiry is to ask if there is a *unique* justification for intervention in the case of an "offender for an offence". It is impossible to say at the outset that there is such a unique justification - I have already considered the possibility that there is no unique justification in the category of offender/non-offence. However, my thesis will be an examination of the question as to whether there is something special about this category which justifies state intervention.

The question is formulated as to whether offender/offence offers a unique ground of intervention. To discover that there is nothing unique about this category is not to discover that the state is unjustified in intervening in the case of offender/offence. But it is to discover that offender/offence is collapsible into one or other of the three remaining categories. To take an example, if state intervention is generally justified by reference to the need to protect the public

from harm then it would seem ultimately unimportant whether the individual in question is an offender. A person with a deadly contagious disease poses more threat than many offenders and a serial killer who has become paralysed poses less threat than a mobile priest. The decision to intervene under this hypothesis would be made solely on the basis of a prediction as to whether the individual posed a threat of harm to others. It may be relevant to such a prediction that the individual in question is an offender. But the point is that it *may not* be relevant and thus there is nothing about the category offender/offence which would offer a unique ground of intervention.

To say that there is something unique about the category offender/offence is to claim that those concepts in themselves contain the necessary ideas to justify state intervention. Or as John Kleinig puts it

the grounds of the justifiability or unjustifiability are already encapsulated in the descriptions of the activities concerned.
(1973, 44)

In the wider scheme this reflects the debate between retributive and non-retributive theories of punishment: those theories which justify punishment simply by reference to the crime which has occurred and those which look to some other end. It is my contention that where some other end is looked to then inevitably the category offender/offence loses its specificity. Either that other end is actually identical to the end of intervening in the case of offender/offence, or else it will justify some instances of intervention which are not offender/offence.

The question then, which may be taken as the basis of this thesis, is as follows: *Is there anything intrinsic to the situation of an offender who has committed an offence that justifies state intervention?* It may be noted that the question so formulated entails that I am considering a justification of something akin to what Hart actually reports as his central case of punishment. The difference is that I am looking only to the features outlined in my question in an attempt to justify intervention, unlike Hart who appeals to a General Justifying Aim (GJA) (1968,

8-11). The appeal to a GJA, in my view, renders the category offender/offence an irrelevance. If a principle *justifies* state intervention then this is so whether or not the individual being intervened upon is an offender. In the case of utilitarianism then, the fact that an individual is a criminal is only of relevance insofar as it adds to or detracts from the utility of punishing him: there is nothing intrinsically important about his criminality. Hart's response is to make a distinction between the GJA and the principle of distribution (POD) (1968, 11-13). This includes the issue of liability or "who may be punished" (1968, 11). But in this case the POD itself becomes no more than a definitional stop designed to ensure that the GJA does not justify harsh treatment of the innocent. Either there will be cases of conflict between the GJA and the POD or they are derived from, or subject to, some third unifying principle. Indeed, in one of his examples Hart posits both a GJA and a roughly retributive POD that are derived from a unificatory utilitarian principles (1968, 11-12). It is by appeal to such a unifying principle that my own justification of punishment will proceed and it is my contention that the category of offence (and hence offender) is a derivation of this principle.

It is to be noted that the inquiry into punishment is an inquiry about intentional behaviour. This follows from the fact that the typology developed above is a typology of state intervention. I am interested only in acts that the state deliberately performs upon an offender. Thus, for my purposes, to talk of an offender whose marriage has split up and who is wracked with guilt as having been punished enough, must be regarded as a metaphorical use of the word "punishment". Albert Camus says that "punishment without judgment is bearable . . . it is called misfortune" (1963, 57). This delimitation of the object of inquiry is not controversial. John Kleinig correctly observes that it only makes sense to demand a *moral* justification of something that is intentional (1973, 21-22). It should be added though that my particular inquiry is limited to punishments carried out by the state or on behalf of the state or under the sanction of the state. While this removes, for example, most cases of intra-family discipline from the investigation it does not necessarily remove all such cases or cases of personal justice, as where I pursue the one who has injured me and exact

retribution on him. What is important here is whether the activity is legitimate under the state law: whether the state could allow that such activity constitutes rightful punishment or whether it would have to prohibit such instances.

I consider the project of inquiry into the justifications of punishment to be a third step in morally grounded theory. The first step is made by Alan Gewirth in establishing that there exist dialectically necessary moral obligations. The second step is made by Deryck Beyleveld and Roger Brownsword in showing that on the basis of this dialectically necessary morality a legal theory can be established, thus resolving the debate between legal idealism and legal realism in favour of the former. The third step then - the one which I am attempting in the field of punishment, and which others have already made in other fields - is to apply this legal idealism to problems of the legal system. The point for current discussion is that because I come from this tradition, the categories of offender (hence non-offender) and offence (hence non-offence) are not problematic in a certain sense. They are non-problematic in that it is not merely a definitional step to talk of offenders and offences. A legal idealist will have a grounded (i.e. non-positive) notion of what an offence is. In conjunction with a limited theory of free-will this gives us a grounded notion of what an offender is.

2 The Moral Foundation

The starting point for my consideration of morally justified punishment is to establish a moral base. The moral base on which I shall be resting my theory is neither unique nor original but runs against the prevailing currents in modern thought by claiming to be categorically binding. That is to say, it is held to be uniformly obligatory for all its addressees. The constituent group to whom it is addressed is all rational beings. The moral theory expounded is within the tradition developed by Immanuel Kant and Alan Gewirth. The terminology used is drawn eclectically from the works of these writers as well as the various commentaries. The project of justifying punishment may be said to be Kantian as it is Kant who attempted to ground retributive punishment and who is recognised in the literature as the model retributivist. However, the immediate derivation of

a categorically binding moral code may be said to be Gewirthian for it is the later scholar who more successfully establishes his supreme principle.

2.1 Rational beings

The argument progresses from the viewpoint of a particular member of the eventual moral constituency. For Kant this is a rational being with a will, for Gewirth an agent (Kant 1987, 40-41; Gewirth 1978, 26-27). The terminology is different but the essential features are the same. In Kant the will is the faculty by which individuals attain ends; it is the ability to act without constraint, that is freely. A rational being is a being with reason, which is the ability to choose one's ends. Thus a rational being with a will is one with the ability to choose and bring about its own ends. Of course all sorts of contingent factors like lack of strength, resolve, or resources may impinge upon such projects but the essential feature is that the individual should have the faculty for choosing and executing its ends.

It must be admitted that this distinction between the ability to choose ends and the ability to execute those ends is not always maintained in Kant. In *Religion within the Limits of Reason Alone* the contrast made is as between *Willkur* and *Wille* (Kant 1960; Sibley 1960, *xcv-civ*). *Willkur* is a cross between will and reason as conceived above; it is the ability to freely choose between two alternatives. *Wille* is an element of reason as detailed above but it does not operate freely or make choices. Rather it is a constant; it is the constant, "purely rational aspect of the will" (Sibley 1960, *civ*). As such it has the potential to operate as a motivation for action on the *Willkur*. We see in this formulation of Kant a difference of emphasis. The important matter, from our point of view, is that the same essential elements of a free ability to choose and act upon various ends are present.

Gewirth agrees with this description of the individual addressee of moral theory, stating that

however much the persons addressed by various moral and other practical precepts may differ in other respects, it is true of all of them that they are assumed to be able to control their relevant behaviors by their unforced choice for reasons and purposes they can make their own. (1978, 30)

These generic criteria of action Gewirth summarises as “voluntariness and purposiveness” and those possessing them are described as agents¹ (1978, 31). For my own part I will most often use, as Kant himself frequently does, the abbreviated terminology of “a rational being” or Gewirth’s terminology of agency, with the specification that these denote the essential features attributed to “a rational being with a will”, or a being with *Wille* and *Willkur*, or “an agent”, outlined above.

2.2 The necessary conditions of agency

Agency as defined above has certain necessary preconditions of its effective operation. The first and most obvious (given the definition of agency it is in the nature of a tautology) is that the agent be free. The agent must actually be able to choose between two given ends and choose an effective means of bringing the chosen end about; these processes must be under his control. This fact is recognised by both Kant and Gewirth.

In Kant two senses of the word “freedom” are apparent, an internal sense and an external sense. Internal freedom is the freedom of the will discussed above. That is the freedom to set ends and formulate ways of pursuing them. External freedom is the practical condition of internal freedom having any sensible effect, defined by Kant as “independence from being constrained by another’s choice” (1991, 63). Gewirth notes in addition that the agent should be free from any illness, reflex or ignorance that cause him to act beyond his control (1978, 31). Freedom is described by Gewirth as a *generic feature* of agency because it is a necessary characteristic of all action (1978, 25).

¹ More strictly *prospective purposive agents* or PPAs connoting not only those who are acting voluntarily and purposively but also those who intend to and those who have the capacity so to act (Beyleveld 1991, xxxvi; see also Gewirth 1978, 62).

Of course, as has been frequently noted, to be free is in the nature of a negative precondition for the effective operation of the will: I may be free but any number of contingencies may prevent me from actually seeing through my projects. As a rich man knows and a poor man understands we are all free to dine at the Ritz. Gewirth's distinctive contribution to the debate, a point which Kant misses and which perhaps ultimately prevents his derivation of the categorical imperative from being entirely successful, is to add a positive precondition for the will's effective operation. This is the agent's well-being (Gewirth 1978, 53 *et seq.*; cf Kant 1991, 63 where he holds that the only innate right for agents is freedom).

[I]t is not the particular purposes and outcomes but rather the generic abilities and conditions that for the agent primarily constitute his well-being, since they are the necessary conditions of *all* his pursuits of his purposes. (Gewirth 1978, 61; my emphasis)

Well-being is a three dimensional concept, consisting of *basic*, *nonsubtractive* and *additive* goods which are necessary to the pursuit of goals. The basic goods are those such as life, food and shelter, physical, mental and psychological welfare. Nonsubtractive goods are the agent's abilities to hold on to whatever other goods he already possesses. Additive goods are his abilities to add to the goods he already possesses (1978, 54-56).

Freedom and well-being (in the three senses above) can be occurrently or dispositionally necessary (1978, 52, 58). A good is occurrently necessary insofar as it is required in order to achieve a particular goal in which the agent is now engaging. A good is dispositionally necessary insofar as it is required in order to achieve goals generally speaking.² It is freedom and well being in this dispositional sense which together form the generic features of agency (see also Beyleveld 1991, 20). It is the identification of well-being (including as it does the need for life, food and shelter etc.) as a generic feature of agency which gives

² Dispositional non-subtractive well-being consists of whatever faculties are generally necessary to retain the goods one already has whatever they may be. Similarly, dispositional additive well-being consists of whatever faculties are generally necessary to add to the goods one already has whatever they may be (1978, 58-59).

Gewirth's scheme a richness lacked by Kant.³ Although the argument is complex it is already possible to see an emerging hierarchy of necessary goods, which will eventually correspond to a developed hierarchy of rights.

2.3 Self-regarding implications of the necessary conditions of agency

Gewirth's argument (and parts of Kant's argument in the *Groundwork to the Metaphysics of Morals*) proceeds on a dialectical basis. This means that it forms an internal dialogue on the part of an agent drawing necessary conclusions from given premises. The argument is furthermore described by Gewirth as producing statements which are dialectically *necessary* because it proceeds from premises with which any rational agent is bound to concur; that is to say, the particular premises employed by Gewirth's agent are not just given but necessarily accepted (1978, 43). The statements produced by such an argument do not have the status of assertoric truth, but nevertheless from the point of view of the agent they must, on pain of self-contradiction, be taken as true. It is an essential feature of rationality that an agent cannot hold two contradictory statements to be simultaneously true. Thus, if the dialectical method binds the agent to hold that statement P is true then that agent is logically bound to accept that -P is false.

2.3.1 The generic features of agency as necessary goods

An agent is one who does, or can, or at some point will, act. This is analytically true given the definition of an agent in 2.1 above. An agent is bound to view the particular end for which he acts as *good*. Gewirth states that

[t]his conception of worth constitutes a valuing on the part of the agent; he regards the object of his action as having at least sufficient value to merit his acting to attain it. . . (1978, 49)

It is to be emphasized that this is a dialectical rather than assertoric truth. What is necessary is that the agent accepts his particular end as good. Anyone else may

³ This feature gives Gewirth's supreme moral principle substantive and formal aspects as opposed to Kant's categorical imperative which is commonly taken as being a merely formal statement of

view his end as wasteful, bad, or vile, but the agent at least values it as being good enough to be worth acting upon. The argument then proceeds from the claim, necessarily accepted by all agents, that “I view the end for which I am acting as being good” to the claim, necessarily accepted by all agents, that “I view the generic features of my agency as being necessarily good”. These generic features - freedom and well-being - together represent

the preconditions necessary to the existence of any agent’s purposive actions viewed generically and collectively. *Hence, since the agent regards his purposes as good, he must, insofar as he is rational, regard these conditions as at least instrumentally good, whatever his particular contingent and variable purposes and evaluations.* (Gewirth 1978, 54; my emphasis)

If an agent must view A as being good and B is a necessary precondition of A then the agent must view B as being “at least instrumentally good”. The point is that whereas A (the particular end of any agent) is a variable, B (freedom and well-being, the generic features of all action) is constant. Therefore all agents must accept, on pain of contradicting that they are actually agents (i.e. beings who do, can, will act), that freedom and well-being are necessary goods. “Necessary” means “necessary to action” and action is something to which the agent is *logically* committed. It is to be noted that the above argument holds true even where the agent decides that his particular end A is bad and so not to pursue it⁴; the conclusion follows simply from the fact that he is an agent, not from particular end A.

2.3.2 The generic features of agency as rights

We have seen above why the agent is logically committed to his freedom and well-being as necessary goods. The next stage in Gewirth’s argument is to establish that the agent is also logically committed to claiming these necessary goods as being his rights. The argument does not yet introduce a moral element; all that is proven at this stage is that the agent must make a *prudential* claim on all others that they do not interfere with his freedom and well-being (Gewirth

what a moral principle is rather than a moral principle *per se*.

1978, 71; also at 79). This necessitates an explanation of Gewirth's concept of a right. He states that "the concept of a right involves the concept of something due to the subject or right-holder, something to which he is entitled" (1978, 73). What Gewirth is attempting to establish at this stage of his argument is that the agent must (prudentially, i.e. in the interests of achieving any ends) hold that he has rights to his freedom and well-being. Synthesizing this with the above definition of right we have it that the agent must claim as an entitlement his freedom and well-being. And the claim to entitlement necessitates the claim that others do not interfere with the agent's freedom and well-being.

Gewirth presents what he calls a "direct argument" for this contention in what I will summarize as four stages (1978, 78-79). First, he points out that a right as he sees it is directly correlative with a duty, that is a strict ought-judgment that others ought to do/not do X, X being the content of the right. Second, if the agent must claim that freedom and well-being are necessary to his agency then he must also claim that it is necessary that others do not interfere with them. It would be self-contradictory (from the point of view of agency) for the agent to hold that freedom and well-being are necessary to him and that it is okay for others to interfere with these goods. Third, because the agent values his particular end he must actually advocate his having freedom and well-being (the necessary pre-conditions to his pursuit of his end). With regard to his *particular* end the agent is bound to the *evaluative* claim "it is good". This is more than a description of fact because for the agent, from the viewpoint of agency, "good" entails that the end lies above other ends in a hierarchy of worth. In a very simple model where the choice is between doing X and -X, if the agent chooses X then at least he holds X to come above -X in a hierarchy of what is worth doing. Because the particular end is held to be good, Gewirth contends, the necessary preconditions of that end must similarly be held to be good. Hence, it is not simply the case that the agent does not *want* others to interfere with his freedom and well-being as the necessary pre-conditions of his particular end, he must value it as good that others do not interfere with them; he must demand that they leave his

⁴ This choice itself forms an end C, which is "pursued" if only by omission.

freedom and well-being alone. This moves the agent's claim from the descriptive "it is necessary that others do not interfere with my freedom and well-being" to the prescriptive "others ought not to interfere with my freedom and well-being". The latter prescriptive claim is necessary in order that the former descriptive claim have any practical import. Lest Gewirth be accused of is-ought subterfuge it is important to remember that the argument is dialectical, proceeding from the internal viewpoint of the agent.⁵ Thus, for the agent, as one who does, can or will engage in action, it is necessary to claim that others strictly *ought* not to interfere with his freedom and well-being if he is to be an agent. The fourth stage of Gewirth's argument, bringing together the first and the third is that the claim "others ought not to interfere with my freedom and well-being" is equivalent to the claim "I have a right to my freedom and well-being".

2.3.2.1 A contingent argument for the pre-eminence of freedom and well-being as rights

This work is about the justification of retributive punishment rather than being a concerted effort to establish Gewirth's rights thesis as sound, albeit that my own argument takes as its starting point a grounded theory of rights. I believe that the "direct argument" presented above is valid but for reasons of space it is impossible to defend it against possible criticism. However, if the argument is queried there is a telling contingent argument to demonstrate the pre-eminence of freedom and well-being as rights. To say that it is contingent is to admit that it is not proof against all objection. That said, the vast majority of people, because of certain presuppositions that they hold to be absolutely true, will be bound to concur with it.

The argument may be stated thus:

⁵ For his own part Gewirth makes the good point that the word "must" can be substituted for both the *is* and *ought* elements in the development of this argument. Thus the claim "my freedom and well-being are necessary goods" is equivalent to "I must have my freedom and well-being" while the claim "others ought not to interfere with my freedom and well-being" is equivalent to "others must not interfere with my freedom and well-being". The word "must" in this context is expressive of the "agent's own requirements for agency" (Gewirth 1978, 81; see also 102; also Toddington 1989, 451).

If an agent is to claim any rights whatsoever *then* he must claim rights to freedom and well-being. (cf Gewirth 1978, 63; Beyleveld 1991, 24)

The contingent clause (*if* . . . *then*) is obvious. The attraction of it is that it must be accepted by anyone who holds that he has rights. This includes the vast majority of people: most accept that there are some rights, disagreement occurs as to their ambit. Having accepted that he has any rights the agent must claim a right to freedom and well-being because together freedom and well-being would be a pre-condition of the exercise of any other right. Just as freedom and well-being are necessary to the pursuit of any goal, so the right to freedom and well-being are necessary to the pursuit of any right. This is also sufficient to establish that whatever other rights an agent thinks he has he must allow that the rights to freedom and well-being are more important.

2.4 Other-regarding implications: the supreme moral principle

So far Gewirth's argument has led us to the statement, made on pain of contradiction, by all agents (or in the second formulation given above, on pain of contradiction by anyone who believes he has rights), that "I have rights to my freedom and well-being". As Gewirth himself points out this is insufficient for a moral theory. For one thing the statement in the form given above has no other-regarding element (cf Gewirth 1978, 3). The next stage of his argument then is to move from the self-interested, prudentially made statement above, to an other-regarding moral principle for action.

The starting point for this stage is the "logical principle of universalizability" which states that

if some predicate P belongs to some subject S because S has the property Q (where the 'because' is that of sufficient reason or condition), then P must also belong to all other subjects S_1, S_2, \dots, S_n that have Q. If one denies this implication in the case of some subject, such as S_1 , that has Q, then one contradicts oneself.

For in saying that P belongs to S because S has Q, one is saying that having Q is a sufficient condition of having P; but in denying this in the case of S_1 , one is saying that having Q is not a sufficient condition of having P. (Gewirth 1978, 105)

Gewirth is at pains to point out that the logical principle of universalizability is not itself a substantive moral principle. Indeed properly regarded it is no more than a treatise on the meaning of the word 'because' when used to import sufficient reason. This is as against Kant. At least the first formulation of the categorical imperative seems like a version of the principle of universalizability, and Kant's theory has faced immanent critique at the point where he attempts to derive substantive duties from what is an apparently formal statement (Kant 1987, 49, 50-51).

The opening manoeuvre then is the logical inference

[S] has [P] because of [Q] > [S_n] which has [Q] has [P]

Substituting terms this becomes

[Agent] has [rights to freedom and well-being] because of [his having purposes he wants to fulfill] > [Any agent] which has [purposes he wants to fulfill] has [rights to freedom and well-being]

This is a logical inference from the statement "I have rights to freedom and well-being because I have purposes which I wish to fulfill" which in turn, as we have already seen, is a statement to which the agent is bound on pain of contradicting that he is an agent. Hence, the agent is also bound to the claim "all other agents which have purposes they want to fulfill have rights to freedom and well-being", similarly on pain of contradicting that *he* is an agent. Given our definition of agency this statement may be simplified to "all other agents have rights to freedom and well-being". All of this follows inferentially because the agent is committed to recognising "having purposes I want to fulfill" as the necessary and

sufficient ground for the claim “I have rights to my freedom and well-being” (see Gewirth 1978, 110).

This later leads Gewirth to a formulation of his supreme moral principle, the principle of generic consistency (PGC) which is given as:

[a]ct in accord with the generic rights of your recipients as well as of yourself. (Gewirth 1978, 135)

A recipient is any prospective agent whose agency is effected by the actions of the agent addressee of the PGC. Because the argument proceeds dialectically from the internal viewpoint of any agent the conclusion reached has equal force for each agent: each agent is bound to accept the PGC on pain of contradicting that he is an agent. There is a transition from prudential to moral principle inherent in this final stage of the argument because the prescriptive statement (the PGC) to which the agent is now bound takes into account the interests of all other agents (Gewirth 1978, 145-147). Once again, it is to be remembered that the argument to the PGC is dialectically necessary. Thus a full statement would run:

From his own internal viewpoint as an agent, each agent is bound to act in accordance with the principle that he should act in accord with the generic rights to freedom and well-being of every other agent.

But given that this is the case for every agent the PGC has the same potency for each agent as it would have if it had been shown to be assertorically true (see also Beyleveld 1991, 45; cf Kant’s use of similar reasoning in Kant 1987, 79-80).

In considering the implications of the PGC Gewirth states that it binds the agent to treating

other persons as well as himself as persons and not as things or objects whose only relation to himself in transactions is that of facilitating his own purpose-fulfillment. (1978, 140)

This has definite resonance with Kant's second formulation of his categorical imperative that demands of the agent addressee:

[a]ct in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means, but always at the same time as an end. (Kant 1948, 91)⁶

The essential features of these moral principles are strikingly similar. They both address a being who acts freely for his own purposes and demand that he treats others (who act freely for *their* own ends) as of meriting respect similar to that which he requires for himself.⁷ It is thus that this thesis is written within the Kantian-Gewirthian tradition. I have only presented Gewirth's derivation of the supreme moral principle because, in my opinion, his derivation succeeds at points where Kant's fails. Thus the PGC, which grants that all agents have rights to their freedom and well-being, is the starting point and constant touchstone for the following consideration of whether retributive punishment can be justified.

3 The Definition of Law

Having established a supreme moral principle we must now consider the legal import of this. Crucially, it is my contention that in accordance with legal idealism (or natural law theory), and as against legal realism (or legal positivism), legal duties are conceptually related to moral duties. This section of the argument is thus a development of Gewirth's proof for the PGC. However, it is Kant rather than Gewirth who is to be typified as the legal idealist and so, much of the following will be a Kantian treatment of Gewirthian material.⁸

⁶ Paton's translation given here is less archaic than that given in Kant (1987, 58).

⁷ The difference as noted is that a Kantian being demands only freedom for himself while Gewirth's agent requires both freedom and well-being.

⁸ John Ladd points out that there are major differences between Kant *qua* natural lawyer and other theorists to whom that label might be applied. Principally, the knowledge of Law comes from within (rather than from God or nature), and is known *a priori* (Ladd 1965, xvii). This much may also be said of a legal theory developed out of the PGC even though Gewirth does see himself as a natural lawyer (Gewirth 1991, xiv).

3.1 Legal idealism outlined

Legal idealism is the thesis that law necessarily has a determinate moral content and hence that an immoral “law” is really no law at all. If we accept the PGC as the supreme moral principle then it follows that only those prescriptions which are in accord with the PGC (in the negative sense of not contradicting it) can possibly be accurately described as moral.⁹ This much may be conceded by a legal positivist. But a (Gewirthian) legal idealist will further insist that only those prescriptions which are in accord with the PGC (in the negative sense of not contradicting it) can possibly be accurately described as *legal*. This amounts to the claim that legal validity is a subset of moral validity and any attempt to legally oblige is an attempt to morally oblige.

This claim is to be seen in Kant’s Introduction to the *Doctrine of Right* where he describes *Jurisscientia* as the “systematic knowledge of the doctrine of natural Right” and states that this contains the “*immutable principles for any giving of positive law*” (1991, 55; emphasis changed; see also Kant 1929, 302; 1956, 32 and esp. 34; 1983, 78-79). To talk of such immutable principles is precisely to hold that the mere positing of a rule does not make it Law. Rather, it must conform to the external canon of morality.¹⁰ Once again, my thesis is not the place to give a full defense of natural law theory. However, I do believe that a justification for the central claims of the theory can be presented in a relatively brief space.

The PGC, which is rationally binding on all agents, prescribes certain norms for action. Primarily it prescribes that every agent act in accordance with the generic rights to freedom and well-being of his recipient as well as himself. In other

⁹ Which is not to say that every prescription which does not contradict the PGC is *ergo* moral: the prescription might be morally neutral.

¹⁰ see further Gregor (1963, 34-49); Weinrib (1987, 472-508); Aune (1979, 133-141) who holds that the derivation of Law from the Categorical Imperative is unsuccessful but grants that this is what Kant was actually attempting; cf. Fletcher (1987, 535-536) who states that Kant’s theory of *Recht* “does not refer to legal rules and principles actually binding at some moment of time”. The point, for current purposes, is that, if my interpretation of Kant is correct, then any rule that failed to conform to *Recht* would not be “actually binding”. Fletcher also attempts to distinguish, in Kant’s work, “two different bodies of thought - . . . the moral and the legal” (1987, 558).

words it grants each agent rights to his freedom and well-being, which are to be respected by every other agent. As the supreme moral principle the PGC must be able to prescribe action for every potential "morally relevant" situation (Gewirth 1978, 199).

In a purported legal order which prescribes that its subjects should infringe the freedom and well-being of others, an agent-subject would be presented with two incompatible prescriptions. The first, borne out of the PGC, demands that he act in accord with the generic rights of his recipient as well as himself. The second, borne out of the purported legal order, demands that he act against the generic rights of his recipient. Logically, the subject-agent cannot observe both demands. He is bound, therefore, to chose one (and ignore the other) as the maxim of his action. But we have already established, on the basis of the dialectically necessary argument for the PGC, that the agent is bound, on pain of denying that he is an agent, to act in accordance with the PGC. Since the agent can only be subject to one of two conflicting "rules" no order positing prescriptions which contravene the PGC can place an obligation on the agent.¹¹ The most that such an order could do is force the agent through exercise of physical or psychological power to act in accordance with its prescriptions. It follows, from the fact that the PGC is the supreme *moral* principle, that such an exercise of power must be morally unjustified, that is, immoral.

It is to be supposed that a legal positivist might go along with much of the above and yet still insist that there is a difference between moral obligation and legal obligation. Hence, he would argue, a rule which is substantively immoral can nevertheless impose obligation from the legal point of view (see e.g. Hart 1961, 84). But this argument is refuted at the conceptual level for the reason that agency is epistemologically prior to both morality and law. By this it is meant that both morality and law are purposive enterprises and therefore presuppose

¹¹ NB if I tell you that you are obliged to do X but the PGC requires that you do -X I put no obligation on you. As you are obliged to do -X you cannot be under any obligation to do X notwithstanding what I might say or threaten. Viewed this way the situation is not one of two conflicting obligations but rather of one obligation which negates any attempt to posit a conflicting obligation.

agency; both assume that there are agents with control over their actions. Therefore any conclusion which necessarily follows from agency is binding on purported moralities and purported laws. To argue otherwise is to fail to recognise that both morality and law presuppose agency and hence its generic features. Agency, with this characteristic of epistemological priority, leads us to the conclusion of binding prescriptions (the PGC). The PGC in turn is *the* supreme practical principle for guiding all agency. It allows of no contender, and this is not merely a capricious *ipse dixit*, but a rationally necessary conclusion of agency. Thus, all agents are bound on pain of denying that they are agents, to rebut any practical principle that fails to conform to the PGC. In turn this means that any law not conforming to the PGC, *from the point of view of the agent*, cannot generate obligation; it cannot operate as a practical principle *for him*.

We have several conclusions about the purported legal system that posits rules in contravention of the PGC. The first is that the rules it posits are immoral. The second is that the attempt to enforce the rules is an immoral use of power. Because the PGC is the supreme practical principle these conclusions lead us to the third conclusion that immoral "laws" generate no obligation on the agent. But, and this is crucial, *conceptually* a law is something which generates obligation. Finally this leads us to conclude that an immoral law, because it generates no obligation, fails to *be* a law. This is the central tenet of legal idealism.

Even if someone were to stubbornly stand by legal positivism there is a more direct argument (which if taken far enough would entail legal idealism in any event) to persuade him of the PGC's priority over positive law. This involves taking the position of a legislator. A legislator, being an agent, is bound to recognise the PGC as categorically binding. This means that he would contradict himself as an agent if he were to attempt to pass any statute incompatible with the PGC. Hence, on all issues involving rights he is bound to pass laws that are congruent with the supreme moral principle.

3.2 Law under the PGC - formal definition

Having shown why it is necessary for laws to be compatible with the PGC it is now required to consider what class of the genus “things compatible with the PGC” constitutes law. At this stage my answer to the question is formal rather than substantive which is to say it offers a definition of law rather than a complete specification of what the laws are.

The PGC is obligatory for all agents whether or not the legal system recognises this is so, or indeed, whether or not there is a recognisable legal system. The PGC allows of no Hobbesian state of nature antecedent to rights and duties. The only state of nature that would not engender rights and duties is a state of nature in which there are no agents.¹² A legal system is legitimate in so far as it doesn't contravene the PGC. Beyleveld and Brownsword, working out the implications of the PGC for an ideal legal system, offer the following definition of a law:

Rule X is a law if, and only if, there is a moral right to posit rule X for attempted enforcement, which is the case if, and only if,

- (a) Rule X is posited by an authorised source;
- (b) The norm posited by rule X is not immoral. (Beyleveld and Brownsword 1986, 180)

Given the primacy of the PGC these two requirements could be collapsed into the single requirement that a law is a rule that doesn't breach the PGC. However, because we have no philosopher king to work out with unerring precision the dictates of the PGC, there is a need to consider the two criteria separately. Specifically, there is a need to allow for “an authorised source” which will have the powers to pass laws in line with the PGC. This is so because there are possible areas for disagreement over the application of the PGC and in such areas there must be some way of settling dispute. An important point to note is that when we talk of an “authorised source” the authority must itself be derived

¹² Even a state of nature with only one agent would entail rights and duties. The individual agent would still be rationally bound by the PGC albeit that as an empirical contingency there would be no other agents to whom he could owe duties and against whom he could hold rights.

from the PGC. The positing of rules is a purposive exercise and therefore comes under the jurisdiction of the supreme moral principle.

Our formal definition of a law then is “a rule which there is a moral right to posit for attempted enforcement” (Beyleveld and Brownsword 1986, 160).¹³ It might be thought that this begs the question as to punishment but enforcement does not necessarily entail punishment (certainly not without further argument). “Enforcement” could mean the setting up of prophylactic mechanisms; for instance, a police force that patrols the streets in order to prevent assaults. Also, given the substantial areas of law which are not criminal law, enforcement could mean the institution of mechanisms for the regulation of relationships; for instance, courts to settle disputes in contract, in marriage, in property etc.

Criminal law directly protects the most important interests of the agent. Positively, the criminal law is generally distinguished (or distinguishable) from the civil law by reference to the types of enforcement deemed appropriate. Thus, assault is held to merit imprisonment while breach of contract is dealt with by the imposition of damages. But the very question we are seeking to answer concerns what it is that justifies punishment. It would therefore be question begging to offer as the distinction between criminal and civil law that the former is dealt with by punishment while the latter is not. Under such a scheme our definition of crime would be “laws which when broken deserve punishment” which vacuously answers the question “what justifies punishment?” with the answer “the fact that a crime has been committed”. The answer is also based on a stipulated presupposition as there is no *a priori* reason why being imprisoned is punishment and being forced to pay damages is not.

However, it is sufficient for the time being to give a loose definition of crime. This can be achieved by considering breaches of the most important goods (which most people and legal systems would agree in any event *do* constitute

¹³ cf. Kant (1991, 55) who defines law as the totality of moral duties which it is possible to frame as external legislation; also Gregor (1963, 34-35).

crimes) and stipulating that these are crimes. Having offered a justification of punishment it would then be possible to return and expand upon this.

3.3 Law under the PGC - substantively outlined

As we have provisionally seen in deriving the PGC, there is a hierarchy of necessary goods and it is the upper end of this hierarchy that the criminal law is designed to safeguard. Gewirth states that

[o]ne kind of action X is more wrong than another kind Y if X is more harmful than Y, in that X tends to lower its recipient's capacity for action more than does Y, where he has a right that such capacity not be lowered. (1978, 236)

It is not necessary for the purposes of my thesis to delineate all those actions in breach of moral duty that constitute crime. All that is required is that it is shown that *some* breaches constitute crime. The rest of this thesis can then be read with a generalised notion of crime in mind; the moral justification of punishment will then apply *mutatis mutandis* to any other action that is criminal. Effectively this leaves the complete explanation of the PGC's application to the criminal law for another project.

As we have seen the PGC grants rights to freedom and well-being to all agents. This is more than an abstract and/or formal notion. In principle it is capable of being given full expression in a charter of defined rights. Obviously, for instance, if there is a right to freedom and well-being then there must be a right not to be killed against one's will. This is so because without life there is (as far as can be empirically ascertained) no agency. Life is necessary for agency and therefore must be claimed as a right by all agents. Gewirth distinguishes between infringements of the right to freedom and infringements of the right to well-being:

[t]o interfere with someone's freedom is to interfere with his control of his behavior, including his participation in transactions; such interference hence affects the procedural aspect of the behavior. To interfere with someone's well-being, on the other hand, is to interfere with the objects or goods at which his

behavior is aimed; it hence affects the substantive or purposive aspect of his behavior. (Gewirth 1978, 251)

The difference is often aspectual. By this I mean a certain breach of duty can infringe against freedom and well-being depending how it is viewed. Thus, killing someone is an infringement of his right to freedom in that it is an action performed on him against his will, and an infringement of his right to well-being in that it prevents him from (ever again) acting for ends which he holds to be good.

3.3.1 Suggested outline for specific rights under the PGC

The following is no more than a suggestion of how rights might be typified in an actual state operating under governance of the PGC. The hierarchy of rights, I believe, is along the lines of freedom - basic well-being - nonsubtractive well-being - additive well-being. Further each of these has dispositional and occurrent aspects which enriches the hierarchy. It is necessary to specify actual rights in order for law to have its practical moment. However, it should be noted that to proceed towards a moral justification of punishment all that is necessary is that the reader accept that in principle there is at least one absolute human right. The subject of this right does not even have to be specified. If we accept that there is a single abstract right X then this is enough to allow us to consider the correct response when X is infringed. Having said that, I want my theory of punishment to be practically relevant so the following hierarchy of rights is a suggestion of the sorts of rights the infringement of which punishment is designed to respond to.

3.3.1.1 The right not to be physically harmed

An agent has the right not to be physically harmed against his will. At one extreme this involves the right not to be killed and along the spectrum includes the rights not to be physically or sexually assaulted. The reason for these rights is fairly obvious: each protects the freedom of the agent. For example, one who is

beaten up is denied the right to control that situation.¹⁴ Furthermore these rights protect the well-being of the agent. Without minimal physical security the agent is unable to set and pursue his purposes. Defining harm Gewirth states that

it is undeniable that the harmful includes whatever adversely affects the necessary preconditions of one's action, as by removing one's life, freedom of movement, physical integrity, or mental equilibrium; I have called these 'basic harms'. (1978, 232-233)

Gewirth's development of rights under the PGC is sophisticated and lengthy (see 1978, 199-271). As well as distinguishing between rights borne out of the generic rights to freedom and well-being he distinguishes between rights born out of the rights to basic, nonsubtractive and additive well-being. Physical assault is defined as an attack upon the agent's right to non-subtractive well-being as it infringes his generic-dispositional ability to defend the goods he already has. However, again an aspectual distinction arises as a physical assault is also an attack on a good which the agent necessarily needs for agency, *viz* his physical integrity. This is offered as no more than a flavour of how rich Gewirth's rights theory is. The point for our purposes is that whether we see assault as an attack on the agent's right to freedom, basic well-being or non-subtractive well-being, the agent does have a right not to be physically assaulted.

3.3.1.2 The right not to be physically coerced

The agent has a right not to be physically coerced against his will. At one extreme this involves the right not to be forced into slavery and along the spectrum includes rights not to be kidnapped or placed under arbitrary arrest.

3.3.1.3 The right to be rescued

The PGC differs from Kant's categorical imperative in several respects one of which is that it accords positive as well as negative rights to agents. Where a negative right states "you must not. . .", taking the form of a demand that the

¹⁴ This is so even if he is strong enough to ward off attack, because the agent does not freely choose to enter a situation in which he is forced to defend himself.

agent refrain from a certain action, a positive right states “you must. . .”, taking the form of a demand that the agent undertake a particular action. Thus, an agent in trouble has the right to be helped, which gives rise to a right to be rescued with its correlative duty to rescue those in peril (Gewirth 1978, 217-219). This right is hewn from the generic right to basic well-being; an agent who is left drowning does not attain the basic well-being necessary for pursuit of his particular purposes. As a matter of fact this right is not one which is generally recognised in the English legal system. However, any posited law that denies an agent the right to be rescued is in direct contravention of the PGC and cannot detract from the agent’s legal-moral duty to nevertheless rescue those in adversity. Having said this there are several restrictions that would apply to a law demanding action which do not generally apply to laws demanding restraint. It is an old axiom that *ought implies can*, so any law demanding that an agent rescue those in trouble must take account of the particular agent’s abilities. To take an extreme example, few agents would be under a legal-moral duty to rescue those on a space mission gone awry for the simple reason that few agents would be in a position to do anything about it. Gewirth offers four restrictive and inter-related criteria for any duty to rescue (1978, 230). By extension these criteria would have to be incorporated into any legislation. The first is the kind or degree of harm: the duty to rescue someone who is drowning is greater than the duty to prevent someone from twisting his ankle because the former disaster impinges much more greatly on the struggling agent’s well-being. While one with a sprained joint may be prevented from pursuing some of his ends for a short period while he recovers, one who has drowned ceases to be an agent. The second factor is the agent’s knowledge of the disaster (that is, the agent to whom a duty to rescue is attributed). This is derived directly from the maxim *ought implies can*. If B is drowning but A does not see him or hear his cries then he is not in a position to rescue B. Therefore, he can have no duty (either legal-moral or non-legal-moral) to rescue B. The third factor is the agent’s ability to avert the disaster and this too is derived directly from the maxim *ought implies can*. Even though B sees A drowning three yards away, if B cannot swim and if there is no rope to hand or no-one else to be called, there is nothing B can do about rescuing

A. The fourth factor is the cost to the agent of rescuing the disaster struck. If B is a very poor swimmer or he knows that there are hopelessly strong currents washing A out to sea, then he cannot undertake a rescue operation without endangering his own life. Given that he is as committed to his own rights to freedom and well-being as he is to the freedom and well-being of others he is not under a duty to rescue in this case. Also, the more trivial the impending disaster the less cost the agent will be morally and legally bound to incur. A strong swimmer may be morally and legally bound to ruin his new suit to rescue a drowning man, but he will not be morally or legally bound to ruin his new suit to rescue another's football from the water, even though that other has substantive property rights in the football. Given these four criteria any legislation positing the right to be rescued would have to be generously drawn up.

3.3.1.4 The right to property as a feature of freedom and both nonsubtractive and additive well-being

The right to property is a feature of the right to freedom in so far as the agent has the right to do with his possessions what he wants unforced by the will of others.

Rights to possess things arise because

having such rights serves to protect both the well-being and the freedom that are needed for purposive action and generally successful action. (Gewirth 1996, 171)

At the most basic level unless I have a right to possess the clothes I wear, the shelter that I live in, and the food that I eat, I will be unable to realise my agency. I will die of exposure or starvation. Even the threat of these harms prevents me from pursuing goals above the level of subsistence. Non-subtractive well-being is protected by the right to property in that, for example, my enjoyment of my art collection is guaranteed if others are prevented from stealing it. Additive well-being is protected by the right to property in that, for example, my ability to accumulate wealth and provide for my dotage is guaranteed if I have the right to the fruits of my own labour. It should be noted on this point that the right to property is not absolute because the positive duty to assist other agents discussed

above requires *inter alia* that agents pay certain taxes when asked (Gewirth 1996, 83). This is not controversial because what we are here developing is a hierarchy of rights in which some rights take precedence over others. Thus, I may not have to pay taxes so that another agent can develop his own art collection, but I may have to pay taxes so that his most basic well-being, including his need for shelter and food, is protected.

The right to property can be occurrent or dispositional. For example, I may have an occurrent property interest in an article such that no other may own or dispose of that thing against my will. The right to property is dispositional in the sense that an agent generally has the right to own property. This is the more important right because it is out of this that particular occurrent property interests are borne. Therefore, to deny a whole class of agents the right to property as in states where women do not have rights of ownership is a more serious wrong than to deny a particular agent his property right in a particular object. The right to property has several applications. At the basic level agents should have the right to ownership. Impaired or partial agency may lead to a corresponding diminishing of the right. Thus a child or someone with a mental handicap may not have the same rights to own goods as someone who is an agent in the fullest sense. Ownership is a right because it is generally necessary in order to pursue one's ends i.e. to be an agent. But those who are not able to pursue certain ends have less need of the right to ownership and are thus done no wrong if they are partially denied what would otherwise be a right.

From the general right to be an owner come a series of more specific property rights. The basic right may be seen as the right to property in a dispositional sense whereas the applications refer to occurrent property rights. The right not to have one's property stolen and the right not to have one's property damaged are both non-subtractive rights as is the right not to be defrauded of what one possesses.¹⁵

¹⁵ The right not to be deceived can also be viewed as a right to freedom. One does not act freely in the full sense if relevant information is deliberately withheld or wrong information deliberately supplied (Gewirth 1978, 250).

3.3.1.5 The right not to be physically endangered

An agent has the right not to be physically endangered over and above certain accepted levels. Of course, life itself is dangerous and the right is not a right to be protected from every potential harm. What I have in mind here are rights against being arbitrarily placed in greater danger than is necessary in every day situations. The delimitation of such a right is not an exact science so what is required is that the legislature act with *bona fides* to protect agents from danger. An example of such a right would be the legitimate expectation of each agent that others use the roads in a safe manner. There is no moral secret to setting the speed limit at seventy miles per hour but it is a genuine attempt by the government to reconcile the conflicting interests in effective transportation and physical safety of pedestrians and other road users. Because it is extremely dangerous to drive while drunk the government is allowed to pass laws against drink driving although there is no uniquely moral cut off point between acceptable blood-alcohol levels and unacceptable blood-alcohol levels. The right not to be endangered is also responsible for *inter alia* laws regulating dangerous activities like the use of guns and fireworks and the possession of dangerous animals, chemicals and machines.

This general right against physical endangerment protects each agent's non-subtractive goods. If the agent is fearful for his future, if he cannot, for example, walk on the pavement in the knowledge that drivers will observe at least some sort of rule (even if it is as basic as driving on a particular side of the road, or driving on the road as opposed to the pavement) then he is generally less able to fulfill his purposes. The need here is for laws, even ones with which the agent may disagree, which give each agent legitimate expectations of the range of possible outcomes when he goes about his daily business. To deprive him of such expectations is to lessen his agency, i.e. his ability to act voluntarily and purposefully. It is to be noted that although we are talking in terms of danger, which is generally contrasted with actual harm, the danger in this case is itself a

harm because it detracts from that which each agent necessarily holds to be good; his freedom and well-being.

3.3.1.6 Additive harms

Additive goods are those which enable an agent to further his agency i.e. his ability to act for and achieve his own ends. In certain cases infringement of such additive goods will be serious enough to constitute crime. Thus, a parent who willfully deprives his child of education is guilty of inflicting an additive harm because the less educated an agent is, the less he is able, generally speaking, to set and achieve his own goals. Other additive harms include racism, sexism and homophobia. To take the first example, an agent is less able to further his agency if he lives in a society that regards him as morally or intellectually inferior because of his skin colour. It is to be noted that racism is most usually legislated against when it is manifested in action which detracts from an agent's basic or nonsubtractive goods. Government, as an institution, may have further duties to educate people against irrational prejudices. With regard to other additive harms such as being addicted to drugs or pornography there is no obviously right answer as to when or how the law should intervene. As in the case of physical endangerment considered above we expect the state to act with good faith and competence in regulating the traffik in drugs and pornographic material.

The above is not supposed to be an exhaustive list of rights under the PGC nor is the grouping of the rights supposed to provide mutually exclusive types of rights: a particular right may well come under more than one heading. What the list is supposed to provide is an outline of the sorts of rights that are derived from the PGC. Further it offers some examples of the type of reasoning necessary to show that a specific right can be fashioned from the generic rights to freedom and well-being.

Given all the rights above it may be asked whether morality and law are in fact co-extensive: is every moral right to be written into the law books? In a sense the legal and the moral *are* co-extensive. As Beyleveld and Brownsword have it the

distinction between the legal-moral and the non-legal-moral (between law under the PGC and non-legal obligations under the PGC) is aspectual (1986, 165). The legal-moral is the set of PGC obligations regarded from the point of view of enforcement¹⁶ while the non-legal-moral is the same set of obligations regarded from some other point of view. There are restrictions on which duties can be viewed from the point of view of enforcement. These are logical restrictions again derived from ought implies can. Kant, in his formulation of law, states that a law is a moral duty which *can* be framed as external legislation (Kant 1991, 55). Law is impotent to make people adopt a moral point of view: although the PGC may require me to look on another agent's generic features of action as favourably as on my own, this is not something which can be enforced¹⁷ (see Kant 1991, 64; Findlay 1981, 324-325). All that can be enforced is that an agent's external actions, which are manifestations of attitude, are *as if* his attitude was one of favourable disposition towards the generic features of others' agency. There is an overlap with the additive harms such as racism discussed above. The argument to the PGC entails that an agent is guilty of irrationality if he sees people of different colour skin as having fewer rights than himself. But no law can demand that each agent actually view all other agents as having the same rights as himself: all it can demand is that he act as if this were the case (on this see further chapter 2).

3.4 Crime

A crime may be defined as the intentional breach of one of the above legal-moral rights. As I have not purported to give an exhaustive list of legal rights I cannot claim to offer an exhaustive list of crimes. However, as has already been noted, all that is necessary for a justification of punishment to follow is that it is accepted there is at least one legal-moral right, R. Similarly, it is only necessary that the reader accept there is at least one crime, C. It is no objection to the justification of punishment to hold that, for example, property is theft and therefore no wrong is done when X takes from Y something which Y claims to

¹⁶ See the formal definition of law given above.

¹⁷ Remember that the legal-moral is the set of moral duties viewed from the point of enforcement.

own. The only implication of such an objection (if shown to be true) is that punishment would be unjustified where one agent intentionally deprives another of an item. My thesis is founded on a moral base that suggests a hierarchy of legal-moral rights and by extension a hierarchy of crimes. However, even if I am wrong on this point the justification of punishment can be thought of in terms of the abstract right R (and by extension the abstract crime C). The argument is therefore sound unless it can be proved that there are *no* legal-moral rights (and by extension no such thing as crime). Furthermore, the argument will be of practical importance unless it can be proved that it is impossible to give any determinate content to the categories of legal-moral rights and crimes.

3.5 The legitimate State

Again, my thesis is a development of what others have written on this subject. Most notably Alan Gewirth himself in *Reason and Morality* and more so in his most recent work *The Community of Rights* deals with the issue of the legitimate state. Between the publication of these two works Deryck Beyleveld and Roger Brownsword wrote *Law as a Moral Judgment* which takes the suggestive material of Gewirth's earlier work and forms it into a coherent Gewirthian theory of law and state. In *Reason and Morality* Gewirth claims that

[t]he main justification of government is . . . to protect the most basic . . . rights; but this shows that the normative existence and implicit claiming of the rights are logically prior to the government that is appealed to for their protection. . . (Gewirth 1978, 75)

The force of this statement is twofold. First it demonstrates that government is in principle justified where it protects at least the most basic rights derived from the PGC. Second it shows, as Gewirth notes, that the PGC is epistemologically prior to the concept of the state. Hence, it is not necessary for there to be developed, or indeed any, institutions of government in order for an agent to claim his rights under the PGC.¹⁸ Government ensures that these rights are respected, it does not

¹⁸ This has the consequence that punishment may be justified in principle even if there is no state or no legitimate state, albeit that it may be practically impossible to carry out punishment where there are no state mechanisms. This is acceptable because I am primarily interested in a moral

institute them in the first place. This sentiment is reiterated in *The Community of Rights* where Gewirth says that

[t]he main justification of the state (and hence of political obligation) is, indeed, that it serves to secure persons' generic rights to both freedom and well-being. (1996, 60)

The ideal state then is one that fully guarantees the like rights of each of its subjects without itself infringing the rights of any individual subject. The legitimate state is one that sets this as its goal and operates with *bona fides* and a certain degree of competence towards this end. States that are openly against the PGC are illegitimate. This would include any state that is oppressive in the sense of denying rights to a section of its subjects. The implications of this vary. The actions that are morally justified against a state which systematically exterminates its subjects are greater than those against a state which, for example, sets a peculiarly high age of majority and thus could be argued to confine certain rights such as voting and choice with regard to issues like the consumption of alcohol, to an artificial section of the population.

The rest of this thesis will proceed on the basis that there is at least in theory a legitimate state. Punishment will be justified by reference to such a state.¹⁹ It is a point for further consideration as to what extent punishment is justified in a state which is either illegitimate or legitimate but some way from ideal.

4 Criminality

Having offered a definition of crime it might seem strictly superfluous to offer a definition of the criminal; definitionally a criminal is one who commits crime. However, it has long been recognised that there is a distinction between the outward actions which constitute crime - the *actus reus* - and the mental state necessary to be held as criminal - the *mens rea*. The purpose of this section is to

justification of punishment. The political applications of this morally justified theory of punishment are of course crucial but secondary to the current object of inquiry.

¹⁹ See also Philips (1986, 405) for the argument that the concept of the justified state is a necessary step in the justification of punishment.

consider briefly what the theories of Gewirth and Kant suggest to one wishing to construct a model of criminality. I take it to be axiomatic that a criminal is one who commits a crime.

As we have already defined crime the remaining "unknown" before we can define a criminal is what it means to *commit* something. In order for a specifically retributive theory of punishment to work it is necessary to demonstrate that the agent can be held responsible for the crime of which he is accused. This then raises the debate between free-will and determinism.

Both Gewirth and Kant (especially in the third section of the *Groundwork*; 1987, 79-97) develop moral theories closely linked to the concept of freedom. There are two important questions to be addressed here. First, what is the concept of freedom involved? Second, is this concept sufficient to ground a theory of responsibility for crime?

Kant, for one, appears to simultaneously rest his theory upon a concept of freedom and deny that freedom is knowable. This paradox can be resolved by asking the question whether his establishment of freedom is assertoric or dialectic. Kant's own pronouncements as to the noumenal nature of freedom would suggest that an assertoric proof of freedom is impossible and that any attempt in this direction will prove futile (see e.g. 1987, 78, 91-92). He is unequivocal in stating that reason would have to transcend its own bounds in order to explain the possibility of freedom. In this case, if Kant is to establish freedom at all it must be upon dialectical grounds. And indeed, this is exactly the nature of Kant's discussion of freedom in the *Groundwork*. He states that

every being that cannot act except *under the idea of freedom* is just for that reason in a practical point of view really free, that is to say, all laws which are inseparably connected with freedom have the same force for him as if his will had been shown to be free in itself by a proof theoretically conclusive. (1987, 79)

Continuing his thread in a footnote he adds that

this method of assuming freedom merely *as an idea* which rational beings suppose in their actions, [is adopted] in order to avoid the necessity of proving it in its theoretical aspect also. (1987, 79 fn.2)

This is a dialectical argument because while it does not prove that rational beings are free, it does go to show that any being which conceives of itself as an agent must believe itself to be free.²⁰ In Gewirthian terminology we might put it that a prospective agent, from his own internal viewpoint, must accept the proposition that he has free-will and must act as if this is the case. Indeed, given our definition of agency this is axiomatic. It is certainly self-refuting from the point of view of agency to deny that one has control over one's actions and it is probably *logically* self-refuting from any point of view to deny that one is free. This latter, logical, self-refutation arises from the fact that anyone making the claim "I am not free" must accept that he was determined to make the claim, which renders its truth value essentially irrelevant.

We have, in this definition of practical freedom, sufficient material to ground a theory of criminality. Specifically, every agent is committed, because he conceives of himself as being an agent, to the proposition that he has the dispositional ability to set and pursue his own ends. A criminal then is one who sets and *succeeds* in the end of crime. It might be the case that some Cartesian devil blinds him to the strings which in fact control his every action but if this is so, as it is unknowable, it is from a practical point of view immaterial. Both the observer and the actor are *prima facie* bound to view the action as occurring under the governance of the actor. It is then down to a consideration of the contingent factors such as disease, physical or psychological coercion etc., to see whether we can hold the actor *responsible* for the crime (see esp. Gewirth 1978, 31). In the absence of such factors we may take responsibility as being shown.

²⁰ This practical demonstration of freedom, as Kant says, explicitly avoids the need for a theoretical proof (1987, 79 fn. 2 quoted above). This seems a wise move given Kant's insistence that transcendental freedom is noumenal to us. This is notwithstanding that in the *Critique of Pure Reason* Kant asserts that practical freedom is meaningless unless we are actually, that is, transcendently, free (1929, 465; see further Beck 1960, 189-190).

This reveals the agent as criminal.²¹ Of course, in individual cases it may be hard to establish how responsible the particular suspect is for an offence but this does not reduce the theoretical force of the above argument.

5 Summary

The scope of this introductory chapter has been to lay out the reference points within which my investigation of the moral justification of punishment will be conducted. First, the object of my inquiry is a particular type of intervention generally conducted by or on behalf of the state. Second, this intervention is that which conceptually, if not empirically, is aimed at individuals who have deliberately committed an offence and the fact that the individual has deliberately committed an offence is taken to be the qualifying grounds for the intervention. Third, the concepts of law (hence crime) and state involved are not arbitrary but derive from application of the rationally justified principle of generic consistency.

The rest of this thesis will take three main steps. First, consideration of how two classic retributivists, Kant and Hegel, have dealt with the justification of state

²¹ Beyleveld and Wiles posit the dilemma (if not paradox) that only moral actions can be freely willed.

Suppose that an agent does B and cites A as the reason for that action. Can we not ask why A was the reason for B, or even what materially caused A to be held as the reason for B? Certainly if it is thought that A is not an adequate reason, that it manifests some degree of irrationality on the part of the agent, we might want to know what caused the agent to act irrationally and accept the explanation of this irrationality as the explanation of B. But suppose that the reason is adequate, that is, is effective in producing the action and justifies the action, is it then relevant to ask for an explanation of the reason?. (1979, 127)

Men have selves only to the extent that they are rational, and free action is rational action. (1979, 141)

The dilemma posited here is that action, to be freely-willed, must be rational (Paton contends that this paradox arises from the problem of explaining the origin of evil in the rational will; 1947, 275). As a corollary of this, action that is irrational cannot be freely willed and must therefore be materially caused or indeterminate. Notwithstanding the epistemological force of this point *we*, as agents who act under the idea of freedom, are bound to treat the criminal as freely willing his action. Similarly the criminal as an agent (i.e. from that point of view) is bound to consider himself as acting freely. To treat the criminal as determined is not just patronising but solipsistic as, for no good reason, it implies that the criminal is not an agent, while the patronising solipsist cannot but act under the idea that he himself *is* an agent.

intervention in the case of an agent deliberately committing an offence (chapters 2 and 3). Second, consideration of how modern thinkers have dealt with some of the shortcomings in the classical retributivism of Kant and Hegel (chapter 4). Third, a justification for punishment presented in terms of the PGC which deals with the shortcomings perceived in Kant and Hegel in a way which I contend modern thinkers have been unsuccessful (chapter 5).

Chapter Two: Kant on Punishment

Kant's place within the tradition outlined in the introduction to this thesis is that of forefather. In this chapter I will contend, against recent work by Jeffrie Murphy, that Kant is to be regarded as a retributivist. This contention relies on a proper understanding of those quotes adduced by Murphy which suggest that Kant's justification of punishment relies on the notion of deterrence. I will contend that as a retributivist the notion central to Kant's theory is that of freedom. For Kant, law is justified by reference to its ability to protect freedom. Punishment is justified by reference to its ability to hinder a hindrance of freedom. I shall consider both of these claims in detail. I will consider Kant's attempts to derive a justification of punishment from the concepts of law (and its opposite crime) and criminality in order to ground the claim that we punish because an offence has been committed. Of course there are certain famous dicta about the meting out of punishment in Kant's work and these will bear close inspection. However, it would be wrong to ignore the context of these passages; many are perhaps no more than polemical or heuristic and secondary to a wider point. Further, it would be wrong to assume that statements sparsely scattered through several works written over a long period of time form (or are meant to form) a coherent theory. The chapter will comprise of three main sections. First, and in order to sort out the confusion raised by some of Kant's pronouncements on punishment, I will consider Jeffrie Murphy's recent querying of Kant's theory of punishment. Having contended that several problematic dicta are reconcilable with a theory of retributive punishment I will go onto the second section. In this I will consider Kant's theory of retributive punishment as contained in the *Metaphysics of Morals*. Third, I will consider what I refer to as "the Shipwreck example" which is found within the *Metaphysics of Morals* and suggests that Kant has a deterrent theory of punishment.

1 Does Kant have a theory of Punishment?

Problems of interpreting Kant's multiple statements on punishment have led one leading Kantian scholar to ask the question at the head of this section (Murphy 1992). In his article Jeffrie Murphy complains

[n]ot only am I no longer confident that the theory is generally correct; I am also not at all sure that I understand (or find understandable) much of what Kant says on crime and punishment. (1992, 31)

Central to his thesis is the contention that all we have in Kant's work is a mixed bag of remarks on punishment which were never developed into, and were perhaps never intended to form, a consistent theory on the subject. This finding is based partly on Murphy's understanding of what it is to have a theory of punishment. At the very least, he states, a theory of punishment must answer the following five questions: What is the nature of crime and punishment? What is the moral justification of punishment? What is the political justification of punishment? What are the proper principles of criminal liability? What are the appropriate punishments? (1992, 32-34). It should be noted that while this list seems a fair requirement of a theory of punishment it is a set of Murphy's own stipulative preconditions. It is broader than the questions I have set myself in order to justify retributive punishment and I will focus on those aspects of Murphy's construction of Kant which deal with the moral justifications of punishment. Murphy's list allows him to draw the conclusion that anyone who fails to answer one of the questions does not have a theory of punishment. Taking the fifth question as an example, we might have a theory of the justification of Punishment (capital P) without completely addressing the question of which punishments (small p) are justified. It is possible to move from a moral theory to a conclusion that Punishment *is* justified and posit this as "a theory of the justification of Punishment" without necessarily going on to settle the issue of which punishments are justified; "a theory of justified punishments". Having made this point I will now consider Murphy's own reasoning as to why Kant does not have a consistent theory of punishment. My emphasis will be on

whichever contentions of Murphy's cast doubt on whether Kant has an answer to the second question: what is the moral justification of punishment?

Murphy's strategy is to reconstruct a theory of punishment from all of Kant's writings on the subject (as well as on the complementary subjects of moral responsibility, state organisation, and rational judgment) *except* for what is found in the *Doctrine of Right* (first section of the *Metaphysics of Morals*) (1992, 36-39). This leads him to reconstruct the following Kantian theory: (all italics direct quotes from the original)

(1) *justified punishment is a deterrence system*

(2) *functioning to maintain a system of ordered liberty of action.*

(3) *To set any more morally ambitious goal for punishment would be to adopt an unacceptable theory of the role of the state and would represent an attempt to play God, revealing a lack of proper insight into our own shortcomings, a lack or appreciation of the role of luck in our own achievements or virtue, and a lack of the posture of humility appropriate to creatures of our sort.*

(4) *Punishment is a necessary evil, but we should inflict and support it with regret and without any sense of having embarked on a righteous moral crusade.*

(5) *Because of the we should be constantly on the watch for factors that excuse or mitigate criminal conduct so that we will not be any harsher than absolutely necessary to accomplish our legitimate state goals - a crime control system constrained and limited by a sense of justice and Christian compassion and humility. (1992, 41-42)*

Having identified this theory of punishment in Kant's work other than the *Metaphysics of Morals*, Murphy contrasts it to the writings on punishment found in that discourse, and concludes that the two are irreconcilable. The two claims which seem to me to be of particular interest are (1) and (2). The remaining three claims have nothing to say about the reasons for, or justifications of, punishment. Rather, they seem to require that once we have justified punishment, we go about it in a particular (humble, compassionate etc.) way.

The claim that "justified punishment is a deterrence system" can be read in several different ways.

- (1.1) As a statement of fact justified punishment *does* act as a deterrent.
- (1.2) Ideally, justified punishment *would* act as a deterrent.
- (1.3) Deterrence is necessary for the justification of punishment.
- (1.4) Deterrence is sufficient for the justification of punishment.
- (1.5) Deterrence is necessary and sufficient for the justification of punishment.

Of these two can be dismissed immediately. (1.4) must be wrong, because Kant expressly denies the utilitarian logic behind the claim. In the *Groundwork to the Metaphysics of Morals*, in other words outside of the *Doctrine of Right* and therefore within the body of works which Murphy is considering in his reconstruction, Kant states that it is illegitimate to use an individual as a mere means to an end (Kant 1987, 58). This would specifically disallow using a particular individual as the subject of punishment in order to deter others from crime. Without other considerations it would also specifically disallow using a particular individual as the subject of punishment in order to deter that individual from crime. The conditional clause “without other considerations” still allows that deterrence could figure in the justification of punishment but does not allow that deterrence could be *sufficient* for the justification of punishment. If deterrence were sufficient for the justification of punishment then we would face the dread prospect of incarcerating the innocent to reduce the crime rate and nothing in Kant’s work suggests that this is a morally legitimate thing to do.²² Because we must reject (1.4), we must also reject (1.5). Obviously, if something cannot be a sufficient justification for punishment, it cannot be a necessary and sufficient justification for punishment. Of the remaining claims only (1.3) addresses the justification of punishment. The other two assume that punishment is justified *before* they go on to consider the relationship between punishment and deterrence. The first claim (1.1) is not of interest to us. The statement that as

²² It is important to remember Hart’s counsel against the definitional stop. It is not enough to say “but, if so and so is innocent then we are not *punishing* him at all”. The point is not what we call it, but whether it is justified, and if I claim that I am justified in doing a *prima facie* wrong to someone *simply* because this reduces crime, then this entails that the guilt or innocence of that person is irrelevant to the justification of my action.

a matter of fact justified punishment *does* deter crime is essentially irrelevant to the justification of punishment (ignoring the spurious factual basis of such a statement). We merely posit a thing (justified punishment) and observe that it has a particular effect. What we are interested in is the thing itself and it is impossible to make any inferences about the justification of punishment from observances about its effect. Statement (1.2) merely holds that deterrence is a desirable aim of punishment. This is true because, being wrong, it is always good if crime is avoided. The statement is also uncontentious because it says nothing as to the justification of punishment: it posits deterrence as a desirable consequence of *justified* punishment (see Kant 1991, 141 quoted *supra*).

Murphy adduces fourteen quotes in support of his construction of Kant (Murphy 1992, 36-39). Of these nine deal with the difficulties of conceptualising free-will and the consequent problem of telling whether someone is morally iniquitous. It should be noted that the force of these quotes is to modify the administration rather than the justification of punishment. They demand that, even though punishment is justified, we are cautious in its application because the motivations behind crime are complex and never entirely deducible. This is the purpose for which Murphy uses the quotes. It is to be noted though that the majority of the evidence offered by Murphy makes no mention of the justification of punishment.

Of the remaining five quotations one seems to deal with neither the justification nor the application of punishment:

The more legislation and government are brought into harmony with the [ideals of justice], the rarer would punishments become, and it is therefore quite rational to maintain, as Plato does, that in a perfect state no punishments whatsoever would be required. (From *The Critique of Pure Reason* cited Murphy 1992, 37; Murphy's modification)

This says nothing about punishment or its justifications. Rather it suggests that ideally *law* and *government* act so as to deter crime. This sort of deterrence is uncontentious because crime is a bad thing to be avoided where possible. The

question is whether punishment can be used to this end and because purely deterrent punishment would use an individual as a means to an end the answer, on Kant's own criteria for moral theory, must be no.

It might also be noted that Kant seemed to have a very optimistic opinion of how easy it is to deter people from crime. For example, in the *Groundwork* he asserts that

[t]here is no one, not even the most consummate villain, provided only that he is otherwise accustomed to the use of reason, who, when we set before him examples of honesty of purpose, of steadfastness in following good maxims, of sympathy and general benevolence (even combined with great sacrifices of advantage and comfort), does not wish that he might also possess these qualities. Only on account of his inclinations and impulses he cannot attain this in himself, but at the same time he wishes to be free from such inclinations which are burdensome to himself. (1987, 87)

Whether or not this is accurate psychology it does suggest why Kant believed that in an ideal state, with perfectly framed and promulgated laws as examples, crime could be eradicated. In such a state punishment would not have the function of deterrence: there would be no punishment. Thus the question of justifying punishment would be rendered academic and hypothetical at best.

Murphy also cites an example from the essay *Perpetual Peace* where Kant states that even a society of devils can be organised under law by appeal to their rational self interest in remaining unmolested by each other. The passage concludes that this problem

does not require the moral improvement of man; it requires only that we know how to apply the mechanism of nature to men so as to organize the conflict of hostile attitudes present in people in such a way that they must compel one another to submit to coercive laws and thus enter into a state of peace, where laws have power. (cited Murphy 1992, 39)

The conclusion Murphy draws from this is that while retribution might be a "laudable moral goal, it is not clear that it is the proper business of the state to aim at this goal" (1992, 40). There are two different issues here. First Kant is

quite clear that law does not attempt to effect the morality of a particular agent. Indeed this follows not only from the works which Murphy cites but from the division of the *Metaphysics of Morals* itself into the *Doctrine of Right* and the *Doctrine of Virtue*. Kant's discussion of right, law and punishment is contained in the first part of the work where he is open that "[n]o external lawgiving can bring about someone's setting an end for himself" (1991, 64). But secondly, to say that law and punishment do not aim at producing a moral agent is not to say that retribution which is morally laudable is unjustified. Law is morally laudable (because it ensures that the (legal-)moral rights of agents are observed) but nevertheless functions by constraining the actions of agents rather than making them better agents. Likewise punishment may be morally laudable (because it brings about a right state of affairs) without having to affect the morality of the individual being punished. Law and/or punishment are morally laudable just because they set for themselves a moral end (conformity with right) and the ideal legislator (some descendent of Plato's philosopher king) has this as his end also. It is possible for this to be the end of law/punishment without law/punishment making it the end of individual addressees. Law achieves its end when individuals refrain from trampling the rights of others even if those individuals remain nasty beings who sincerely wish they could harm others. The fact that punishment may not affect the morality of individual agents does not preclude its being morally justified. It remains to be seen whether it *is* morally justified but this has not been disproved on the strength of the evidence adduced so far.

The quotes which do suggest that justified punishment is deterrent are as follows:

It is . . . a duty of virtue not only to refrain from repaying another's enmity with hatred out of mere revenge but also never even to call upon the world-judge for vengeance [. . .] this must not be confused with *placid toleration* of injuries [. . .], renunciation of the rigorous means [. . .] for preventing the recurrence of injuries by other men . . . (From *The Doctrine of Virtue* cited Murphy 1992, 39; ellipsis in brackets my own)²³

Punishments are, therefore, a means of preventing an evil or of punishing it. Those imposed by governments are always deterrent.

²³ NB The *Doctrine of Virtue* was designed to be complement the *Doctrine of Right* so it is highly unlikely that Kant himself considered the sentiments expressed in this quote to contradict what is written in the *Doctrine of Right*.

They are meant to deter the sinner himself or others by making an example of him. (From *Lectures on Ethics* cited Murphy 1992, 36)

In the first case Kant states that we should not tolerate crime because this would lead to giving up the means of preventing further crimes. In the second Kant asserts that governments always aim at deterrence when they punish. Certainly this would appear inconsistent with Kant's assertions in the *Doctrine of Right* that the criminal

must previously have been found *punishable* before any thought can be given to drawing from his punishment something of use for himself or his fellow citizens. (1991, 141)

However I believe that these statements can be reconciled and that Kant achieves this himself in another quotation given by Murphy. In this Kant, writing to J.B. Erhard, tells him:

you are right in saying that the *poena meremoralis* ["ethical penalty"] (which perhaps came to be called *vindicatiua* ["avenging punishment"]) for the reason that it preserves the divine justice), even if its goal is merely medicinal for the criminal and the setting of an example for others, is indeed a *symbol* of something deserving punishment, as far as the condition of its authorization is concerned. (From *Letter to J.B. Erhard* cited Murphy 1992, 37-38; translations given in Murphy)

This is based in a tradition traced through Plato to Kant of distinguishing between a theoretical, ideal world and the real, less-than-ideal world (see e.g. Plato 1974, 260-263). In Kant the ideal is termed, whether through genuine faith or political expediency, as the *divine*. I believe that Kant's overall position is this. Punishment is ideally retributive (a point which Murphy seems to concede (1992, 41)). In the real world governments impose punishment for many and varied reasons (prevention, deterrence, rehabilitation) all of which are good insofar as they tend to reduce the amount of crime. But what *justifies* punishment, that which allows us to enforce these goals without the consent of the criminal and yet without treating him as a mere means to an end, is that punishment is retributive: it reflects the type of punishment which God would impose under divine justice. Thus Kant says that even if the goal of punishment

is the deterrence of others, which on the face of it is utilitarian and allows of scapegoating, punishment is only *justified* because it is a “symbol” of a situation in which punishment is otherwise deserved. Kant is explicit that this symbolism is “the condition of its authorization”. By analogy, what justifies the law is that it reflects virtue albeit that it may achieve this without making anyone set virtue as his end. What justifies real punishment is that it reflects ideal punishment albeit that the actual goal which a particular punisher may have in mind is non-ideal (cf Sullivan 1989, 243-244 who makes a similar point). This reading of Kant ties in well with those passages cited by Murphy where Kant states that the actual motivation for an action is unknowable and thus we cannot judge the iniquity or otherwise of a particular criminal. We are not perfect and thus our punishments will never be perfect but the fact that we have a concept of ideal punishment means that we should strive to conform to that ideal. We achieve this by judging someone’s moral iniquity on the basis of his external actions which, without some excuse, justification or mitigation, may be taken to manifest his moral worth (see especially Kant 1929, 474, 477). Recognising this imperfection, as Murphy contends in his reconstruction of Kant, we should be cautious in the imposition of punishment. To take an example from an actual penal system it might never be known with absolute certainty (in a way that an omnipotent God might know something) that X did Y. We can nevertheless demand that it is “known” *beyond reasonable doubt* that X did Y before we punish X. Kant’s contention that punishment is justified because it is retributive can actually be squared with Murphy’s reconstructed Kantian claim that “justified punishment is a deterrence system”. In the real world punishment is used as a deterrent but it is justified because ideally it is retributive. This is the difference between the claims that “*justified punishment is a deterrence system*” and “*punishment is justified because it is a deterrence system*”. Even more importantly this interpretation of Kant can be reconciled with his treatment of punishment in the *Doctrine of Right*. Remember that Murphy’s strategy was to reconstruct a Kantian theory of punishment from the works outside of the *Doctrine of Right* and then contrast this with the pronouncements found therein. This is supposed to lead to inconsistency which I do not believe exists. As we have seen above,

Kant concedes in that work that deterrence is a justified goal of punishment but not what justifies punishment. He also asserts that punishment

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*. (1991, 140)

Because the *Doctrine of Right* is the one point where Kant deals exclusively and in detail with the problem of punishment we must attach proportionately more weight to the findings of that work. I will later consider one example (which Murphy does not employ himself) contained within the *Doctrine of Right* which itself suggests a deterrent justification for punishment and reconcile this with the more orthodox retributive sentiments expressed in that work. It should be remembered that there were two elements to Murphy's reconstruction of Kant which dealt with the justification of punishment. The first of these I have dealt with above. The second goes on to say that justified punishment functions "to maintain a system of ordered liberty of action" (Murphy 1992, 41). This I take to be the equilibrium which Kant believes punishment restores and I will focus on this in considering Kant's justification of punishment in the next section. It is now time to turn to the *Doctrine of Right* and Kant's more familiar statements on retributive punishment.

2 Kant as a Retributivist

The orthodox interpretation of Kant is that he is a paradigmatic retributivist. As a reason for punishment retributivism entails that we punish the criminal *because* he has committed a crime. Pronouncements to this effect are to be found within Kant's work (e.g. 1991, 140). Retributivism is also commonly linked with theories of "just deserts", that is theories which dictate how the punishment should be made to fit the crime. These usually involve some notion of proportionality. Kant's name is most often linked with the *lex talionis*, which requires that the punishment be exactly equal to the offence (Kant 1991, 141). There are then, two strands to retributivism, one of which states *why* we punish,

the other of which states *how much* we punish.²⁴ Kant is generally held to be a retributivist in both senses (but see Scheid 1983, 262; and generally Byrd, 1989).

It seems to me that the *if* the claim that *we punish because the individual has committed a crime* is established then the claim that *we should punish in proportion to that crime* has some force. In other words, given the first strand of retributivism, the second strand has an inherent logic. This is not to say it has a water-tight logic. All that I am contending at this stage is that if the reason for which we punish is solely that the individual has committed a crime then it would be strange if we looked to factors entirely extraneous to the crime in setting the amount of the punishment. However, *without* further presuppositions (such as “crime is a bad thing”, “when somebody does a bad thing they ought to have bad things done to them”) there is no inherent logic in the claim that we should punish someone because he has committed a crime. It is certainly not analytic. If, as Hawkins contends, “[f]or Kant it is an evident moral principle requiring no justification outside itself that crime deserves punishment” (1971, 15) then Kant’s retributivism is an arbitrary reason for punishment. It is no more logically compelling than saying we punish people because they have blue eyes, albeit that it may be more intuitively compelling. In an attempt to reconstruct and understand Kant we must look for any presuppositions upon which the claim that we punish because of an offence may be based.

Both strands of retributivism must be considered in order that punishment has any practical relevance: it is hollow to show that punishment is justified in theory if we do not have any reference by which to measure the justification of individual instances of punishment. For the moment though the link we must forge is between the notion of equilibrium and punishment’s general ability to restore it. This can be achieved by consideration of the first strand of retributivism without yet delineating the precise amount of punishment necessary in any given case. My emphasis in considering Kant is therefore primarily upon

²⁴ It will be noted that Murphy’s reconstruction of Kant considered above deals largely with the second strand of retributivism. Most of the quotes adduced by Murphy deal with how we should go about punishing once (or perhaps *if*) we conclude that punishment is ever justified.

the first strand of retributivism because this covers the questions of what the equilibrium is, why it is to be restored and how punishment could be said to achieve this in general.

2.1 The equilibrium to be maintained by punishment: Kant's notion of right as an external lawgiving protecting freedom

In chapter one we encountered Kant's claim that the distinction between the legal-moral and the non-legal-moral lies in the possibility of framing the legal as entirely external legislation. That is to say the action required must be entirely external and the motive for the action must at least be potentially external.²⁵ "Right is the totality of laws for which an external lawgiving is possible and positive Right is the totality of laws for which an external law *has* been given" (1991, 55). One must be careful when dealing with the wording of this formulation. First, when Kant states that Right is the totality of *laws* for which an external lawgiving can be given, the word "laws" is in fact synonymous with *moral duties* rather than legal duties. If this were not the case the definitions of Right and positive Right would be circular (i.e. we would be defining our legal duties in terms of our legal duties). Second, the fact that Right is the *sum of* (pre-existent) *laws* (moral duties) for which an external law-giving is possible emphasises that for Kant the legal is not a distinct category from the Moral. Without the inclusion (and proper understanding) of the italicized clause it might seem that Right could include a law such as "All Scots must be killed" simply because this *can* be formulated as an external law. However, if such a "law" were drawn up it would never be included in the doctrine of Right, because the killing of all Scots does not correspond with any moral duty (i.e. *law* in the sense in which Kant employs it in this context). In other words, before we can consider whether something can be framed as an external law we must consider whether it is a moral duty. If it is not, then no matter how we frame it, it can never be called a legal duty within the doctrine of Right. Kant's legal theory is a distinctly

²⁵ I may choose to drive on the left because I am scared of punishment (an external motive) or because I am a genuinely virtuous person (an internal motive). I can only choose to be charitable because I am a genuinely virtuous person. For Kant, the law does not *have* to be obeyed because of an external motive but it must be possible to obey it because of an external motive.

natural law theory based on the premise that a law (in the orthodox sense) is a moral duty for which an external lawgiving is possible and a positive law is a moral duty for which an external law has been given.²⁶ This then shows us why, for Kant, right is important and to be maintained. The category of Right is coextensive with a class of rationally justifiable and imperative moral duties. As it is the case that these duties *ought* to be fulfilled it must also be the case that their concomitant rights ought to be upheld. The equilibrium exists as between rational beings with a will who have duties towards each other and hold rights against each other. The equilibrium is breached where a particular being fails to fulfill his duty towards another and hence breaches that other's right.

The possibility of an external law-giving for a moral duty is what distinguishes a legal-Moral duty from a non-legal-Moral duty. What then does it mean to say that an external lawgiving is possible for a certain moral duty? Kant is not referring to the potential for *articulating* a particular moral duty. It must be possible to express any moral duty, otherwise it would not be a moral duty for the simple reason that it would be noumenal. Nor, is Kant interested directly with external *actions*: while, it is true to say that the juridical law *is* concerned with external actions this would not be a manner in which we could necessarily distinguish it from ethics. Ethics will also dictate certain external actions albeit that what defines the rule as ethical is the motive from which the action is carried out rather than the action itself (see Gregor 1963, 23 and 27). For example, the ethical rule that I give to charity is ethical because it requires that I do so out of a sense of duty. Nevertheless, it does require that on at least some occasions I perform the external action of giving money to charity. Rather, what is at stake in external lawgiving, is the notion that another, i.e. a legislator, can *compel* me to a certain action. The lawgiving is external because both the action and the motive are external to the subject. In the *Metaphysics of Morals* Kant states that ethical rules cannot be the subject of external lawgiving

²⁶ This premise is the metaphysical definition of law. Together with the metaphysic of morals, which allows us to determine *a priori* the nature of moral duties, we have an *a priori* criterion for judging the lawfulness of a positive law (Gregor 1963, 34).

simply because they have to do with an end which (or the having of which) is also duty. No external lawgiving can bring about someone's setting an end for himself (because this is an internal act of the mind), although it may prescribe external actions that lead to an end without the subject making it his end. (1991, 64)

The distinction between ethics and law is a subtle one because they are not distinct types of duties but only distinct as to the types of reasons for which the duties are carried out. The distinction lies in the fact that ethics is the sum of those duties for which the only incentive *offered* is the fact that it is my duty. Positive Right then, or Positive juridical law as Kant sometimes calls it, is concomitant with the sum of those duties for which an incentive other than the fact that it is my duty is offered. It must be emphasised that in such a case the duty remains my duty whether or not such an external incentive to conform is offered. As Kant claims, all duties *can* be typified as ethical duties; i.e. I ought to perform all my duties and I *need* no more reason or incentive for this than the fact that they are my duties.²⁷ This is what it means to say I ought to do something within a deontological moral theory and can also be seen from the fact that all duties are derived from the supreme moral principle and are hence *categorical*. It is the morally consistent possibility of offering an external incentive to obey a moral duty which narrows it to a legal-moral duty and distinguishes it from a non-legal-moral duty. As Gregor puts it "the moral necessity of the actions is independent of our inner attitude of will" (1963, 28). By this she means that certain actions or omissions are required in themselves whereas others are required only insofar as they are manifestations of an inner attitude of will. Refraining from murder fits into the first category while giving to charity fits into the second. A law-giver cannot prescribe the reason ("inner attitudes of will") for which a person must act but can only prescribe the *way* in which he must act. A law which stated that all citizens must drive on the left hand side *because it makes them happy* or *because this is the law* would be unenforceable because we would never have any way of telling whether those driving on the left were doing so for such reasons. Furthermore, it would be

²⁷ Ethics in this sense is a synonym for Morality. This is because I owe all my duties for no greater reason than the fact that they are my duties. However, ethics in the narrower sense which Kant uses it excludes those duties for which additional incentives are offered.

pointless (within the scope of juridical law) so to prescribe because all that is necessary for the protection of the freedom (and well-being) of all in accordance with a universal law is that people actually *do* all drive on the left not that they do so for any particular reason. Juridical law does not prescribe a reason for which people *must* act but rather offers (dis)incentives for them to act in the prescribed way. Kant says:

[s]trict Right rests . . . on the principle of its being possible to use external constraint that can coexist with the freedom of everyone in accordance with universal laws. (1991, 57)

If certain freedoms are rights then it must be the case that we can act to ensure that those rights are upheld in any relevant situation. Hence, Kant talks in terms of constraints serving to protect rights to freedom. The example that Kant gives us of this “external constraint” is that of a creditor taking his debtor to court so as to have the debt enforced. It should be noted that this is not yet punishment either as reported in Hart’s core definition or as in my own starting point of a category of intervention specific to an offender who has committed an offence. Enforcement of law is wider than these notions of punishment. For one thing, to force me to fulfill an obligation which I have is not punishment. Punishment as we commonly conceptualise it, operates where it is too late for me to fulfill a duty which I have previously ignored. This notion of law enforcement would include such mechanisms as deterrence, prevention, self-defence and restitution all of which act as constraints on the will of an offender (or non-offender or potential offender) to produce the desired consequence of conformity with the law.

2.1.1 Deterrence as a constraint on the will

A person is deterred from an action where a sign of some sort operates on his mind so as to cause him to decide against that action. An example of this sign might be a threat as to what will happen should the individual perform the action. Each sign is an external factor (i.e. one which the agent is forced to take into account, whether or not it is decisive) operating internally so that the

individual tends to make the (constrained) choice to refrain from the action. Provided only that the constraint is justified there is no problem with deterrence in this sense. The constraint is justified for Kant where the action which the individual would otherwise perform is a wrong. Given that nothing is actually done to the individual in this case of deterrence there is little problem with how he is actually deterred. It does me no wrong if I am deterred from stealing by a threat that I will be burned at the stake if I do so, albeit that I would be wronged if anyone actually burned me at the stake for an act of theft. The reason why it does me no wrong is that I have not had my freedom impinged. I am not free to steal in the first place.

Of course, as has long been recognised by punishment theorists, problems do arise where the harsh treatment of *others* or the torture of myself in advance are used to dissuade me from future actions. In the case of threat my internal freedom is not actually hindered (I am not free in this sense to steal) and my external freedom is only hindered insofar as it would be an expression of wrong (i.e. the external act of stealing is wrong). But in these cases of public flogging or deterrent torture external freedoms are actually hindered in a sense which is itself *prima facie* wrong because they are contrary to Kant's injunction against treating someone as a mere means. Therefore, these examples could not form the basis of justified constraints of an individual's will. The fact that the equilibrium of rights to freedom is justified does not justify its maintenance at any cost. Specifically it cannot be maintained where the action itself leads to a greater abuse of right than the initial breach or threatened breach. It will also be noted that while deterrence as viewed here tends to maintain the equilibrium of right by preventing an initial breach, retributive punishment purports to restore the equilibrium once the breach has occurred. Although this is a justifiable goal - given that the equilibrium of rights to freedom is important - it is not yet clear how punishment achieves it. Effective deterrence on the other hand clearly serves to maintain equilibrium by preventing disequilibrium in the first place.

2.1.2 Prevention as a constraint on the will

A person is prevented from an action where external circumstance contrives to constrain him from performing the action. Examples of this would be where I stop someone from mugging me by employing bodyguards or stop him from burgling my house by putting a one hundred foot fence around my property. Here the individual is not deterred from doing wrong, but rather he is physically prevented from doing it. Deterrence operates internally in the sense that it is always open to the individual to commit the crime and then face the consequences if he should choose. Prevention operates externally in the sense that (where it is effective) it makes no difference what the individual chooses, he will be disabled from committing the crime. Obviously there is a degree of overlap between deterrence and prevention so defined. For example, while a one hundred foot fence might prevent me from burgling your house, a seven foot fence, which I *could* surmount, but only at considerable effort, might deter me from the attempt. However, a hard distinction between the two is unnecessary for the purposes of the present argument. What is important is that some forms of deterrence and prevention will require no justification because no-one is done any wrong. Thus, just as the threat of reaction in the event of crime does the contemplative criminal no wrong, neither does the placing of an insurmountable object in his way do him any wrong. Again, his freedom is not limited because he is not free to commit crime. And again as with deterrence, the tendency is to maintain the equilibrium of rights to freedom by averting a contemplated breach.

Excessive actions taken to prevent a crime are themselves wrong because they make a more basic interest in freedom and well-being give way to a less important interest in freedom and well-being. As in the case of deterrent torture, simply because some deterrents/preventions are justified it does not follow that there is *carte blanche* to deter or prevent crime by any available method. Neither deterrence nor prevention are ends in themselves. Rather, for Kant they serve the end of protecting the freedom of each compatible with the freedom of all and are therefore unjustified when they infringe this end.

2.1.3 Self-defence as a constraint on the will

Self-defence is a particular occurrent example of crime prevention and as such has the same justification. It should be added though that when a crime is actually taking place rather more is justified than would normally be the case with prevention as a class. This is because the problems of predicting that crime is going to occur are obviated in the particular case.

2.1.4 Enforcement of the law as a constraint on the will

We have now considered the instances of deterrence/prevention which Kant may have in mind when he states that Right is inextricably linked to an authorisation to use coercion. He also, explicitly, uses the example of law enforcement in connection with coercion. It would seem uncontroversial that where there is an independently justified law there should be a correlative right to see that the law is followed. If a law requires that my external actions conform to the imperative "whoever uses the road must drive on the left" then it follows from the very fact that this is an imperative, that there should be a right to enforce it. This conclusion is reached by consideration of two statements. First, something cannot be simultaneously imperative and optional. Second, I am not free to disobey an imperative, and therefore I am done no wrong where I am coerced into obedience. Again the same proviso that was attached to the instances of deterrence and prevention must be given to this second statement. Namely that a coercion which is more than the mere enforcement of an imperative will require independent justification. Thus, if I owe someone five pounds I can be forced by law to pay this debt but, *prima facie*, I cannot be forced to pay one hundred pounds and I cannot be forced to comply by the selective shooting of my family. But the problem for a Kantian theory of punishment is precisely this; punishment does go over and above what is required for the purposes of deterrence, prevention or enforcement thus conceptualised. It does go beyond what is required to stop a breach of the equilibrium. I cannot use deterrence as a justification for punishment because punishment *prima facie* constitutes a wrong

(i.e. an infringement of my freedom). To deter by threat only infringes my non-existent "freedom" to commit crime; to deter by torturing me so that I will think twice before committing an offence infringes my very real freedom to bodily integrity. To prevent me offending by putting obstacles in my way again infringes the non-existent "freedom" to commit crime; but to prevent me by cutting my legs off, or incarcerating me, infringes my existent freedom to walk about so long as I am not infringing the freedom of others. To attempt to justify punishment by reference to prevention would, at the very least, involve the controversial claim that we *know* that an individual will commit a crime unless locked up. And where this claim can be made, for example in the cases of the mentally disordered or someone bearing arms in an apoplectic rage, the actions which we may take to prevent wrongs are by way of treatment or a brief preventative custody in the name of self-defence rather than punishment as normally understood. It certainly appears irrelevant in such cases that the individual is an offender or is actually committing an offence. The justification for prevention is by reference to defence of a right vested in another and this type of intervention is justified against an innocent automaton as well as a dangerous criminal. Both cases are justified on the basis of a contemplated and practically certain wrong and independently of the actual commission of a crime which would seem to be a conceptually inherent aspect of punishment. If punishment is to be justified on the basis that an individual has committed a crime - and certainly this is how Kant himself sees it - then prevention is irrelevant to the issue of whether the punishment is *prima facie* justified. To attempt to justify punishment by reference to enforcement would again be misleading. The whole point of enforcement is that it does me no wrong because I have a duty to do that which is being enforced anyway. I have a duty to pay my debts, I have a duty to keep off of someone else's private property etc. But I have no *prima facie* duty to be locked in a cell or to let myself be punitively fined unless such things are otherwise justified - and this is precisely what is at stake.

The problem so far is that while certain practices and mechanisms can clearly be seen to maintain an equilibrium of freedom between individual agents it is not

immediately clear how punishment could function to this end. There is one feature of punishment which renders it problematic in this context. Punishment operates after disequilibrium has been introduced. It is, as it were, a cure rather than a prevention., and it is not apparent how it operates as such. While it is easy to see why making a criminal give back some stolen goods restores an equilibrium between the criminal and the victim, it is not so easy to see what, if anything, is restored by fining, incarcerating, torturing or censuring the criminal.

2.2 How does punishment restore the equilibrium: punishment as a hindrance to a hindrance of freedom

In the *Metaphysics of Morals* Kant states that

[r]esistance 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.* (1991, 57; emphasis changed from original)

This suggests a conceptual link between the notions of right and punishment.

Both enforcement, as discussed above, and retributive punishment aim at the restoration of a pre-existent equilibrium of right. However, enforcement need aim at no more than the restoration of the material equilibrium, restoring to the victim the things of which he was wrongly deprived. Because enforcement does no more than restore the rightful position it is unproblematic. If a certain situation is rightful then there must be a right that that situation be maintained. In other words the right that a thief be deprived of his ill gotten gains stems entirely from the right of the true owner to possess them. This is a logical connection because to say that I have a right to possess a thing entails that no other has a right to possess it. However, retributive punishment is (*prima facie* at least) over and above what is required to restore the rightful situation, or seemingly

irrelevant to restoring the rightful situation. In the former case, retributive punishment is over and above that which is required for the purposes of restitution because not only do we make the criminal undo his wrong action but we impose a supplementary wrong upon him. Thus, where an individual has stolen 5 pounds, enforcement will make him give it back, while retributive punishment will make him give it back *and* impose a fine of 5 pounds. In the latter case, retributive punishment is irrelevant to restoring the rightful situation, because that situation is beyond restoration. Thus, where an individual is punished for murder his punishment (however else it may be justified) does nothing to bring the victim back to life. Likewise it is impossible to un-stab or un-rape someone.

Another way of viewing this distinction between enforcement and retributive punishment is to focus upon the harm done to the criminal. In the case of enforcement the criminal is done no wrong because he has no right to profit from his crimes. Thus, the thief who is forced to return stolen goods is done no wrong because he has no right to those goods in the first place and his immoral action (theft) cannot give him a moral claim upon them. As the criminal is done no wrong in such cases and as enforcement maintains a rightful situation, the act of enforcement does not require any further justification. It would seem, *prima facie*, that a wrong is done to the criminal in the case of punishment.²⁸ Thus, not only is he deprived of that to which he has no right (i.e. his criminal profits) but he is also deprived of that to which he has a right. In the case of a fine he is deprived of his property, in the case of prison he is deprived of his liberty and in the case of execution he is deprived of his life. All of these are things to which agents are generally entitled.

Again, the material differences between enforcement and retributive punishment are seen when the claims that they entail are considered. The enforcer says "you have no right to this thing, and therefore I do you no wrong when I take it from

²⁸ Of course, *if* punishment is morally justified then no "wrong" is done to the criminal. But it can be admitted that on first inspection punishment is a wrong done to the criminal in a sense that enforcement is not.

you". The punisher says "although you would normally have a right to this thing by virtue of your criminal actions I now have a greater right to deprive you of this thing, and therefore, I do you no wrong when I take it from you". Given an understanding of the meanings of the terms used, the first claim is logically self-contained because it is *analytically* true. The second claim is also logically self-contained in that the conclusion follows from the premises, but it is only *synthetically* true. In this case for the conclusion (that retributive punishment is justified) to be other than arbitrary, the premises themselves must be established as independently true. Most pertinently, the conclusion will not follow unless we can show that *by virtue of a criminal action, the punisher gains a right to deprive the criminal of that to which he otherwise has a right*. And a theory of retributive punishment framed in terms of equilibrium must show that the exercise of such a right is necessary in order to restore the equilibrium in order for the punishment to be justified.

Kant's talk of a hindrance to a hindrance of freedom does not, on the face of it, add anything to the argument because the victim's right has already been hindered and it is too late to "hinder that hindrance". This line can only make sense with regard to punishment if there is some ongoing hindrance perpetrated by the criminal which remains in need of "undoing" after the commission of the immediate offence. We have already seen that the claim that a criminal ought not to keep the profits of his crime is true. Maybe, then, there is some ongoing profit which the criminal may rightly be deprived of. In this case we need to consider what might constitute profit apart from the immediate material booty of an offence against property and how the criminal might be deprived of such profit. If there is such an ongoing harm then punishment would be analogous to a combination of self-defence and enforcement: defence of a right which is being breached in an ongoing sense and/or enforcement of a duty which is being ignored in an ongoing sense.

In the case of *any* crime the criminal gains an advantage over everyone else in that he takes for himself a freedom to act outside of the constraint of the law.

Law, we have seen, is that which guarantees the freedom of each, compatible with the freedom of all. Crime is that expression of external freedom which denies the freedom of each, compatible with the freedom of all. Because of the law, every individual is specifically prohibited from expressing his external freedom in this manner. The criminal takes an advantage which is not open to other members of society and has no justification for doing so.²⁹ This is to say that, given a valid theory of law, and given the presupposition that individuals are free to act other than they did, the criminal act is always wrong. And, importantly, it is wrong independently of the material profits that are gained from it; if I beat you up this is qualitatively wrong in exactly the same sense as if I beat you up and steal your wallet, albeit that quantitatively the latter offence may be worse. The "profit" ("advantage" might be a better word) gained from *any* criminal act is that the criminal has overridden the rights of another. The problem now is that the right to property, which when overridden leads to a material profit, is more easily restored than the right to freedom, which when overridden leads to a non-material profit.

Enforcement constitutes a hindrance to a hindrance of freedom because (taking theft as our example) the criminal's retention of his victim's goods is an ongoing hindrance of the victim's freedom to dispose of them as he will. Therefore, when the law forces the criminal to return such goods this is a hindrance to a hindrance. With regard to the non-material profit we must first establish whether it is "ongoing" in any meaningful sense, because if it is not then it would appear to be too late to hinder the hindrance. The sense in which we could say that the benefit is ongoing is that the sum total of external freedoms which the criminal

²⁹ Which is not to say that any one criminal may not have a very good reason for doing so. The classic example in this vein is the starving waif who steals a loaf of bread. However, no matter how good a reason such a person may have, the generic features of his action - arbitrarily taking a freedom which is denied to all other members of society - are not justified. Thus his action may properly be characterised as a crime, albeit that the exceptional circumstances may lead us to treat him differently from other criminals. It would be particularly pertinent to consider the fact that one facing starvation as his only alternative does not act as freely as one whose choice is between, say, remaining relatively poor and stealing to become extremely wealthy. (cf. Kant 1929, 477; we view criminals in two ways: one as empirically determined, the other as absolutely free.) An alternative way of interpreting such a situation is to say that the starving waif is not a criminal because his actions in this instance are not arbitrary but rational: the freedom to take food is open to all agents *who are starving*.

has over the course of his life has been increased by the amount of freedom which he claimed in acting criminally. In terms of freedoms, the criminal has increased his freedoms in proportion to his offence, because in acting as he did he acts as if he has a right, greater than the right which he has overridden.³⁰ Whichever way we conceive of it the criminal has an ongoing benefit in that *without* punishment he has effectively claimed for himself these extra freedoms. That is, without punishment, the criminal has not only acted *as if*, he had those freedoms as rights, but has *actually* enjoyed them as if they were rights. Punishment would then be justified by the need to ensure that the criminal does not have these unfairly gained freedoms. The very act of punishment is what negates the freedom which the criminal has purported to take for himself. In denying him external freedom by "hindering" him through punishment we take back the freedoms which he stole from the sum total of possible human freedoms.

Presented thus the justification of punishment as a hindrance to a hindrance of freedom is highly problematic. For one thing, all we do in punishing is to *further* diminish the sum total of possible human freedoms. If C murders V he has claimed a huge freedom for himself and taken a huge freedom from V. If we punish him we take a huge freedom from him but we do not restore the huge freedom to V. As Richard Posner has it, once a crime is committed the costs become "sunk costs" (1980, 74). It is not even true to say that we take back exactly the same freedom from C as he took from V unless we resort to the crudest form of the *lex talionis*. Even then in the case of crimes such as rape, fraud and destruction of property it is not necessarily possible to do to C what he has done to V. This does not mean that punishment is wrong but it may mean that the metaphor of taking back a freedom wrongly appropriated from the sum total of possible human freedoms is obfuscatory. Having said this, such a theory of punishment is one of the most commonly defended in recent jurisprudence. It therefore merits a much more extensive consideration and this will form the basis of chapter four of my thesis. For the time being it is sufficient to conclude

³⁰ The criminal cannot gain *rights* because nothing he does alters the truth status of the categorical

that this is a potentially fruitful line of inquiry but one which Kant himself does not develop beyond the point of stating that punishment is justified because it is a hindrance to a hindrance of freedom. Put in this way Kant offers not much more than a formal declaration of what an equilibrium theory looks like. It remains for those working within the Kantian tradition to add substance to the theory and demonstrate how it is that punishment restores an equilibrium of freedom.

2.3 Punishment as right

For Kant punishment is by definition the right response to crime. However, he is clear that

[n]o 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. . . (1991, 143-144)

By way of a caveat to this quote he says:

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. (1991, 144)

This is to say that punishment is right and that as a (potentially) rational being the criminal must consent to the fact that it is right. In this sense it might be said to be *his* right (as to which see further chapter three section two). However, it is unimportant for Kant whether or not the criminal does consent to his punishment or regard it to be right. The right involved in punishment is vested in the legislator/judge (in Kant, the “sovereign”) who not only has a right to punish those disregarding his laws but has a categorical duty to do so (see 1991, 141). This is tempered by a right to grant clemency in exceptional circumstances i.e. where greater right is endangered by punishing (1991, 145).

3 The Shipwreck Example

Kant says that

there can be no *penal law* that would assign the death penalty to someone in a shipwreck who, in order to save his own life, shoves another, whose life is equally in danger, off a plank on which he had saved himself. For the punishment threatened by the law could not be greater than the loss of his own life. A penal law of this sort could not have the effect intended, since a threat of an evil that is still *uncertain* (death by judicial verdict) cannot outweigh the fear of an evil that is *certain* (drowning). (1991, 60)

This section of *The Metaphysics of Morals* (Appendix II to the Introduction *The Right of Necessity*) merits considered attention for three reasons. First, it suggests a deterrent theory of punishment which is contrary to orthodox interpretations of Kant. Second, as he mentions it in the same work as his main writings on punishment, it cannot be dismissed as an example which Kant had forgotten or overlooked. Third, he mentions the same example elsewhere - the earlier essay *On the Proverb: That May be True in Theory, But is of No Practical Use* ("Theory and Practice") (1983, 80 in footnote to page 79) - so the conclusions that Kant draws from the shipwreck example are something which he consistently held to be true.

I have already contended, while discussing Murphy's work at the outset of this chapter, that deterrence cannot be a sufficient reason for punishment within Kantian moral philosophy. The second formulation of the categorical imperative commands that one "[s]o act as to treat humanity, whether in thine own person or in that of any other, in every case as an end withal, never as means only" (1987, 58). This rules out the punishment of an innocent in order to prevent public disorder because to do so would be to use a human being "as [a] means only" (to the end of maintaining public order). Hence deterrence as *sufficient* for the justification of punishment is ruled out. This is not to suppose that Kant is never inconsistent, but it is safe to assume that his intention when discussing the shipwreck example is not to contradict his own categorical imperative. This particular assumption of consistency must be a significant pointer in interpreting the shipwreck example.

Even where the deterrent aspect of punishment treats the individual as a means to an end we do not treat him *merely* as a means *if* some other justification of punishment allows for the individual to be treated as an end in himself. A formulation of this might be that the conditions for justified punishment are (at least) that

- (1) the individual has committed a crime³¹; *and*
- (2) the punishment must have some deterrent effect

Thus in the shipwreck example Kant maintains that the survivor *has* committed a crime - *is* culpable - but is unpunishable because the punishment could not have a deterrent effect.

Against this reading Kant also states that

if a civil society were to be dissolved by the consent of all its members . . . the last murderer remaining in prison would first have to be executed, so that each has done to him what his deeds deserve. . . (1991, 142)

In this case it is clear that the punishment could have no deterrent effect. Once civil society is dissolved the (legal) categories of crime and punishment are immaterial. Neither the criminal nor anyone else will be deterred from crime by the thought of a prior punishment because the possibility of that punishment being repeated "dissolves" along with civil society. This is a rejection of the idea that punishment requires some deterrent effect in order to be justified. Hence, deterrence of another is ruled out as a justification of punishment. The criteria for justified punishment can now be redrawn as

- (1) the individual has committed a crime; *and*
- (2.1) the (threat of) punishment *could possibly* have had some deterrent effect on the criminal

³¹ Assuming what has yet to be proved, that to punish someone because he has committed a crime is to treat him as an end in himself.

If this is what Kant is saying in the shipwreck example we are not presented with a deterrence theory in any orthodox sense. For most deterrent theorists the institution of punishment derives legitimacy from its ability to prevent contemplated breaches of the law and the act of punishment derives legitimacy from its ability to prevent further breaches of the law. Thus, punishment is supposed to actually reduce the number of crimes. But for Kant the justification of punishment is not that it actually deters anyone from crime. Instead an individual act of punishment is only justified if its threat *could have been* deterrent, in the sense that it could have forestalled a bad outcome. Rather than being legitimated by deterring people from crime, punishment is *delegitimated* where it could not possibly have changed the material outcome of a situation. In the shipwreck example the threat of punishment could not possibly have removed the inevitability of one life being lost. Hence, the second criterion operates as a constraint on the imposition of punishment. Criterion 2.1. can be rewritten as

(2.2) the absence of such circumstances as would have rendered the threat of punishment necessarily operationally ineffective in preventing the evil effect of the crime.

The use of the word "necessarily" in this formulation emphasizes that the test is not subjective with respect to the individual involved but is objective by reference to the possible responses of any individual in the given circumstances.

In the context of Murphy's reconstruction of Kant this makes good sense. We impose legal punishment as a reflection of moral punishment, the goal of legal punishment being to produce an external set of circumstances which conform to right, the goal of moral punishment being retribution for moral iniquity. Where the goal of legal punishment is unachievable, as in the shipwreck example where the wrong (death) is unavoidable, we do not punish. Not because punishment is unjustified but because we do not have the necessary faculties to bridge the gap

between legal and *purely* moral (i.e. between legal-moral and non-legal moral) punishment. To do this we would need to know how evil the survivor had been in letting his companion drown to save himself. Because the survivor has been to some extent evil punishment is justified but we, as imperfect judges, have no criteria to judge him on. We cannot tell his moral worth and the harm of his external action, which is the criterion normally employed as a reflection of his moral worth, is non-existent: he does no harm over and above what was inevitable and thus unpunishable. Ideally we punish mental state M. Really we punish external action A which is presumed to be a manifestation of M. Where A does not occur, where an evil person chooses not to offend, we do not punish because M is unknowable and thus, on its own, unjudgeable by us. Where A produces an outcome which it was impossible for the actor to avoid in any case we do not punish because A is effectively an act of nature and M remains unknowable and thus, on its own, unjudgeable by us.

When expressed in such a manner deterrence is seen as hardly relevant in the vast majority of cases. In a murder case, for example, it is not open to the defendant to claim that he was so provoked that no punishment could have deterred him. Indeed, it is illegitimate, under the sort of circumstances that Kant discusses in the shipwreck example, for him to argue that the victim provoked him to such an extent that most people would have reacted murderously.³² The relevant circumstances are that the individual had *no real choice* and thus 2.2 talks about preventing the evil effect of the crime rather than the crime itself. The case of provocation gives a choice between being severely provoked or facing punishment whereas Kant's example gives a choice between dying or facing the death penalty, i.e. between two outcomes ("evil effects") which are supposed to be equivalent. In this case the sense in which punishment is "deterrent" is that it must present an outcome which would be materially different from the natural outcome of not committing the crime. So what Kant requires is that punishment *could not possibly* reduce the evil effects of crime. And even here, as we have already seen, Kant is not looking at any *prospective*

³² It may be that such an individual would be excused for a different reason.

deterrent effect of punishment. So the fact that a punishment could not possibly prevent *future* evil effects (as in the case of civil society being dissolved) would not delegitimize the punishment. The last murderer in prison is punishable because his actions A are presumed to be a manifestation of (evil) mental state M and thus he is judged and punished on this criterion. All that can be said on the evidence of the shipwreck example is that if we could show that the threat of a particular act of punishment could not possibly have altered the material outcome of the case then we are not justified in executing that punishment. This much conforms with Kant's definition of law as that which constrains the will so as to bring about a desirable set of external circumstances. Law is impotent where a consequence is inevitable.

It is to be noted that the "material outcome" talked about here does not refer solely to the material outcome for the defendant. If this were the case then where the choice was between death and a lesser punishment (i.e. between two outcomes which are not materially equivalent) the defendant would stand to be punished despite the fact that in this case his crime must have been less serious than in the shipwreck example. Rather than material equivalence for the defendant we are looking for material equivalence with regard to all morally relevant beings involved in the situation. Thus in the shipwreck example there is a material equivalence between one man dying and the other dying because both are rational beings with a right to defend their own lives. In this situation no matter what punishment (or none) is threatened the outcome is materially equivalent because nothing can detract from the inevitability of one of the men dying. In such a case Kant holds that the imposition of punishment would constitute a gratuitous infliction of additional evil.

If an aim of a practice is distinct from its justification then we must hold that the practice *can*, in principle, be justified without achieving that aim. This would seem to be a probable representation of Kant's thought given the extent to which his moral philosophy is based on *a priori* statements leading to a series of claims about duty. Duty justifies action without reference to any particular end state.

None of this is to say that for Kant deterrence is *not* an aim of punishment. Deterrence is a good thing for the simple reason that people ought not to commit crimes. Therefore, if they are deterred from crime this is good. However, we must be careful to avoid the teleological conclusion that therefore anything which deters them from crime is also good. For Kant it is good if punishment happens to deter people from crime but punishment still stands in need of an independent justification which is conceptually distinct from any good end state that it may bring about.

An alternative reading of the shipwreck example is extracted from a short section headed *The Right of Necessity*. Kant is considering the arguments for a right to do undeserved evil to others in extraneous circumstances. In doing so he is unequivocal in stating that such undeserved evil is morally *and legally* wrong (1991, 60-61). Kant maintains that I have in fact no *right* to take the life of an innocent other, even in the conditions of necessity (“[i]t is evident that were there such a right the doctrine of Right would have to be in contradiction with itself”) and he maintains that such an act is a breach of the law: “there could be no necessity that would make what is wrong conform with law” (1991, 60). However, despite the fact that there is no right to take the life of an innocent other anyone who does so in circumstances materially equivalent to those outlined by Kant will not be punished; the killing is “not to be judged *inculpable* but only *unpunishable*”. Kant explains this by drawing a distinction between the *objective* law which maintains that the killing is wrong and the *subjective* sentence which ostensibly declares the killer innocent.

Kant is taking the pragmatic line that where the law leads to an absurd result we would be wrong to apply it. When considering the same hypothesis in the earlier essay *Theory and Practice* he states that

teachers of universal civil right justifiably authorize such measures as are required in emergencies. For the authorities cannot attach any *punishment* to the prohibition since that punishment would have to be death. It would have to be an absurd law that threatened anyone with death if he did not freely give in to death in dangerous circumstances. (1983, 80 fn.)

Here Kant addresses the fact that law is the application of an Ideal concept (Justice) to the material contingencies of human nature. The Ideal of justice dictates that the hypothetical killer does wrong and is therefore culpable. However, as we have already seen, the contingencies of the shipwreck situation are such that there is no real choice: the alternatives are death or death. This is borne out by Kant's distinction between the objective law and the subjective sentence. By an objective standard - the Ideal of justice - the survivor has broken the law; he has failed to fulfill an absolute legal-moral duty to another in a situation where he logically *could* have done otherwise. The duty in issue is the duty not to kill but it would be an absurd law which faced with a situation where one death is inevitable anyway (i.e. through no morally culpable action of the actors) then kills the survivor as well. Hence, the subjective sentence grants him impunity. There is no inconsistency between saying that the justification of punishment is to repay evil with evil and holding that punishment should be suspended in situations where, although the defendant has technically committed an evil, that evil could not have been avoided. To punish in such a case would be an absurdity because the mischief which punishment seeks to avoid (*viz.* the infliction of evil) could not have been avoided anyhow.

A germane question of the shipwreck example is, if the survivor has truly done *a* wrong, why not punish him in proportion to this wrong? If the shipwreck survivor does not merit the death penalty why not punish him to some lesser extent in proportion to his wrong (which is less than that of murder). This conclusion would seem to follow from Kant's insistence that the survivor is culpable because he has failed to conform with the law. But the point is that the only criteria we, as judges, have for judging the survivor's moral desert are his external actions, his mental state (i.e. how evil he is) being unknowable. The external actions in the shipwreck example merely reproduce the naturally inevitable and cannot therefore be taken as *prima facie* manifestations of evil. This makes sense of Kant's claim that punishment for man is merely a hypothetical imperative. It is hypothetical upon our desire to bring the real world into conformity with the ideal world. Where this is impossible, for example

where the external outcome which law seeks to avoid is inevitable, and because we do not have the skill to judge moral worth except as manifested in actions, punishment is unnecessary for us albeit that the perfect judge (God, the philosopher king, the perfect rational being) would be categorically bound to punish on the basis of the (knowable to such a judge) moral iniquity of the actor.

4 Conclusions on Kant's Theory of Punishment

Kant offers the reader a theory of punishment without being entirely convincing as to the justifications of punishment. It is clear that he himself thought punishment to be justified by analogy with an ideal derived from pure reason. It is debatable how successful he was in deriving such ideals in the field of moral philosophy generally and concerning punishment in particular. We are left with a suggestive legacy to be developed, reconstructed and defended against criticisms of Kant's theory in its original form.

4.1 Kant is clearly a retributivist. His statements as to the deterrent nature of punishment found both outside and within the *Doctrine of Right* can be reconciled with the retributivist sentiments found within that work. To state that justified punishment is a deterrence system does not entail that the justification (or even *a* justification) of punishment is its deterrent effect. Confusion on this issue is avoided by bearing in mind the distinction between the ideal and the phenomenal. To observe that particular punishers act from a motive (even where they believe this motive is the justification of the act) is not to say that the act is in fact justified by reference to the motive. For Kant punishment is justified because it is a necessary expression of the supremely important aspect of rationality, *viz.* freedom. Punishment is seen as a hindrance to a hindrance of freedom and hence in itself an expression of freedom. Kant is retributivist because punishment (as a hindrance to a hindrance) is justified by reference to no more than the initial crime (i.e. the initial hindrance of freedom).

4.2 Kant gives us a formal statement of punishment as justified in terms of its ability to restore some rightful equilibrium. The notion central to Kant's

derivation of right is freedom and the equation central to his justification of punishment is that of a hindrance to a hindrance of freedom equalling an expression of freedom. Kant's notion of a hindrance to a hindrance is formal. He correctly observes that if there is a right R there must be a concomitant right to enforce R. However, the examples of enforcement which most obviously fit within this model - deterrence, crime prevention, the restoration of stolen goods and debts - are qualitatively different from punishment. They ensure that V (or a potential V) is able to exercise a right which he ought to be able to exercise. In the case of punishment it is, on the face of it, too late to let V exercise his right: it has already been ignored by C. Nothing in Kant's work tells us why or how punishment is supposed to actually restore or maintain freedoms in a meaningful sense.

4.3 Kant's attempted justification of punishment is problematic because it is framed exclusively in terms of freedom and the right to freedom. It is not at all obvious how a freedom, once taken away, can be restored. If anything punishment seems to be a gratuitous additional diminishment of freedom (the criminal's freedom). While this may be intuitively appealing and even poetically just, Kant doesn't offer us a principled justification for this removal of freedom in terms of his supreme principle which demands that freedom be maximised.

Chapter Three: Hegel on Punishment

The purpose of this section is to consider Hegel's contention that in some way or other punishment restores right. The equilibrium restored by punishment is thus directly tied to the concept of right and the justification of punishment is in terms of punishment's ability to restore right. This takes us beyond the point reached by Kant. It will be recalled that a central difficulty with his justification of punishment was that it was framed exclusively in terms of freedom and failed to show how punishment was to be conceptualised as other than a diminution of freedom. By introducing an equilibrium in terms of right Hegel offers us a potential solution to this problem. The first point to be made concerning Hegel's notion of right is semantic but extremely important and concerns the meaning and translation of the German word *Recht*. The problem is that no single English word captures the richness of the German term. Thus, while Right is certainly not a mistranslation, it does miss some of the connotations of the German. Most importantly *Recht* can mean something akin to "the realm of Right" and can equally imply a "claim to Right". Furthermore, *Recht* refers to *just* law (rather than any law) and hence talk of *Recht* suggests our own rendering "natural law" rather than "positive law".³³ This might lead to the conclusion that "justice" is the nearest definition of *Recht* but even that term is tainted in English by the fact that some would see the application of a properly published positive law as "justice" (see further *Translator's Foreword* to Hegel 1967, vi; *Translator's Preface* to Hegel 1991, xxxviii; *Translator's Note on the Text* in Kant 1991, x-xii; Inwood 1992, 221, 259-260).

In the face of these difficulties it is hard to offer any solution other than "beware". Over time the best answer might be for English to assimilate the word *Recht* into its own vocabulary although given the direction of English-speaking jurisprudence this is unlikely to be recognised as a necessary (or valid) addition. I will generally avoid using the German in my own thesis for the simple reason that my inability to expand on the term satisfactorily means that its use could end

up obfuscatory rather than illuminating. However I will have recourse to the terminology of a "realm of right" or "justice" where I believe that this is the most helpful way of unpacking the Hegelian baggage. Certainly when I come to develop my own theory of punishment I will use these terms with more confidence as I will no longer owe a debt of complete fidelity to the original text.

1 Punishment as the Annulling of a Wrong

Central to Hegel's theory of punishment, and of prime interest to us, is his claim that punishment amounts to the annulment of crime. The claim is analogous to Kant's notion of punishment as a hindrance to a hindrance of freedom. This is important because Hegel argues that in the very process of annulment punishment restores right. First, bearing in mind the above warning about the ambiguities of *Recht* it is important to establish exactly what Hegel means by his claim that punishment restores right. Second, having clarified this point we must consider exactly how punishment is supposed to effect such a restoration.

First then, in what sense of the word "right" is right restored by punishment? Hegel, defining crime states that

[t]he initial act of coercion as an exercise of force by the free agent, an exercise of force which infringes the existence of freedom in its concrete sense, infringes the right as right, is crime - *a negatively infinite judgement* in its full sense, whereby not only the particular (i.e. the subsumption under my will of a single thing . . .) is negated, but also the universality and infinity in the predicate 'mine' (i.e. my capacity for rights). (1967, 67; emphasis added)

In the translator's notes "negatively infinite judgement" is interpreted as signifying a "total incompatibility between . . . [the] . . . victim and rights" (1967, 331 note to paragraph 95). In so far as the predicate "mine" is not limited to the victim - rather it could pertain to any rational being with a will - I would take the negatively infinite judgment to be a total incompatibility between *rational beings* and rights, which in Kant's theory, for example, is an unsustainable

³³ cf. Inwood (1992, 260) for the suggestion that there is a distinction between legal and moral rights in the work of Hegel. See further Riedel 1971, 137 *et seq.*

contradiction. However, for immediate purposes this distinction is not especially important. What is to be noted is that for Hegel not only the particular right is negated by crime but also the victim's (or "all beings" on my reading) *capacity for rights*. It would be arbitrary for the criminal to say that "the victim has no right (say) not to be assaulted". Rather the criminal must claim that "the victim has no rights". This is because the victim's right not to be assaulted is no more than a particular instance of the victim's ability to bear rights.³⁴ Certainly there are many instances where a person may not have some particular right while maintaining the ability to bear rights; children, the mentally impaired and prisoners are all pertinent examples. However, in each case there is an independent reason for the denial of particular rights. The criminal does not have recourse to such an independent reason as to why, for example, his victim should be denied the right not to be assaulted, because if he were then, by definition, his action would not be a crime.³⁵ The argument then is that V has a right (R) because V has the ability (A) to have rights. This may be restated as

If A then R

The criminal's action is a denial of R; i.e. he is saying that V does not have R. By the application of logic if V does not have R then this entails that V does not have A. This can be restated as

-R therefore -A³⁶

³⁴ Employing Gewirthian terminology here the right not to be assaulted is no more than a particular application of the victim's right to freedom and well-being and these rights arise because of the fact that he is an agent.

³⁵ It is important to remember that we are talking about an Ideal-type here and not a particular criminal standing trial. Criminals in the dock offer all sorts of defences but what interests us is a hypothetical figure who we can be certain *is* a criminal. In order to be defined as a criminal such a person must, because of the definition offered in the introduction to this thesis, have infringed the rights of someone.

³⁶ Of course, I have already pointed out above that there are more reasons than a general inability to bear rights as to why V may be denied a particular right. If these other reasons such as youth, infirmity etc. are denoted j,k,l etc. then the logical progression is as follows

If A and -(j,k,l) then R

-(j,k,l) and -R therefore -A

Thus, for Hegel a crime denies both the particular right and the ability to bear rights. For purposes of the above I said that it was not especially important whether we meant the ability of V to bear rights or the ability of all beings to bear rights because the argument to “-R therefore -A” follows in either case. However, I will now offer reasons why I think that “-A” is properly unpacked as implying that no beings have rights rather than merely being V has no rights.

(i) I have already suggested that the predicate “mine” could pertain to any being with a will. The negation of the “*universality and infinity* in the predicate ‘mine’” (1967, 67; emphasis added) implies a total denial of the category “mine” rather than the denial of a particular being’s use of the word. Therefore it implies that all beings with a will cease to have a capacity for rights (NB in the original text the predicate “mine” is an equivalent for the phrase “my capacity for rights”).

(ii) It would in any event be arbitrary to limit the criminal’s action to denial of all the rights in a particular being V. There is no relevant difference between V and other rational beings (NB we have specifically excluded from the hypothesis the notion that V lacks the relevant rights for some reason such as age or mental impairment). All rational beings have rights because of their relevant similarities i.e. the fact that they are all rational. Therefore to deny that one rational being has rights (all other things being equal) is to deny that all rational beings have rights. To see this we may consider the following syllogism;

(PI) All rational beings have rights

(PII) V does not have rights

(Conc) Therefore V is not a rational being

We know that the conclusion is false because V is a rational being. Therefore, one of the premises must be false. In fact we know that the false premise is (PII)

I am taking the unit -(j,k,l) as read, because we are assuming that C is in fact a criminal (i.e. that

because we know that V does have rights: there would not be an issue of crime if V did not have rights. However, it is not open to the criminal to say that (PII) is false: the whole point is that by his very actions he has denied that V has rights and this much we have already established on the basis of the argument to “-R therefore -A” above. By his very action C is committed to the truth of PII. In consequence, the false premise from the perspective of the criminal must be (PI). In effect, by his action the criminal is committed to the claim that “(PII) is true” *and* to the claim that “(Conc) is false” (because the Ideal criminal offers no defence such as “V is not a rational being”; he is a heuristic stooge who accepts that he has committed the crime but is asking “even so, why ought I to be punished?”). Therefore he must also be committed to the claim that “(PI) is false”. This cuts the link between rationality and rights. It denies the stage of Gewirth’s argument where he applies the logical principal of universalizability to move from the prudential claim that “I have rights (because I am an agent)” to the moral claim that “all agents have rights (because they are agents)”. All of this is by way of showing that a criminal, in denying that V has any rights, must be committed to the claim that relevantly similar (i.e. rational) beings do not necessarily have rights (i.e. he is committed to the claim that rights do not flow from rationality). In Gewirth and Kant rights are derived from rationality; the criminal has committed himself to a denial of this and hence he has committed himself to a denial of rights *per se*.

(iii) Finally, I think the very point that Hegel was making was that an attack on the particular right of a particular being is an attack on *Recht* in the sense of the entire realm of right. The pertinent dicta here are “crime - a negatively infinite judgement in its full sense” and “[t]he initial act of coercion . . . infringes the right *as right*” (1967, both at 67; emphasis added). Crime attacks not a particular right but “right as right” or in other words the very notion of right. Inwood in his explanation of Hegel’s theory of crime and punishment states that

[w]hat is wrong with crime is thus not that it is unpleasant or inconvenient for its victims, but that it is an *Unrecht*, an attack on

he has no defence such as V did not have particular right R because of particular reason j).

right as such, usually in the form of an attack on a particular person or his property. (1992, 233)³⁷

This has resonance with those theories considered in the next chapter which are couched in terms of an equilibrium wherein crime is wrong, not because it hurts the victim, but because it upsets a pre-ordained balance of benefits and burdens. However, the point for Hegel is that an attack on an individual and an attack on the entirety of rights are one and the same thing: V does not have R entails V does not have A entails no rational being has A. He does not discount the victim; in fact the wrong against the victim *must* be righted otherwise the whole realm of rights is thrown out of kilter. It is perhaps such an understanding which Kant had reached when he made his seeming overstatement "if justice goes, there is no longer any value in men's living on the earth" (1991, 141).

My first conclusion then is that when Hegel talks of punishment restoring right he means both the particular right attacked and *Recht* in the sense of a realm of rights. This can be seen two ways. On the one hand the restoration of the realm of rights guarantees the particular right which dwells within it. On the other hand the restoration of the particular right guarantees the security of the realm just as the mending of a hole assures the safety of a ship. This dual sense in Hegel's theory perhaps hints at the richness of the term *Recht*.

Having established Hegel's meaning when he asserts that punishment restores right, we must ascertain exactly how punishment is supposed to do this. In order to conduct such an inquiry it is necessary to understand the Hegelian concept of *negation*. For Hegel, crime negates right and the function of punishment is to negate this negation and thus reassert right. This is a three stage process: assertion; negation; negation of negation. The assertion - that there are rights - we can take as read. The second stage or the negation we can also take as read on the basis of the paragraphs immediately preceding this one. We have already seen how a crime (or the maxim of a crime) denies not only the particular right

³⁷ Hegel, unlike some of the theorists considered in the next chapter, does not let this metaphysical conceptualisation of crime blind him to the fact that a crime is also a particular crime against a particular right (1967, 67-68).

of the particular victim but by inference denies all rights of all individuals. What remains to be seen is how punishment amounts to the negation of this negation and hence to a reassertion of right.³⁸ In fact, put like this, such a question may be seen as *the* question facing those retributivist theories based on a notion of equilibrium. The key passage in Hegel is the following:

In so far as the infringement of the right is only an injury to a possession or to something which exists externally, it is a *malum* or damage to some kind of property or asset. The annulling of the infringement, so far as the infringement is productive of damage, is the satisfaction given in a civil suit, i.e. compensation for the wrong done, so far as any such compensation can be found. . . . But the injury which has befallen the *implicit* will (and this means the implicit will of the *injuring* party as well as that of the injured and everyone else) has as little positive existence in this implicit will as such as it has in the mere state of affairs which it produces. In itself this implicit will (i.e. the right or law implicit) is rather that which has no external existence and which for that reason cannot be injured. Consequently, the injury from the point of view of the particular will of the injured party and of onlookers is only something negative. The sole positive existence which the injury possesses is that it is the particular will of the criminal. Hence to injure [or penalize] this particular will as a will determinately existent is to annul the crime, which otherwise would have been held valid, and to restore the right. (1967, 69; in this quotation only the ellipsis have been added)³⁹

This has several stages:

(i) a crime in its external aspect (the "mere state of affairs") is undifferentiated from an accident or an act of God. Of course it is a bad thing if someone is killed but if we look merely at the outcome of events then we can say no more about a murder than we can about a disease or a volcano which kill. So in this sense the

³⁸ It is to be noted that it is not sufficient here to insert logical units. The progression assertion: negation: negation of negation = assertion is sound in such terms because it amounts to $-(-X) = X$ which is true. However, what we need to demonstrate is that punishment *is* the negation of the negation of right. If this can be shown *then* by the application of $-(-X) = X$ we can show that punishment also corresponds to the reassertion of right.

³⁹ In this passage Hegel neatly spots and sidesteps one of the difficulties in Kant's theory of punishment. He states that the annulment of material damage is effected by compensation after a civil suit. This, it will be recalled, was the most obvious interpretation of the Kantian notion of enforcement. By indicating that the injury which punishment annuls is something other than the immediate material damage Hegel opens the way for consideration of a metaphysical sense in which punishment might be said to restore equilibrium. It might be added that Hegel sidesteps one difficulty and moves into the path of another considerably larger problem. Ultimately though, my own justification of punishment, offered in the conclusions of this thesis, will take the same deviation from the Kantian path as Hegel takes at this point.

bombing of my car no more attacks my right to possession than if a tree were to fall on it.

(ii) furthermore the crime is not committed against right as a concept (“the implicit will”) because nothing in a crime diminishes right (i.e. as a concept).

(iii) but what necessitates action against a crime is that it is a *willful* attack on right as an actuality. That is, it is an operation of the explicit will (of the criminal) against the realised implicit will (of the criminal and others).⁴⁰ Left unchecked this operation is “held valid” and right is given no actuality; right is effectively killed off.

(iv) therefore that which must be the target of response is the will of the criminal. For Hegel this response takes the form of an “injury” or “penalty”.

In his own notes Hegel sums this up as follows:

the only important things are, first, that crime is to be annulled, not because it is the producing of an evil, but because it is an infringement of the right as right, and secondly, the question of what that positive existence is which crime possesses and which must be annulled. . . (1967, 70)

His argument that if we merely look to the “producing of an evil” there is nothing to distinguish between a crime and an accident is good. We are not punishing acts, we are punishing actors. Furthermore, I think Hegel is right to identify the explicit will of the criminal as the seat of the crime; the explicit will is the “positive existence . . . which must be annulled”. Without such a response the attack on right would be “held valid” i.e. it would be legitimated. Hence we can see why the negation of a negation (i.e. the negation of crime which is the negation of right) amounts to a reassertion. By quashing any attack on right we

⁴⁰ It is the implicit will of all rational beings that there are rights. By realised implicit will I mean an *actual* right i.e. a right which is actually possessed and enjoyed by an individual.

reassert the legitimacy of right; we refuse to recognise the legitimacy of anything which seeks to usurp right.⁴¹

For Hegel, the annihilation of an attack on right takes the form of positive action against the criminal. It does seem to me that Hegel's argument leads us to this point but that this is the most difficult hurdle to cross: exactly why is it necessary to "injure", in Hegel's terminology, in order to quash this uprising in the realm of right; is this perhaps an example of the metaphor dictating the reality, talk of a *realm* and an *attack* on that realm making the logical conclusion that only a *forceful* repression will be successful? What Hegel has shown is that it is the *will* of the criminal which has thrown the continued existence of rights into danger; if we legitimate one breach of right then by implication other breaches are legitimate and so rights cease to exist as anything other than a nebulous concept. So action must be *against* the will of the criminal; this alone is the legitimate target. Therefore the action must be something which directly impacts the will of the criminal. Insofar as the will expresses itself through freedom it must be something which affects the freedom of the will. This may sound like a spurious derivation of the prison service in one easy step but it would be missing the point if it were thought that only a limitation of freedom of the body affects a limitation of freedom of the will. A limitation (to a greater or lesser degree) on the ability to exercise choice can take many different forms. It remains to be proved that this is the correct response to crime.

2 Punishment as the Criminal's Right

We have seen above how Hegel considers punishment to be the necessary annulment of a crime. However, he goes on to claim that not only is punishment necessary to reassert right but it is also in fact *the right of the criminal to be punished*. He states that

⁴¹ This argument doesn't always follow. In quashing resistance the Nazis did not establish the legitimacy of their own regime. The point is that we have already made an argument for the legitimacy of right. What is now at stake is the recognition or non-recognition of anything which stands against right.

[t]he injury [the penalty] which falls on the criminal is not merely *implicitly* just - as just, it is *eo ipso* his implicit will, an embodiment of his freedom, his right; on the contrary, it is also a right *established* within the criminal himself, i.e. in his objectively embodied will, in his action. The reason for this is that his action is the action of a rational being and this implies that it is something universal and that by doing it the criminal has laid down a law which he has explicitly recognized in his action and under which in consequence he should be brought as under his right. (Hegel 1967, 70; Hegel's own addition to the text)

There are two parts to this, one of which I think is uncontentious (certainly uncontentious from a Kantian view-point) and the other with which I disagree. The two parts correspond with two senses in which it might be said that the criminal wills his own punishment. First, the criminal *implicitly* wills his own punishment. This is because (the rest of Hegel's theory being taken as read) punishment is the rational response to crime and therefore as a rational being the criminal is bound, on pain of self-contradiction, to accept that his punishment is justified. Under this formulation we say that all rational beings are bound to agree that punishment is the right response to crime and therefore any rational being who commits a crime implicitly consents to his punishment. As we have already seen Kant, using different terminology but with similar intent, states that

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. (Kant 1991, 144)

I find this aspect of Hegel's assertion that the criminal has a "right" to be punished relatively uncontroversial. To say that punishment is justifiable is precisely to say that in principle any rational mind can consent to its propriety in given circumstances. Hegel takes this to mean that the criminal has a right to be punished: "it is *eo ipso* his implicit will, an embodiment of his freedom, his right" (1967, 70). I am not sure that the mere fact that someone must rationally consent to something makes that thing a right: I must rationally consent to the assertion "I exist" yet it is not a right that I exist. Be that as it may, the claim found in both Kant and Hegel that the criminal is rationally bound to consent to

his punishment, is sound and indeed uncontentious given a proper understanding of “justification”.

Second, Hegel believes that by the very act of crime the criminal has *explicitly* consented to his own punishment. This is because the logical inference from his crime is that he, the criminal, does not have a right of the relevant nature; if he has denied to V the right not to be assaulted he must deny himself this right. Again there are Kantian overtones: Kant’s moral theory is famously grounded in the dictum “[a]ct only on that maxim whereby thou canst at the same time will that it should become a universal law” (Kant 1987, 49). However, the point for Kant is that a criminal *cannot* will that the maxim of his crime should be a universal law. For example, with theft the maxim might be “all people are free to help themselves to whatever property they want”. For Kant a rational being could not consistently will this because it would jeopardize his own property as much as that of his victim. Furthermore, to will such a maxim leads to a *logical* inconsistency because it ultimately entails that there is no such institution as property. Thus the criminal would simultaneously claim to take ownership of something while, by virtue of his actions, denying that ownership is possible. Schematically, the difference between Hegel and Kant on this point may be shown as follows:

A Hegelian Progression

- (i) You have committed act A predicated on maxim M.
- (ii) By the principle of universalisation if we act on maxim M we will be allowed to commit a similar act A on you.
- (iii) Therefore “punishment” (that is the commission of act A on the you) is willed by you.

A Kantian Progression

- (i) You have committed act A predicated on maxim M.
- (ii) By the principle of universalisation if we act on maxim M we will be allowed to commit a similar act A on you.
- (iii) You could not consistently will (ii) above; therefore you ought not to have committed act A; act A was *wrong*.
- (iv) Since act A was wrong, and on the basis of an independently justified theory of punishment which allows punishment of those who do wrong, we are justified in punishing you.

The important difference to note is that for Kant the only thing which follows from (i) and (ii) is that the act is wrong whereas Hegel employs (i) and (ii) as the basis for his claim that the criminal explicitly wills his own punishment. Of course for Kant (i) and (ii) lead to (iii) which for him is the basis for saying that a criminal as a rational being *implicitly* wills his own punishment but this is a quite different claim. For one thing it is still necessary to show that a rational being is bound to accept that punishment is the correct response to crime; although Hegel also shows this, his progression, as I argue below, actually sidesteps the need to do so by saying that the criminal has willed his own punishment.

My objections to the idea that the criminal explicitly wills his own punishment are several.

- (i) For Kant (and I think he is right on this point) the reason a crime is wrong is precisely that the agent *cannot* consistently universalize the maxim on which his action is predicated. It is therefore gratuitous and, more importantly, illogical, to justify punishment by means of universalizing the criminal's maxim. Thus Kant's account of crime negates this particular attempt by Hegel to say that we can justify punishment by reference to the explicit will of the criminal. For what it is worth I do see the force in saying to the criminal "if we treated *everyone* as you have treated the victim then surely you must acknowledge that this would be a bad state of affairs". This is simply a restatement of the classic reprimand

“would you like it if I did that to you?” Its force lies in its potential to make the criminal realise *why* his action is wrong and perhaps even *why* he ought to be punished. However, it would be inconsistent to say on the one hand “action A is wrong” and on the other “you the criminal have, by your action, said that action A is okay therefore we will do it to you”. Why should (and how could) the criminal’s statement that, for example, murder is okay, have any bearing on actuality: why should (and how could) it mean that murdering the criminal is justified? This can be seen even more clearly when it is realised that even those theorists who stand by the death penalty do not consider it to be murder, but rather a lawful killing. Therefore, even in this extreme case we cannot say that the criminal’s maxim (“murder is okay”) has been universalised. The bottom line is that if the criminal’s action cannot be consistently universalised then we would be wrong to treat it as if it could. Crime, by definition, is wrong and justified punishment, by definition, is right. I think that Hegel himself comes close to this when he makes the point that

[i]f crime and its annulment . . . are treated as if they were unqualified evils, it must, of course, seem quite unreasonable to will an evil merely because ‘another evil is there already’. (1967, 69-70)

Here Hegel recognises that crime and punishment are qualitatively different actions, in which case the attempt to derive the justification for punishment from some maxim of the criminal is fruitless.

(ii) To say I have a right to something is not the same as saying that I will it, because in many instances I have a right to waive my right. For example, if I, an ascetic, inherit a fortune I may well have a right to the money, but it is not an action of my will (it is literally the action of another’s will) which has given me that right and, more significantly, it might be directly against my will to receive the money. This is by way of a demonstration that we could theoretically hold that the criminal has a right to be punished, notwithstanding that he might will that he is not punished. Conversely the fact that I will something does not necessarily give me a right to that thing. For example, I might strongly will that someone kill me but this does not give me any right that another kill me. This

might sound far fetched but it is no more so than the proposed derivation of a right to be punished from the willing of a crime.

Furthermore such a derivation (of a right to be punished from willing a situation) has positivist ramifications. The argument is that we punish those carrying out action A; X does A; therefore X wills that he is punished. However, if A is "resisting fascism" can we say that X has a *right* to be punished in this situation? The answer, surely, is that we cannot, because, there could hardly ever be a right (let alone a duty) to punish someone for resisting fascism. If it is *wrong* to punish someone for resisting fascism, then *ipso facto* that person cannot have a right to be punished. The argument from *we will punish those carrying out action A to X does A, therefore X wills that he is punished* only follows if action A is a punishable action. This can only be established on grounds independent of "X has done A and we said we would punish him if he did A". Only an independent justification for punishing action A will avoid the positivist trap.

(iii) Even if we can say that the criminal *wills* his own punishment I do not think that this is in anything more than the weak implicit sense already considered above. To say that treating the criminal in a bad way is a logical inference from the way in which he has behaved does not necessitate that he has *explicitly* willed his own punishment in the way that Hegel claims it does. At best punishment is implicitly willed; i.e. at best we can say that the criminal has by his actions suggested that action A is okay, therefore, *by implication*, it would be okay if we did action A to the criminal. On this point I think Kant has it right when he states that

[s]aying 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. (Kant 1991, 144)

Even accepting that the implication of the criminal's action is that A is okay we can hardly use this as the justification of punishment given that we are saying to the criminal "A certainly is *not* okay and you shouldn't have done it".

(iv) Finally, we might ask *if* the criminal has explicitly willed his own punishment is there any need to *justify* punishment as such? I have already hinted that the answer to this question is yes, because the mere fact that someone wills something does not necessarily give him a right to that thing. However, it does seem that to a certain extent Hegel's pronouncements on this issue would sidestep the need to justify punishment independently. In addition to the passages already quoted Hegel says that

what is involved in the action of the criminal is not only the concept of crime, the rational aspect present in crime as such whether the individual wills it or not, the aspect which the state has to vindicate, but also the abstract rationality of the individual's *volition*. Since that is so, punishment is regarded as containing the criminal's right and hence by being punished he is honoured as a rational being. (Hegel 1967, 71).

Here, the very fact that the criminal has willed a crime makes it his *right* to be punished and, Hegel claims, punishment is necessary so that the criminal is "honoured as a rational being". This makes punishment not only a right of the criminal but gives the state a concomitant *duty* to punish. So the objection that simply because C has willed something does not necessarily mean that we have to do it is now overcome; we have a duty to carry out C's will in order that he be respected as a rational being. But I find this argument unsustainable. Hegel's line is that (a) C has explicitly willed that he be punished (b) if we don't punish him we treat him as if he is not fully rational (i.e. we say that his explicit will is *irrational*) (c) therefore we must punish C. However, once again this only follows if it is in fact the case that it is rational for C to will that he is punished. This needs an independent justification for punishment to avoid circularity in the argument. And if we have an independent justification for punishment it seems neither here nor there that the criminal has expressly consented to his punishment.

2.1 Qualification

The statement "punishment treats the criminal as a human being" begs the question: it only treats the criminal as a human being *if* the correct response to a human being who offends is to punish him and this contingent clause is exactly what we are trying to demonstrate in coming up with a theory of justified punishment. However, there is a sense in which this statement does have a meaning and is to that extent helpful to us in our inquiry. What I mean now when I refer to a human being is a being with the capacity for rational action, a being who can freely chose to perform a particular action or not as the case may be. In this sense punishment, as opposed, for example, to medical intervention or quarantine, does treat the criminal as a human being. Punishment as a response is predicated on the idea that the criminal has the capacity for rationality (i.e. that he could be responsible for his actions). Treatment need only be predicated on the idea that the criminal is sick or malfunctioning and this could be equally true of an animal or a computer. Punishment then treats the criminal as a human being in the sense that it is a response which is only appropriate to beings who have the capacity to control their own behaviour whereas other responses would also be appropriate to non-rational beings.

3 Modern Interpretations of Hegel

At the end of section 1 of this chapter I stated that the most problematic step in Hegel's purported justification of punishment is the contention that a forceful attack on the will of the offender is necessary in order to restore right. This present section is a consideration of how Hegel's theory of punishment has been interpreted in the more recent literature. This is necessary because Hegel himself is rather terse and at times frustratingly chimerical in his own pronouncements. My focus will be mainly upon how writers have dealt with Hegel's idea that punishment is legitimated as a necessary attack on the deviant will in order to restore right. There is another preliminary point, connected to the interpretation of this. J.D. Mabbott claims that while he is himself a retributivist he differs from Hegel in that the latter maintains "that the essential connection is one

between punishment and moral or social wrongdoing” while he treats “punishment as a purely legal matter” (1971, 43).⁴² By now it should be clear that this is not a valid distinction within my own thesis. When I talk about the restoration of right I am referring to the restoration of rights which are simultaneously moral and legal, that is morally legitimated legal rights (rather than merely positive legal rights).

J.F. Doyle in *Justice and Legal Punishment* states that

an offence is a challenge, in some degree, to the integrity and justice of a legal community - in Hegel's words, it is an 'infringement of the right as right'. Nothing could be more contrary to just law than simply to ignore such an act. *But it is far from obvious that the appropriate or just response of a legal system to an offence is the infliction of evil.* (1969, 163; emphasis added)

This follows the progression (and problems) of my own interpretation of Hegel; the argument to “an attack on a particular right amounts to an attack on the realm of right” is sound and even relatively uncontentious. But it seems like a leap of fancy to conclude that the defence of right necessitates the infliction of (painful) punishment. For both Hegel and Doyle the solution to this problem rests (initially, at least) in viewing the actions which comprise punishment as something other than an evil. As Hegel points out it is the source of some confusion to view punishment as an unabated wrong done to the offender; to do so is to leave oneself defending the thesis that two wrongs *can* somehow make a right. The question now rising is, *if* we are not to view punishment as an evil then what *are* we to view it as? Doyle says that

we should describe justifiable punishment (if there be such) as the satisfaction of all the just claims invoked by the commission of criminal offences. (1969, 163)

I think that Doyle is on the right track here, i.e. that punishment is to be viewed as the satisfaction of a justifiable claim. However, his thesis suffers from being

⁴² Mabbott's boast that as disciplinary officer of his college he enforced rules on chapel attendance with which he disagreed (“They had broken a rule; they knew it and I knew it. Nothing more was

purely positivist. Within a closed system which says “action C (crime) will be met with response P (punishment)” and that “some person has a claim to enforce P” it is true to say that action C gives rise to an *internally* justified claim that P be carried out. But this is an unsatisfactory account because it does not tell us what justifies the closed system itself; it leaves us no get out clause when action C is “refusing to obey orders to kill someone”. When Doyle does come to deal with this issue his response is a rather circular quasi-utilitarianism. For example the claim that “[t]here would be no point in legally proscribing certain types of conduct as criminal . . . unless there were good reason to discourage its incidence” (1969, 167). This ascribes rather more *bona fides* to those in power than is always merited. In any case it misses the point that any government (logically) may proscribe activity which it ought not to. In this case we need to be able to demonstrate that the proscription and hence punishment are unjustifiable notwithstanding that the government *does* proscribe and punish the activity in question. The state may have pragmatically good reasons for proscribing behaviour even though it has no moral warrant for so doing. The account is also unsophisticated because it does not attempt to tie the notion of justified punishment to the concept of a right. Rather it posits a system where punishment *is* (stipulatively) tied to the concept of right. When I come to deal with this issue myself I will discuss whether simply by consideration of what it means to have a right we can show that a breach of that right gives rise to a claim for punishment. Given that rights can be grounded we would then have a grounded justification for punishment.

Specifying exactly which claim he considers punishment to satisfy Doyle states that

[w]ith other members of the legal community the immediate victims share a right to secure public censure of legal wrongs, as well as possible protection from the repetition of such wrongs. (1969, 165)

necessary to make punishment proper” (1971, 45)) was made in 1939. It would be interesting to know if the following six years changed his mind.

To this it might be objected that public censure does not necessarily require the infliction of pain (or indeed anything); certainly it seems more logical to say that punishment necessarily involves public censure than public censure necessarily involves punishment.⁴³ Further, "protection" is a wider concept than punishment in the sense that we may need protection from someone who is not a criminal - a madman for example. The need for protection is not entirely unrelated to punishment but as Hegel points out it is more in the nature of an afterthought than a justification:

[t]he various considerations which are relevant to punishment as a phenomenon and to the bearing it has on the particular consciousness, and which concern its effects (deterrent, reformative, &c.) on the imagination, are an essential topic for examination in their place, especially in connexion with modes of punishment, *but all these considerations presuppose as their foundation the fact that punishment is inherently and actually just.* (1967, 70; emphasis added).

Doyle's consideration of the Hegelian idea that punishment is justified as an attack on the will of the criminal necessary in order to restore right is unsatisfactory. First because, for Doyle, punishment constitutes the restoration of no more than a closed system; Doyle himself does not purport to answer fully the question of why we should adopt his closed system (1969, 168). Second because the claims that he identifies as the reasons behind punishment (public censure and protection) are contingent effects of punishment; *if* we justify punishment by reference to the need to uphold standards or protect the public then our focus will be upon a considerably wider constituency than criminals. Doyle employs as a justification for punishment something which in the original is no more than an incidental benefit of the practice.

It should be added at this point, of Hegel's own account, that while it must be right to say that justified punishment is not an evil, this is not the reason that punishment is justified. Thus, if punishment is justified it is *de facto* right and therefore a mistake to call it evil. But merely to stipulate that justified punishment is not an evil does not tell us why punishment is right: it does tell us

⁴³ Unless we view them as the same. But this would involve a very strange conception of either

why justified punishment is to be regarded as right but it does not tell us why punishment is to be regarded as justified which is the issue at stake.

Ted Honderich also considers, albeit briefly, this idea of Hegel's. He is dismissive of the idea but I believe this is based on a misreading of Hegel. Crucially Honderich focuses on the *physical* pre-crime situation and (correctly) noting that this cannot be restored says that Hegel is wrong to talk of the annulment of crime and the restoration of right. However, Hegel is clearly referring to the restoration of the realm of right rather than any physical state of affairs; his is a metaphysical restoration. Honderich replies

[t]he question then arises as to whether the restoration of a moral principle so conceived is something sufficient to justify the practice of punishment. I take the answer, given our commitment to the avoidance of unnecessary suffering, to be no. (1969, 37)

But this is to beg the question: the suffering is only unnecessary if we have decided that punishment doesn't restore right or that it is not ultimately important to restore right. Honderich, it might be said, has an overriding commitment to the avoidance of suffering while Hegel has an overriding commitment to the maintenance of right.⁴⁴ Thus for Honderich any maintenance of right which involves suffering is unjustifiable while for Hegel any avoidance of suffering which allows right to be breached is unjustifiable. No doubt this is to simplify considerably but it does make the point that Honderich's rejection of Hegel's thesis only follows if we share Honderich's overriding commitments. Given the epistemic grounding of right on which I base this thesis the reason suffering is bad is that we generally have rights not to have our wills coerced. This makes the possession of right epistemologically prior to the wrongness of suffering. The conceptual debate is as between the Hegelian (and Kantian and Gewirthian) line that "because we have a right not to suffer, suffering is bad" and Honderich's line (which dates back at least as far as Bentham) that "because

public censure or punishment.

⁴⁴ If we typify this as Honderich's overriding commitment for the sake of argument, then the word "unnecessary" before "suffering" is strictly superfluous; presumably all suffering which can be avoided is unnecessary.

suffering is bad, we have a right not to suffer". It is only the latter line which would render right subservient to the avoidance of suffering.

J.P. Day defends the thesis that retributive punishment restores right. However, if anything, Day's is an even bigger misreading of Hegel than Honderich's, especially given that Day, unlike Honderich, is claiming to work *within* the Hegelian tradition. For Day a retributive punishment is exacted when an authority requires that the criminal returns to his victim that which he has deprived him of or an equivalent. Thus, retributive punishment is effectively reduced to compensation or restitution (1978, 503; cf. Benn and Peters 1959, 177). Day claims that the retribution involved is a by-product of the restitution, "[f]or R cannot, logically, make C give back X or EX [equivalent to X] to V without depriving C of X or EX" (1978, 502 abbreviations in quotes from Day changed to conform with my own shorthand). Further, he claims that punishment is justified if and only if the good in the restitution (C returning X or EX to V) outweighs the bad in the retribution (C being deprived of X or EX). This is poor philosophy for the simple reason that C never has any title to X and therefore no moral wrong can be done to him by taking it away from him. Certainly we might identify a psychological harm to C but we are looking for a moral rather than a psychological justification of punishment. If we are simply to weigh psychological goods and harms then we are lost in the middle of Benthamite territory; but Day is clear that he does "not agree with Bentham's thesis that 'the end of (criminal) law is to augment happiness'" (1978, 510). Having claimed that Day's thinking is confused it is imperative to point out that Hegel is not guilty of this fault. Hegel expressly says that the wrong done as far as it concerns a possession is a *malum* the satisfaction for which is "given in a civil suit, i.e. compensation for the wrong done" (Hegel 1967, 69). As we have seen, he goes on

[b]ut the injury which has befallen the *implicit* will . . . has as little positive existence in this implicit will as such as it has in the mere state of affairs which it produces. (1967, 69)

Concluding

[h]ence to injure . . . this particular will as a will determinately existent is to annul the crime, which otherwise would have been held valid, and to restore the right. (1967, 69)

This does not tell us exactly what Hegel means by that idea but it should be clear that for him compensation is something very different from retributive punishment. Crucially, Hegel regards the physical effects of crime as distinguishable from the metaphysical effects. While it is no doubt important to undo the physical effects (insofar as this is ever possible), punishment is aimed at a metaphysical restoration. Day's philosophy on the other hand is steeped in the language of economics. He claims that

if C puts out V's left eye, monetary compensation by C (partly) restores V's human right to immunity from bodily injury. (1978, 504; cf. Nozick 1974, 135 who draws a distinction between compensation and punishment).

It is a moot point whether C ought to compensate V for his crime. But the vital issue is that Day thinks the compensation *is* punishment. It is central to Day's thesis that *if* somehow C could give V his eye back then he ought to do this rather than compensate him and that *this would then constitute retributive punishment* (1978, 504). This line of thinking leads Day to conclude that imprisonment can never be justified as retributive punishment because it could never constitute a restitution or compensation for a wrong done (1978, 508; cf. Garcia 1989, 266 who makes that the point that there are reasons other than the need to punish for making the criminal return the material profits of his crime).

In summary then, Day thinks that punishment is justified because it returns to V whatever he has been deprived of (or an equivalent). If we are to interpret Hegel in this purely materialist manner (and to know the first thing about Hegel is to know that we ought not to!) then Honderich is right to say that many crimes cannot be undone. Now in a sense, for Hegel, punishment *is* justified because it returns to V whatever he has been deprived of (or an equivalent) but this is in a strictly *non-material* (metaphysical) manner: V has a right restored to him. Unless we agree that the payment of compensation (NB as opposed to a fine)

restores, for example, V's right not to have his eye put out, we must reject Day's interpretation of Hegel.

Benn and Peters interpret Hegel's philosophy of punishment as "veiled utilitarianism" (1959, 177). It should be noted that this claim is based on the flawed logic that; Utilitarianism is forward looking; their interpretation of Hegel's justification of punishment is that it is also forward looking; therefore Hegel is a utilitarian. This is a classic *non sequitur*. This may seem like a summary dismissal of Benn and Peters but it is roughly as long as their treatment of Hegel (on this point) and so it is hard to read them with any more sympathy. They go on to question the Hegelian notion that "the idea of right, which law embodies, would be denied, unless it were reaffirmed through the machinery of punishment" (1959, 177). Their criticism is the argument that the idea of right could be reaffirmed simply by denunciation of the crime. We might admit that this is a sound *ad hominem* argument but the notion which it attacks is *not* Hegel's. Most significantly it is not the idea of right which punishment is to restore but rather right itself or the *ideal* of right.⁴⁵ Again the tendency to resort to metaphor raises this confusion. It is perhaps natural to think of crime as a *symbolic* attack on right for the reason that an average criminal does not commit his crime because he wants to intellectually challenge the philosophical concept of right but rather because he has a besetting appetite to possess the loot. By extension it is tempting to frame punishment as a symbolic restoration of right. However, while the initial attack is a symbolic attack on the *idea* of right it is an *actual* attack on right itself; it is an *actual* attack on the victim's right which logically commits the criminal to an attack on right *per se*. This might seem strange because right is not a material thing to be attacked: *are* right and the idea of right really separable in this manner? The answer to this lies in acknowledging that rights are concepts rather than things but that they are concepts which do impinge upon our relationships with one another. A derivation of categorical rights which proves to all rational beings that they have certain rights and that

⁴⁵ NB if it is only the *idea* of right that we are restoring then a mere assertion or censure seems well suited. However, if it is something more substantial than the idea of right that is in need of restoration perhaps something more than merely articulating that the right exists is required.

they ought to respect the like rights of others has real implications for the lives of those rational beings. A crime does not directly attack the derivation of those rights or the concept of those rights (see also Cooper 1971, 164). Rather it attacks the human relationships which have been shaped by those rights. A system based on rights gives rise to legitimate expectations by members of the system which in turn effect their interactions with one another. It is at the intersection of these relationships that crime strikes its wicked blow. For example, I have legitimate expectations of all others that no one steals my car and that no one beats me up. When one person does either of these things he has, temporarily at least, removed any practical import which those legitimate expectations have on my life; he has, as it were, cut the tie of right and duty between him and me. Punishment is aimed not at restoring a concept of right (which in any event remains immune from attack) but rather the practical outworking of the right. This is a point taken on and recognised by Bernard Bosanquet in his *Philosophical Theory of the State*. He typifies the offender as

a responsible person, belonging to a certain order which he recognises as entering into him and as entered into by him, and he has made actual an intention hostile to this order . . . In other words, he has violated the system of rights. . . (1951, 208)

Elsewhere he states, in agreement with Hegel, that

in a matter of criminal law there is involved an infraction of right as such, which by implication is a denial of the whole sphere of law and order. (1951, 214)

On these issues I think Bosanquet is absolutely in accord with Hegel. However, the question remains *how* is punishment supposed to restore the breached right? Bosanquet's answer to this is problematic. There seems to be a confusion in his writing as between a preventative theory of punishment and a deterrent theory of punishment (see especially 1951, 212-213). This is crucial because there is a sense in which Hegel's is a preventative theory of punishment (although certainly not in the orthodox sense) but his justification of punishment is clearly *not* made in terms of its deterrent effects (Hegel 1967, 70). First then, Hegel's theory is preventative in that punishment is justified because it is necessary in

order to prevent a successful breach of right. As Bosanquet has it, crime is “a denial of the whole sphere of law and order” (1951, 214) and for Hegel *unless* we punish this denial is held to be valid. In other words punishment is preventative in that it prevents the validation of un-right. Confusion arises because the normal understanding of prevention is that it is forward-looking: we punish in order to prevent further breaches of right and here there is an obvious overlap with deterrence theory. Of course, prevention in this forward-looking aspect and deterrence are both desirable things in themselves but Hegel is clear that they are only to be considered once we have established that the punishment is in fact just (Hegel 1967, 70). And punishment is just because failure to punish *retrospectively* validates a breach of right and therefore effectively validates future breaches of right. Bosanquet states that

[t]hat which is to be prevented by punishment is a violation of the State-maintained system of rights by a person who is a party to that system. . . (1951, 212-213)

This is ambiguous: the addition of the word “ongoing” before “violation” would bring it closer to accord with Hegel. As it stands the prevention intended could be against a future violation of right, a past violation of right, or both. That Bosanquet has at least some form of future violation in mind is revealed when he goes on to say that

[i]f a lighter punishment deter as effectively as a heavier, it is wrong to impose the heavier. For the precise aim of State action is the maintenance of rights; and if rights are effectively maintained without the heavier punishment, the aim of the State does not justify its imposition. (1951, 213)

Thus, for Bosanquet, at least part of the “maintenance of rights” is tied up in punishment’s ability to deter crime. This is not true of Hegel for whom punishment is justified solely by retrospective considerations.⁴⁶ Confusion can be avoided by use of stricter terminology; what we are preventing is not, after all, the infringement of right but the *validation* of an infringement of right.

⁴⁶ Any retrospective theory *can* be framed in prospective terms, e.g. we want to bring about a future state of affairs in which the abrogation of right is no longer recognised.

Synthesizing the considerations of Benn and Peters and Bosanquet we can state the following about a Hegelian justification of punishment:

- (i) What is being restored by punishment is not the *idea* of right but the experience of a right.
- (ii) This restoration takes the form of preventing the validation of an infringement of right.

We are closer to an understanding of what Hegel means when he claims that punishment annuls crime but still lacking in a sound exposition of *how* punishment does this. One attempt to explain this is that of David Cooper in his article *Hegel's Theory of Punishment*. His attempt proceeds from an analytical definition of right which stipulates that right is only upheld if the criminal is punished for his offence. Given that most people would agree with the importance of upholding right this would provide a neat but stipulative solution to the problem of why we punish. Cooper states, for example, that

[i]f it is important that men have legal rights, it is important that there be punishment - for without the latter, there could not, logically, be the former. (1971, 163)

It so happens that I think a satisfactory rendering of Hegel will draw us towards such a conclusion. However, the important point is that it will be a conclusion rather than a premise. When we unpack what Cooper himself means by this idea it becomes clear that he is relying on a viciously circular argument; he says earlier that

[i]f a man is not liable to punishment for an action, that is strong reason for supposing he committed no crime, that he infringed no rights. (1971, 162; also at 164)

There are two ways of understanding this statement in the context in which it is made. It may be an appeal to intuition; Cooper says that where "no one would dream of censuring" a purported wrong-doer then this is a good reason for saying that no right was infringed (1971, 162). On the other hand it may be an attempt to define right in such a way that Hegel's ideas follow; a right is that which,

when infringed, gives rise to a legitimate claim for punishment. This may (or may not) be true but in order to establish its truth a considerable amount of argument is necessary. This is simply not provided by Cooper although it must be said that he at least acknowledges this point (1971, 163). One further criticism of Cooper's thesis is that it does not make a strong case for the necessity of *punishment*. For Cooper when a right is infringed *something* is necessary in order to display that right has been infringed; if we do nothing we effectively say that there was no right. This much is in accord with Hegel. But if the something is *merely* there in order to prove that there *was* a right as opposed to somehow *re-establishing* that right then surely anything will do (cf. Cooper 1971, 167 where he states that hypothetically it *is* possible that other practices could replace punishment in restoring right). Not only might we use censure as an alternative to punishment but a circus or a feast. Cooper's thesis appears to be that every time a right is infringed there must be a set response which will be taken as an indicator of the fact that a right has been infringed, which is why I claim that the response could be almost anything. The difference in Hegel is that punishment is necessary not in order to show that right has been breached but in order to rectify that breach. Rectification is something which neither a circus nor (in every case, at least) mere censure can achieve.

For the moment we might add the following to our list of conclusions:

- (iii) Neither stipulation nor intuition are satisfactory ways of showing why punishment is necessary in order to restore right.⁴⁷

Cooper's argument starts with a rejection of H.A. Reyburn's contentions in his work on Hegel's ethical theory (Cooper 1971, 160; Reyburn 1921). That which Cooper rejects is the notion that crime is self-contradictory and therefore demands to be sublated. Cooper fails to see why inconsistency merits punishment; it might well merit a warning, or even psychiatric intervention, but

⁴⁷ NB what is at stake here is not why we restore right but why (or in what sense) punishment can be said to restore right.

why punishment?⁴⁸ The key to this is in understanding that crime is not *merely* an inconsistency; rather it is inconsistent-with-right and is therefore wrong. It demands to be punished, the argument goes, *because* it is wrong. Thus, inconsistency does not lead us directly to the conclusion that a crime ought to be punished (although it is one step along the way); rather it shows us that crime is wrong (i.e. inconsistent with right) a point on which Hegel is clear (1967, 67; also Reyburn 1921, 149).

4 Reconstruction of a Hegelian Justification of Punishment

This reconstruction does not purport to be a totally comprehensive rendering of Hegel's theory of punishment. For one thing I have already rejected a significant part of that theory (*viz.* that the criminal has a right to be punished). Notwithstanding that, I do think that this argument follows one thread of Hegelian thought from start to finish.

4.1 Punishment as a restoration of equilibrium on a metaphysical level

The starting point for a Hegelian justification of punishment is the acknowledgment that it is a highly metaphysical account. Hegel frames his justification of punishment in terms of right and the need to protect the realm of right. There is nothing bad in the fact that the theory is pitched at a metaphysical level - right itself is a metaphysical concept - but it must serve as a caution in developing the report. For one thing it must make one wary, in mathematical parlance, of "confusing the units". If a non-metaphysical step is to be added to the argument (for example, an account in terms of the deterrent effects of punishment) then it must be explained how this fits. Second, because the metaphysical is often best explained in terms of metaphor there is a danger of confusing the reality for the metaphor.

On account of the metaphysical nature of Hegel's justification of punishment it is crucial to draw a strong distinction between a right and the recognition of that

⁴⁸ Furthermore, if crime is already self-contradictory it seems gratuitous to sublimate it.

right. A right *per se* exists exclusively in the metaphysical realm. However, that right defines relationships which exist in the physical world. A crime does not have the power to effect the normative status of a right: it does nothing to modify the truth value of the statement "you ought not to do X". But crime, in a conspicuous manner, strikes at the relationship defined by the right: it breaches the benefit/protection which the right ought to afford. This creates disjunction between the metaphysical and physical worlds since the latter no longer accurately reflects the former. Although the victim still has his right intact, the physical situation is *as if* he did not have that right. It is important to emphasise this distinction in order that I am not interpreted as employing a recognitional concept of right. Thus, a right is a right whether or not anyone acts as if this is the case. But at the same time if people do not act as if this is the case, the right-holder is stripped of any benefit or protection which ought to be afforded by that right. Given that the benefit/protection conferred by a right is the practical moment of the right as an ideal, the *practical* situation will be as if the right does not exist at all. This distinction between right and the recognition of right is echoed in Hegel who discusses the difference between the *principle* of right and the *appearance* of right (1967, 64). ..

I have said that the Hegelian account of punishment is conducted largely on a metaphysical level. It is important to consider then what implications crime has at this level given that crime cannot actually effect the normative impact of rights. A law, as a normative statement, purports to shape individual relationships (which shaping is the recognition of right). But crime also purports to shape individual relationships. Critically, law shapes relationships to reflect [that which is] right, while crime shapes relationships to reflect [that which is] un-right, or wrong. A law which is consistently recognised accurately mirrors the metaphysical realm of right. But a crime which is consistently recognised obscures that realm. It creates a physical world which is *as if* it is based on a metaphysical realm which took that which is in fact wrong to be right. It echoes the occult mantra that fair is foul and foul is fair. Thus, the effect which crime has on the metaphysical realm is that it attempts to subvert it; it attempts to

portray the metaphysical as that which it is not, rendering that which it is as practically irrelevant. Crime is based on a principle which is in direct opposition to the metaphysical principle of right. It is my contention that for Hegel punishment is necessary because the state cannot allow such a situation to occur. It must maintain right.

4.2 The need for the State to do *something* in the face of crime

In the face of crime the state must act to restore right and the realm of right. I have stated that crime shapes relationships in accordance with un-right. A relationship can be shaped either by the mutual choices of the parties or by coercion (either by one of the parties or by an outside agency). Coercion is morally justifiable where it is in accordance with right; this is not a great insight, being in the nature of a tautology. The state - which is coercive in that it implements laws which do not allow of choice on the part of subjects - derives its moral legitimacy from the extent to which it strives to uphold right.⁴⁹ A state which *advocates* crime is morally bankrupt in that it promotes an ongoing situation where relationships are defined by un-right. At a metaphysical level one might say that the principles from which the state derives its legitimacy (right) are incompatible with the principles that individual relationships are shaped by (un-right in this hypothesis). For both Kant and Hegel the state founded on un-right is not just undesirable, it is an unsustainable nonsense; it is an attempt to morally legitimate that which is expressly illegitimate. Central to this is the insight that the state itself is not an unproblematic concept. It is a coercive mechanism and therefore in need of *prima facie* justification. The state is that coercive mechanism which is justified because it upholds right in a way which individual citizens are unable to. Therefore, to talk of a state which lays down laws in accordance with un-right is a contradiction. This is not to ignore the reality of states which do pass immoral laws (almost inevitably, all states do so at one time or another) but it is to conceptually link the idea of the state with moral

⁴⁹ In one sense moral legitimacy is related to the *attempt* to uphold right rather than success in doing so. This is because of the near impossibility of being right all of the time. In the face of this, the most we can require on the part of the state and its officials is a certain level of competence in conjunction with a certain level of *bona fides*.

legitimacy. It is to be noted that crime is a coercive way of shaping relationships. Coercion is to be justified only where it is in accordance with right. Therefore, a state which is founded on crime must claim as being in accordance with right that which is (by definition) un-right.⁵⁰

A state which does nothing about crime (as opposed to openly advocating it) is also morally bankrupt. This is because it tolerates a situation in which relationships are defined by un-right. Such a state does nothing to bring relationships into accord with right and this is incompatible with the state's commission, i.e. that it does attempt to act in accordance with right. What if the state passes and promotes laws but does nothing when those laws are breached? Again we can say that such a state tolerates relationships defined by un-right. That which defines a relationship (right or un-right) is entirely at the whim of the individual. This is the same situation which would exist if there were no state at all, i.e. a situation of anarchy. Whether this is referred to as the golden age, as in Locke's account, or the brutish world which Hobbes famously envisaged, is contingent upon one's degree of optimism or pessimism concerning human nature. The point is that a government which does nothing to define the conduct of relationships between individuals adds nothing to the good, bad or indifferent state of nature. Relationships which are defined by un-right (crime) are met with precisely the same response as relationships defined by right (obedience to the law). The situation is suggestive of a metaphysical realm where right and un-right have an equal status. This situation is incompatible with a state which derives its legitimacy from a metaphysical realm where right alone is to be upheld.

To do nothing accords to crime a normative status which it does not have. This is so because inactivity allows crime to have at least as much influence in defining relationships as the law has. It is a tacit acceptance of a situation where it is as if

⁵⁰ Of course, in practice such a state is unlikely to give two figs about moral legitimacy, or will have its own skewed definition of moral legitimacy which allows crime to proceed. This is not the point though. A state which *is* to be regarded as morally legitimate - and therefore able to justify its title to punish - must make the claims attributed in the text. We are dealing with the notional *ideal* state and considering what is true such a domain.

it were okay that crime shapes people's interaction. The reason for this is that abeyance does not actually alter the relationship between the criminal and the victim whereas the initial crime *has* actually altered their relationship. The ongoing situation would still be *as if* the victim did not have the right, for all the state's hand-wringing. The protection which the right is supposed to afford becomes merely theoretical if the state is not prepared to do anything when its laws are ignored. Again it is to be emphasised that *morally* there is nothing discretionary about the observation of legal duties; but the state in failing to align the legal system with the moral law has created a system where it is as if there is discretion.

Hegel does not make any empirical claims for his model. It is not necessary to demonstrate that a system where nothing is done in response to crime would thus have more crime than a system where something is done.⁵¹ To demand such a demonstration is to miss the point. Hegel does not look to create a system in which crime is effectively controlled but a system in which right is effectively recognised. My claim so far is that a state which does nothing about crime does not recognise right because it allows a system where it is as if individuals do not have rights. But, it might be objected, it is not failure to do anything which creates such a system but the fact that crime occurs in the first place. It is not the state's failure to punish a criminal which puts the victim's right in question so much as the initial breaching of that right by the criminal. In a way this point can be conceded; if the victim's right was not disrupted in the first place then there would be no question of systems where it is as if there are no rights. But this is merely a restatement of the Platonic truism that in a state of saints there would be no need for laws at all.⁵² The point is that we are not a state of saints, people do offend, and the resulting dilemma is what, if anything, do we do about that. Certainly, it is the act of crime which first puts the victim's right in question, but failure to do anything effectively meets that question with a shrug which at worst

⁵¹ Although, as a point of interest, it seems intuitive to say that there would be a higher incidence of crime in the former system. Of course, even this depends on what the *something* done is: if we actively reward crime then common sense suggests that this system would achieve an even higher incidence of crime.

⁵² At least, it seemed a truism to Plato.

is a sign of complicity. If we are to avoid a system where it as if no rights exist it seems we must do something when rights are trampled. The point is to create a system where rights are effectively recognised. This is a state which accurately (or as accurately as possible) reflects the moral reality.

4.3 Punishment as the attempt to uphold right

Punishment is justified *because* it upholds right. The something which the state is to do is aimed at realigning the skewed relationship. It achieves this by ensuring that in practice un-right is ineffective at defining relationships. The key to a Hegelian justification of punishment is the contention that attacking un-right is itself right. It seems to me that Hegel regarded this to be self-evident. He therefore offers little of persuasive value to those who regard the attack on the criminal as itself un-right, being content to describe this view as “superficial” (1967, 70). The counter question would be “why is the attack on the criminal wrong?” This affords at least two instant answers; it is wrong because it is an unnecessary infliction of pain; or it is wrong because it is an infringement of the criminal’s right.

To describe punishment as wrong because it inflicts unnecessary pain is either a sound response or “superficial”, as Hegel put it, according to one’s starting point. As we have already seen in considering Honderich, it is a sound objection *if* one’s overriding principle is the avoidance of suffering. That this is not Hegel’s overriding principle should be clear. In attempting to justify punishment someone who takes pain-avoidance as his Grundnorm is limited to some form of utilitarianism. His units of calculation must be whether one course of action avoids pain more effectively than the contrary course of action; a restatement of this line of thinking forms the basis of theories which justify punishment in terms of its deterrent value or its ability to do good to the offender. Undoubtedly pain is an undesirable experience but crucially, for a rights based thesis, pain may be inflicted legitimately where this is in pursuit of right.

Generally though, individuals have rights not to have pain inflicted on them. Thus, some object that punishment is an infringement of the criminal's right. Exceptionally this will be an absolute objection (such that punishment is never justified) but more commonly it is posed as a problem demanding a further justification for punishment. One way round this would be to claim that the criminal, on account of his crime, has lost his title to right. But we do not need to go this far because in any event the right to avoid suffering is not an absolute right. For instance, the right not to be killed is a more important right than the right not to be assaulted and so I do you no wrong if I push you out of the way of a swinging axe. It is central to a Hegelian account that punishment is necessary to uphold right *per se* and this is held up as more important than the general right to avoid pain.

4.4 Hegel's failure to demonstrate how punishment restores right

It may be conceded that something needs to be done about crime if the state is to avoid the charge of allowing a morally illegitimate situation to prevail. It may even be conceded that, for the reasons just highlighted, the state does the criminal no wrong if it inflicts punishment. However the mere fact that I would do no wrong in committing a certain act does not mean that I ought to do it; at least some issues are morally optional and considerations such as humanitarianism might play a role in settling them. Many therefore look for some purpose which punishment serves before they are willing to allow that it has been justified (for example Hart and his general justifying aim of punishment, 1968, 8-11). It is here that one must remember the warning against "confusing the units". It is poor analysis to conduct a study in the metaphysical terms of right and then make one's justification in the physical terms of (for example) pain and suffering or the psychological units of fear and desire. This would be the case if one were to proceed with an argument similar to that outlined above and then demand that punishment serve the additional purpose of being deterrent before it can be justifiably imposed. There must be a hierarchy of principles in any theory which claims to be justificatory. To allow of ambiguity is to allow that punishment is ultimately an issue of whim. Considering only the

principles of right and deterrence, there must be some way of determining which is the conclusive rationale; otherwise we are given no good way of settling when we ought to punish, which is to say, we are given no good justification of punishment. The method of determination may be a third principle with the result that neither right nor deterrence but this third principle (or the rationale behind it) is the ultimate justification of punishment. All of Hegel's discussion as to why crime creates a situation where something ought to be done has been conducted in terms of right and the fact that a just state is one which attempts to uphold right. Although there is overlap, deterrence is a distinct consideration; to deter is not necessarily to uphold right (as in scapegoating) and to uphold right is not necessarily to deter (as in Kant's hypothetical last execution in the state which is about to dissolve itself⁵³). Crucially, if punishment can be shown to uphold right then it will be justified as an end in itself and require no further purpose or aim (also Bradley 1962, 27). I think, however, that Hegel is merely suggestive of the right answer and that considerably more argument is needed than he offers. As the consideration of the various thinkers above has shown, there is no obviously satisfactory answer to the problem of how punishment is supposed to restore right. As J.N. Findlay puts it "the theory is an impressive arch from which the keystone is missing" (1958, 313).

5 Conclusions

The following points are laid out in the form of statements to show what I think Hegel has successfully established, and where I think his theory is lacking.

5.1: An attack on the particular right of a particular being logically commits the criminal to an attack on right *per se* (or *Recht*).

5.2: The attack on right *per se* stems from the explicit will of the criminal to override the legitimate rights of others.

⁵³ I myself am not yet making the claim that Kant's proposed course of action in this hypothesis is correct.

5.3: Left unchecked such an attack is itself legitimated and hence right *per se* is effectively *delegitimated*.

5.4 Hypothesis: In order to prevent the delegitimation of right we must “attack the attack” i.e. attack the will of the criminal.

I term this as an hypothesis at this stage because I find it the most problematic part of Hegel’s theory (in the sense that I find it the most difficult to understand) and hence I will not accept it as a conclusion until I have tested its soundness further.

5.5: While there is some force in the claim that the criminal must implicitly consented to his punishment the claim that he has *explicitly willed* his punishment is misleading. In any event it is not necessary to show that the criminal has explicitly willed his punishment in order to show that punishment is justified and the focus of this thesis is simply to show that punishment can be justified.

It is my contention that Hegel has successfully established conclusions 5.1, 5.2 and 5.3. Further, it is my contention, having considered what Hegel and Kant have to say on the matter that conclusion 5.5 is sound and that Hegel is wrong to assert that the criminal has explicitly willed his own punishment (or at the very least pursuing an unnecessary line for the purposes of justifying punishment). Finally, it is my dual contention that if Hypothesis 5.4 is discovered to be sound then we have effectively established a retributive justification of punishment, and that Hegel fails to prove the Hypothesis.

Chapter Four: Benefits and Burdens Theory

In this chapter I consider a major contemporary Kantian theory of punishment. Although it would be misleading to suggest that all the writers dealt with in this section see themselves as working from the Kantian tradition or that they all agree with one another as to the fine points of their individual theories, there is a common thread running through each of the discourses. This can be briefly stated as follows: there is an equilibrium, upset by crime, which is restored by punishment. These theories can be described as Kantian being in the same vein as the "hindrance of a hindrance" notion discussed previously. Personally, I think the theories could just as accurately be described as Hegelian although none of the scholars discussed traces himself to Hegel. Indeed, given that many of the theories are expressly couched in terms of rights rather than freedoms it would be fair to say that these theories are *more* accurately described as Hegelian.⁵⁴ The idea involved is the mathematical one of two negatives combining to give a positive. Thus an action which does wrong against a wrong is itself right, whatever generations of parents have told their warring offspring. The difference is the movement from a purely formal construct of what it means to negate the negation of right to a substantial idea of what it is that crime and punishment consecutively reverse. This chapter will be in three main sections. These will deal with the equilibrium which punishment is supposed to uphold; the justification of that equilibrium; and the manner in which punishment is held to maintain the equilibrium. Benefits and burdens theory, like many theories which have come in and out of vogue, is more often criticised than defended in the literature. Therefore, while considering each stage of the argument to the justification of punishment I will deal with the criticisms of the particular stage as developed in the literature. My aim in this is to highlight those parts of the benefits and burdens theories which stand up to criticism and can therefore take us on within the tradition so far traced through Kant and Hegel

⁵⁴ Finnis suggests Thomastic origins for his version of Murphy's theory (Finnis 1972, 134). Burgh objects that the theory is not retributive at all because "desert is tied to the forward-looking consequence of reestablishing an equilibrium" (1987, 319 in footnote to 318). This point is trivial because any theory can be re-expressed in consequentialist terms. What is important is that balance theory demands punishment because it is *just* which makes it an explicitly deontological rather than teleological theory.

1 The Equilibrium which Punishment Maintains

1.1 An equilibrium of the benefits and burdens of social living

The theories I am dealing with in this chapter posit an ideal equilibrium between the benefits and burdens of living in community. Punishment is regarded as one of potentially many mechanisms by which that equilibrium is upheld. A model statement for many of these theories is that of Herbert Morris who talks in terms of punishment which

restores the equilibrium of benefits and burdens by taking from the individual what he owes, that is, exacting the debt. (1971, 79)

Similar statements are to be found through the works of many theorists (Honderich 1976; von Hirsch 1976; Murphy 1979; Finnis 1980; Sadurski 1985; Sher 1987). It is important to outline exactly what the "benefits and burdens" discussed by the theorists are. Here again Morris provides a typical comment:

[a] person who violates the rules has something *others* have - the benefits of the system - but *by renouncing what others have assumed, the burdens of self-restraint, he has acquired an unfair advantage*. Matters are not even until this advantage is in some way erased. (1971, 79; emphasis added)

The benefit of the system is that the agent is offered freedom to pursue his ends unmolested by other agents. The burden of the system is that the agent is required to forbear from molesting other agents pursuing their ends. Put like this the legal system is too simplistic. There is obviously a need for a concept of *legitimate* ends as it is impossible that each agent should be allowed to pursue *any* end without this entailing that at some point he will come into conflict with the ends of other agents. Merely by adopting a particular end an agent can ensure that it is impossible for others to forbear from thwarting him. For example, if A joins a bizarre cult which believes that those with red hair should give themselves up for sacrifice, redheads with an instinct for self-preservation will necessarily offend against A's desire to follow his creed. This is to emphasise that any legal system will need to delineate which goals or which sorts of goals it is legitimate to

pursue. However, given such a system of legitimate ends, placed in a hierarchy with respect to each other so that it can be judged which of two legitimate but conflicting ends is to take preference, it can be seen what sort of benefits and burdens Morris has in mind. The benefit is that legitimate ends can be pursued and the burden is that illegitimate ends (i.e. those which tend to thwart the legitimate ends of others) cannot. The unfair advantage obtained by the criminal is that he retains the benefit of the legal system while renouncing the burden. The criminal is afforded the protection of a community which respects his legitimate ends while trampling the legitimate ends of others. Thus Murphy sees punishment

as *a debt owed to the law-abiding* members of one's community. . . (Murphy 1979, 101; emphasis added)

and says that

[i]f the law is to remain just, it is important to guarantee that those who disobey it will not gain an unfair advantage over those who do obey voluntarily. (1979, 100)

Similarly, Honderich;

penalties imposed must re-establish an equal sharing of welfare and distress with respect to offenders and non-offenders. (1976, 177)

Von Hirsch;

[w]hen someone infringes another's rights, he gains an unfair advantage over all others in the society . . . The punishment - by imposing a counterbalancing disadvantage on the violator - restores the equilibrium. . . (1976, 47; also note 4 at 161)⁵⁵

While all the writers talk in terms of benefits and burdens gained over others in society it is apparent that they do not have exactly the same "unfair advantage" in mind in each case. This point is raised by Richard Burgh who draws a distinction between Morris's notion of unfair advantage and Murphy's notion of unfair profit

⁵⁵ It should be noted that this is von Hirsch's interpretation of Kant rather than the crux of his own theory. But von Hirsch has eclectic taste in moral theory and although he is not a Kantian he does rely partially on this rendering of Kant to justify his theory of punishment (see von Hirsch 1976, 54).

(Burgh 1982, 206-207; also Falls 1987, 27 fn. 5). While Morris is held to look to a sphere of non-interference guaranteed by the law, Murphy looks to the renunciation of the burden of self-restraint. Morris's unfair advantage is the fact that the criminal enjoys non-interference with his rights on the part of others. Murphy's unfair profit is the fact that the criminal gets to exercise a freedom denied to all others. For Morris, the advantage is that C has the general protection of the law afforded to all without paying the necessary price. For Murphy the profit is that C has exercised a freedom over and above the general autonomy protected by the law. This distinction is aspectual highlighting two ways of viewing essentially the same action. Morris as we have already seen talks of the unfair advantage gained by renouncing burdens which others have voluntarily assumed (1971, 79). Murphy claims that

[i]t is important that no man profit from his own criminal wrongdoing, and a certain kind of "profit" (i.e., not bearing the burden of self-restraint) is intrinsic to criminal wrongdoing. (1979, 100)

Thus, for Morris the function of punishment is to remove the benefits of non-interference which the criminal has illicitly enjoyed, while for Murphy punishment's function is to remove the excess freedom which the criminal has illicitly enjoyed. While Morris's theory looks only to the fact that the criminal has experienced less burdens, Murphy's says that in the very act of experiencing less burdens the criminal enjoys an extra benefit which is intrinsic to crime. Sher clearly takes the latter line, holding that

[t]he benefits-and-burdens account regards punishment as justified not merely by the wrongdoer's receiving the benefits of others' self-restraint, but by his having these benefits *plus* the benefit of his own lack of self-restraint. (1987, 80)⁵⁶

Expanding on Murphy's theory, Finnis comments that

what the criminal gains in the act of committing crime . . . is the advantage of indulging a (wrongful) self-preference, of permitting

⁵⁶ It should be noted, for the sake of fidelity, that Sher is defending Morris rather than Murphy against Burgh here. However, it does seem within the distinction made by Burgh that the line Sher is defending should be regarded as Murphy's rather than Morris's.

himself an excessive freedom in choosing. . . (1972, 132; see also Finnis 1980, 263; Duff 1986, 206-207).

Thus Murphy's theory allows for a *double* benefit of (a) not experiencing burdens and (b) experiencing extra freedom. Sadurski on the other hand takes Morris's line when he asserts that the criminal

has acquired some of his victim's benefits and he has renounced some of his own burdens (namely, burdens of self-restraint). The offender has arrogated benefits without bearing the burden of self-restraint and, in consequence, the general balance of benefits and burdens has been upset. (1985, 226)

Although Sadurski does talk in terms of specific benefits which the criminal has acquired as a consequence of his crime these are the incidental benefits he has acquired from his victim such as the victim's money. There is no notion of Murphy's *intrinsic* benefit of acting freely.

As the distinction is aspectual it is not a crucial one for much of the following discussion. The essential features of an equilibrium between the benefits of the legal system and its concomitant burdens are the same. For what it is worth I think that Morris is closer to the truth on this particular issue. What is wrong with crime is not the exercise of freedom *per se* (as in Murphy) but the failure to 'pay the price' (i.e. to restrain oneself from infringing the rights of another). Ultimately though I will contend that even Morris's notion of crime is lacking. The wrong in crime is not *merely* that the benefits of the legal system are unpaid for. If this were the case punishment would be gratuitous because we do not restore the pre-crime equilibrium; we do not make C assume the burdens which he originally failed to assume, we make him assume different ones. For the time being though, we have the notions of an equilibrium of benefits and burdens and an unfair advantage/profit which is accrued over the law-abiding members of society when the criminal fails to adopt the burdens.

1.1.1 Rights and the benefits of having rights

The question must be as to exactly what the law-abiding members of society are supposed to have lost in the commission of a crime. Again I will turn to Morris in the first instance and his notion of rights. This is best explained in terms of a distinction between a right and the benefit of that right. Without such a distinction many of his assertions appear contrary and confusing. For example, he states that the

right to be treated as a person is a fundamental human right belonging to all human beings by virtue of their being human. (1971, 92)

He adds that this particular right

is a right we have *apart from any voluntary agreement* into which we have entered. (1971, 94; my emphasis)

However, Morris also asserts that

rights exist when there are recognized rules establishing the legitimacy of some acts and ruling out others. (1971, 99)

On first inspection, the last of these statements appears to contradict the other two. How can a right be fundamental, existing by virtue of one's humanity and simultaneously exist only where there are recognized rules to back it up? The point is that a human right of the type Morris initially identifies exists independently of any rules and, indeed, exists whether or not anyone actually recognizes its existence. To have it otherwise would be precisely to make rights conditional upon some sort of voluntary agreement (recognition, as in "recognized rules", being a necessarily voluntary action). This is something which Morris has already denied. The way round this, I believe, is to hold that in his first two statements Morris is referring to rights properly understood, and in his third statement he is referring not to rights, but rather to the benefits of, or the enjoyment of, those rights. It is not, of course, *strictly* true to say that the enjoyment of a right is guaranteed by having rules prohibiting the right's

infringement. But given that Morris talks in terms of rules which are "recognized" this is a plausible interpretation. That is to say, where a rule which protects a right is recognised then the enjoyment of that right will be guaranteed, to the extent that such a guarantee is ever possible. On this interpretation, Morris is holding that while rights are fundamental and inalienable, the benefits of enjoying those rights (and the burdens of non-interference with the enjoyment of those rights by others) are the subject of distributive justice. Where one has cast off the burden of self-restraint then punishment restores the distributively just balance by imposing an equivalent burden (of enforced restraint). This links Morris with Murphy notwithstanding the distinction drawn by Burgh. The benefit of a right is the freedom to act (in accordance with that right).⁵⁷

I will argue that the theorists fail to show a meaningful sense in which punishment can be said to restore the benefit of a right. I hope to make up for this shortcoming in the final section of this thesis. While there is a conceptual distinction between a right and the benefit of enjoying a right, it is not immediately obvious how the two can actually be separated for the purposes of distribution. For the time being though we can see that this notion - of benefits and burdens (whether of rights or freedoms) as the fitting subjects of distributive justice - is repeated throughout the theories.

Murphy says that

[c]riminal punishment . . . attempts to maintain the proper balance between benefit and obedience by insuring that there is no profit in criminal wrongdoing. (Murphy 1979, 77)

Sadurski discussing the "ideal, hypothetical balance" states that

[i]t is an equilibrium in the sense that, in the situation of full respect for each person's sphere of autonomy, all enjoy equally the benefits of autonomy and the burdens of self-restraint. (1985, 104)

⁵⁷ Significantly the converse is not necessarily true: the freedom to act is not necessarily the benefit of a right. Some freedoms/acts are wrong. This confusion, a major weakness in the theories, is due to the lack of a strong or grounded notion of right. In particular it impinges on Murphy's version of the theory; if rights are important it is untenable to hold that freedom should be protected no matter what.

Finnis asks

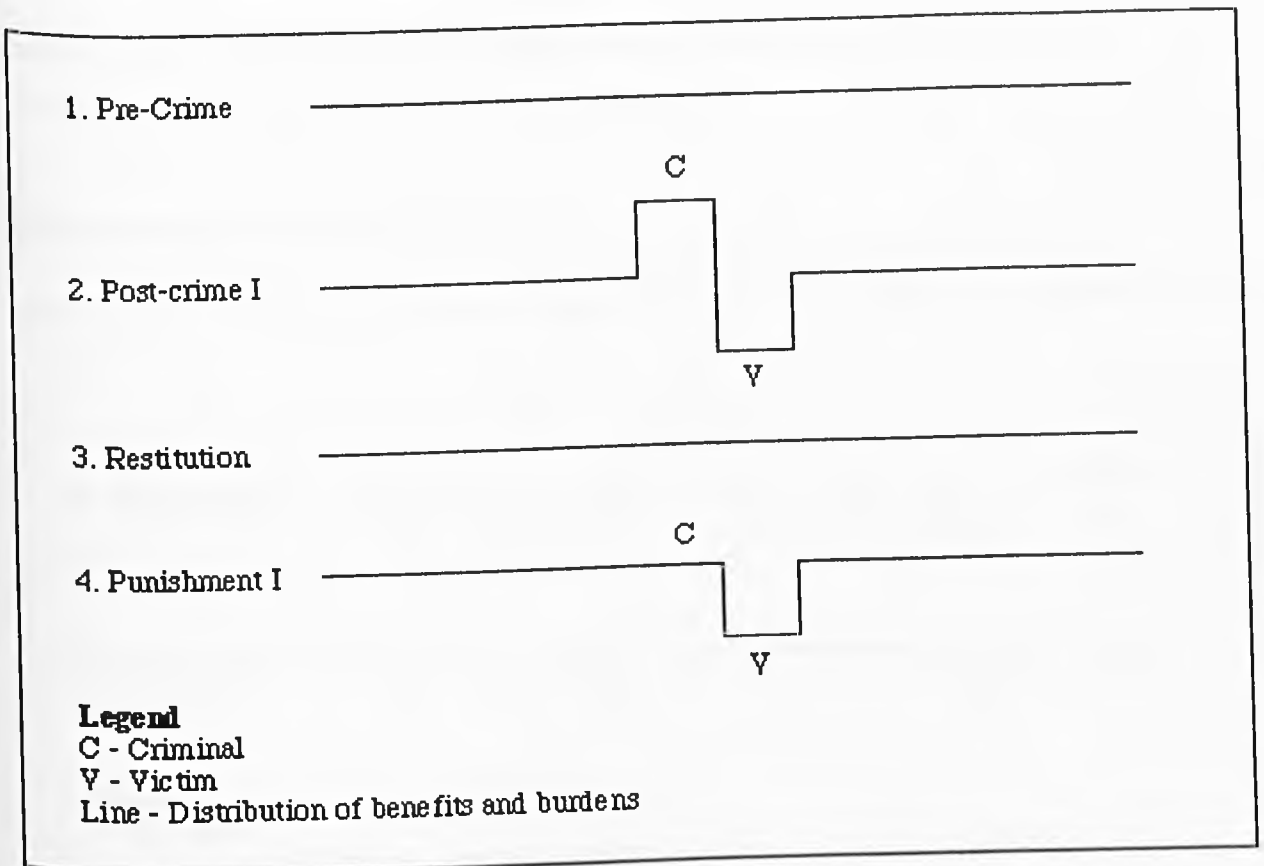
is not the exercise of freedom of choice in itself a great human good? . . . Punishment, then, characteristically seeks to restore the distributively just balance of advantages between the criminal and the law-abiding. . . (1980, 263).

Again the subject of distributive justice is conceived in terms of the benefit of exercising free choice and its converse, the burden of self-restraint. The currency of distributive justice in the realm of punishment, then, is the benefits of having rights and the concomitant burdens of observing duties.

1.2 An equilibrium spread across the members of society

Where punishment is imposed so as to restore balance this is held to return the status of the criminal, C, to the same level as before the crime. However, the status of the victim is still reduced. Even if restitution has secured the return of any material loss, V has still experienced the crime (an experience he ought not to have undergone) and suffered any consequent psychological "loss" (cf. Wasserstrom 1980, 145).

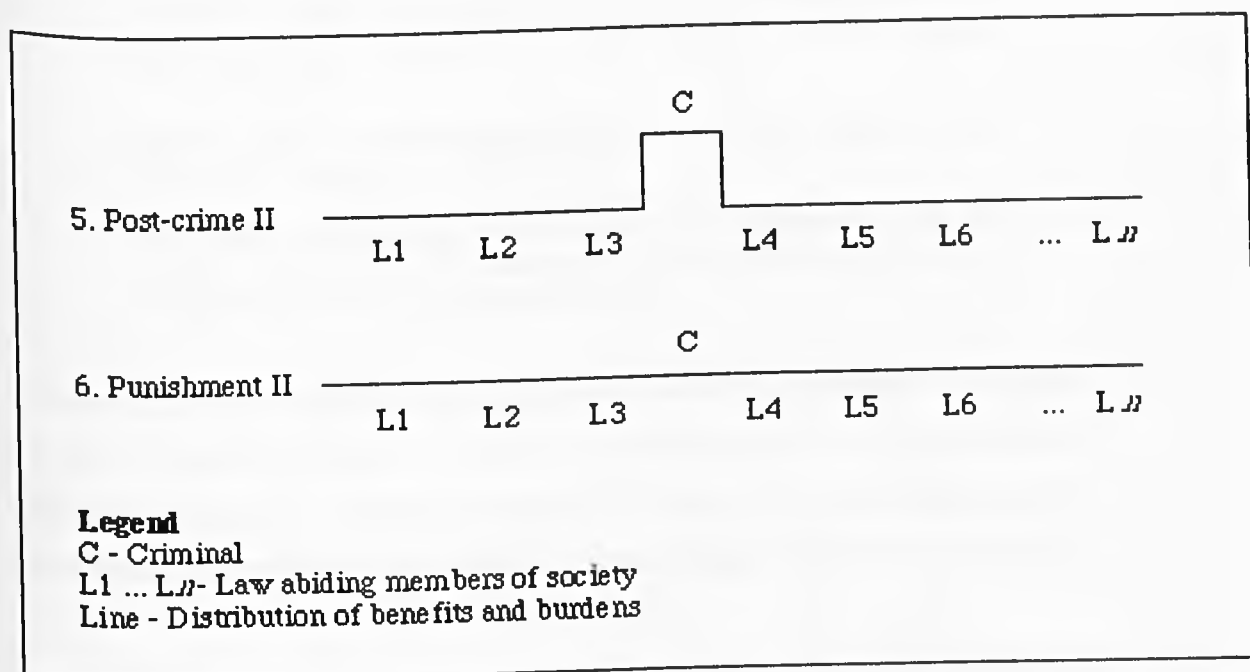
Diagrammatically this may be represented as shown on the following page:



It is seen as impossible to restore the victim to his original position (1) in anything other than purely material terms (3) and even then not all crimes lend themselves to material restitution. The best that could be offered in material terms for crimes such as assault is *compensation* and this hardly returns V to the pre-crime position. It is perhaps because of the perceived impossibility of undoing what has been done to V (4) that most balance theorists write the immediate victim of the offence out of the punitive equation. Instead the wrong is conceived in terms of the loss experienced by society. The attitude towards V is apparently summed up by Finnis' aphorism that "[w]hat is done cannot be undone" (1980, 263; see also Nozick 1981, 379; Sadurski 1985, 227 who say substantially the same thing). The law abiding members of society (represented by L1, L2 . . . Ln) lose out *relative* to C because by taking the opportunity to offend which others have denied themselves C gains an unfair advantage over those others (5, overleaf). In the language of the theorists, he has taken the benefits of others' forbearance without accepting the reciprocal burdens of self-restraint. Because the loss of L1 etc. is a

relative loss the equilibrium can be properly restored simply by imposing a burden on C equivalent to the benefit he has taken to himself (6).

Diagrammatically this is shown as follows:



In these diagrams the post-crime position shows that L1 etc. have not lost out in real terms; their position is only diminished relative to C who has taken an advantage for himself. Therefore, simply by removing that advantage - by imposing a countervailing disadvantage - the position of C is returned to equality with that of L1 etc.

The advantage is conceived in terms of what C has gained over the law abiding. This is unorthodox as it is more common to conceive of crime as an advantage gained over the immediate victim of the offence. Of all the theorists Wojciech Sadurski is closest to conceiving the upset equilibrium as existing not between the offender and society but between the offender and the victim. He is clear that

the harm inflicted on another person by an offender constitutes the infringement of that other person's rights to liberty, life, property etc. (1985, 226)

However, it is nevertheless true that Sadurski's justification of punishment looks to an equilibrium between the offender and *society*. The following three quotes are instructive here;

[t]he fundamental benefit that a criminal acquires from violating the rules of criminal law is a benefit of non-self-restraint, that is to say, a benefit of freedom from burdens imposed by the criminal law. (1985, 229)

Criminal justice is concerned with the distribution of rights among members of a society. . . (1985, 225)

The general justification of punishment is analogous to that of rewards: it is a method of restoring an *overall* balance of benefits and burdens. (1985, 225; emphasis added)

The message here seems to be that (i) the illicit benefit is a freedom from restraint of the criminal law; that (ii) the criminal law ensures to exist a fair distribution of benefits and burdens amongst members of society; and that therefore (iii) punishment is *justified* because it restores the net balance disturbed by the crime.

I have shown that the equilibrium which the theorists have in mind exists between the benefits of having rights and the burdens of observing duties (or in some cases the benefits of exercising freedom and the burden of being restrained). I have further shown that the equilibrium exists across the various members of society. I will now consider the criticisms of this stage of the argument developed in the literature.

1.3 Criticisms

1.3.1 The notion of equilibrium is metaphorical

In *Censure and Sanctions* Andrew von Hirsch states that

[i]t is arguable (although still debatable) that the offender, by benefiting from others' self-restraint, has a reciprocal obligation to restrain himself. It is much more obscure, however, to assert that if he disregards that obligation and does offend - the unfair advantage he supposedly thereby gains can somehow (in other than a purely metaphorical sense) be eliminated or cancelled by punishing him. (1993, 7-8)

In a similar vein, discussing the issue of someone who through lack of qualifications is forced to take a demeaning or undesirable job, Sadurski says that "it is not 'equality' in any other than a metaphorical sense that economic compensation brings him back to" (1985, 102). The crux of this argument is that when two incommensurable things (in this case being demeaned and being paid) are weighed a finding of equivalence between the two is necessarily arbitrary. In the school room parlance we have already encountered, this criticism points to confusing the units (see also Wasserstrom 1980, 145).

It is to be conceded that in one sense the theory is metaphorical. There is no actual equilibrium just as it is frequently pointed out there is no actual social contract. However, this is not to say that there can be no ideal notion of equilibrium against which disequilibrium is measured.⁵⁸ If such an ideal notion can be *rationaly justified* then the admission that equilibrium is metaphorical in nature does not entail the admission that equilibrium is necessarily conventional. Furthermore, provided that there is sound argument as to why a special sort of privation is to be imposed in response to crime the theory is saved from the charge of confusing the units. The crux of the issue is whether sound argument is made as to why punishment is the correct response to crime, not the fact that on the face of it punishment appears to be a qualitatively different action to crime. On this point Galligan distinguishes between two versions of the balance theory. The first "strong" version, which he holds to be untenable, is that punishment wipes out the advantage obtained by the criminal. The second "weak" version, which he supports, is that punishment wipes out the unfairness of that advantage. In the first case we can see why punishment and crime are said to be unmatched: nothing in the punishment can remove the actual advantage which the criminal enjoyed in expressing his freedom. In the second case however, all that is required is that the disadvantage imposed by way of punishment is *equivalent* to the advantage gained so that over the period of time including the crime and its punishment the criminal "is in no better position than those who did not offend" (Galligan 1981, 157). This

⁵⁸ see Kant (1991, 124-125) for the parallel argument that the social contract is an ideal against which actually existing societies are to be judged.

is a valid defence but even so I would contend that the theory's real deficiency lies in its failure to fully explicate what is at stake in a crime. It is this deficiency which leaves it open to the accusation of being metaphorical. Galligan is right to say that the best the theory can hope for is to remove the unfairness of the crime rather than the crime itself. Unless the wrong perpetrated against the law abiding is properly delineated the notion of punishment undoing this wrong, if not metaphorical is certainly tenuous (see further Walker 1991, 73). Is it, as the next criticism of the theory suggests, merely sour grapes to punish the criminal if all we achieve is to ensure that he has had no more freedom to act than anyone else? Significantly, only where there is some concept of an *ongoing* wrong done against the members of society can the wrong be undone in a real (i.e. non-metaphorical sense). It is too late to *undo* a wrong that is entirely in the past. The real problem with the theory is not so much that the notion of equilibrium is metaphorical as that the things which the theorists claim to be the subjects of equilibrium are not the fitting subjects of social distribution. In my own reconstruction of the theory I will suggest an ongoing wrong which is actually undone by the act of punishment.

1.3.2 Balance theory has a distorted view of crime

A further objection to benefits and burdens theory is that it presents a distorted view of crime. In particular it necessitates that the opportunity to rape or murder is an advantage or profit (Falls 1987, 31). Duff makes the point that

the idea that the wicked profit unfairly by their wrong-doing reflects a grudging and less than whole-hearted commitment to the values which they flout... (1986, 214; also Wasserstrom 1980, 143-144; von Hirsch 1986, 58)

Duff's objection, which is valid in its own right, is to the sort of sentiments which suggest that really we would all like to rape and murder and it would not be fair to let just a few partake in these pursuits. Sadurski's response to this is to say that even if one does not perceive self-restraint as burdensome it is nevertheless a burden because "criminal law restricts an absolute freedom to do as one pleases" (1985, 228). This ties in closely with Murphy's point that the unfair profit derived from crime is the opportunity to act in a way from which others have restrained

themselves. But it must be questioned whether this sort of "absolute freedom" is in fact a good thing; certainly the theorists must contend that it is a good thing if *profit* or *advantage* are to be shown (see Finnis 1980, 263). Pertinently Falls states that according to at least one viewpoint "in doing evil one harms oneself rather than profits" (1987, 30). This criticism echoes (ironically given the number of balance theorists who claim to be Kantian) the assertion that "whatever undeserved evil you inflict upon another within the people, that you inflict upon yourself" (Kant 1991, 141). A core problem here relates to the nature of free-will as discussed in the introduction to this thesis. It is not good to act totally freely in the sense of randomly. What *can* be described as good is the ability to control one's actions. However, it is still not the case that this is an *unqualified* good. If there are rationally binding moral laws which demand that the individual does not act in certain ways then it is good if he does not act in those ways and, logically, bad if he does act in those ways. Insofar as the objection is to a misrepresentation of crime as being a "good thing" there are two things to be said. First, a begrudging attitude which holds that the law-abiding would all like to offend but choose not to, and that the function of punishment is to stop the law-abiding from becoming disillusioned, is misplaced. However, secondly, I do not think that the theory necessitates such a conclusion. All that it entails is that the *criminal* view his crime as good. This does not mean he sees it as *morally* good - he may view it as expressly wicked - but he sees it as at least good enough to cause him to act upon it. Under the balance theory, punishment is committed to removing a wrongly exercised proportion of the generic good of being able to pursue one's own ends. The criminal realises a greater portion of this good than the law-abiding member of society because he refuses to acknowledge the cases in which refraining from pursuing his ends is the *morally* right thing to do. On balance though I would say that the argument raised by Duff, Falls, Wasserstrom and von Hirsch in this context holds good for the reason that the benefits and burdens theories themselves are couched in terms of an advantage taken over society. As expressed in the theories the law abiding have not particularly lost out. Alternatively the law abiding have only lost out in a relative sense in which case it is correct to object that punishment constitutes a begrudging attitude towards the

value of right. Moreover, it is difficult to see exactly what is restored to the law abiding by punishing the criminal. In this sense it is right to say that the notion of crime which the theorists have is unsatisfactory because it does not allow them to prove the necessary point that equilibrium is restored by punishing the criminal.

1.3.3 Balance theory does not account for the victim

Another major criticism which can be levelled at balance theory, and one which I hope to meet as my own theory develops, is that it has no place for the victim in its equations. We have seen that this is not entirely true of all the theories (Sadurski 1985, 226) but as a class the theories do fall foul of this objection. We have already seen why this should be so: a theory which conceives of just distributions as existing between the offender on the one hand and law-abiding members of society on the other, will accord no special status to the victim. Obviously none of the theorists would deny the *harm* done to an individual who suffers at the hands of a criminal, nor would they deny that a *wrong* is done to that individual. However, the *crime which punishment is supposed to address* is never regarded as a crime against society and the victim is only of relevance in so far as he is a member of the particular society. Thus Murphy writes,

[i]f a man does profit from his own wrongdoing, from his disobedience, this is unfair or unjust, *not just to his victim, but to all those who have been obedient.* (Murphy 1979, 84; emphasis changed from original)

To accentuate that this was in fact Murphy's position at the time of writing *Marxism and Retribution* we need only look to his own statement when he comes to recant his earlier work:

I have persuaded myself (and have been persuaded by others) that it is morally implausible to think that the primary evil of a crime such as murder is that it involves the taking of an unfair advantage. . . (Murphy 1992, 47-48)

One of the others is Duff who in *Trials and Punishment* acutely pin-points the fault with the balance metaphor in this respect:

what is wrong with rape is that it attacks another person's interests and integrity, not that it takes unfair advantage of the law-abiding. (Duff 1986, 212)

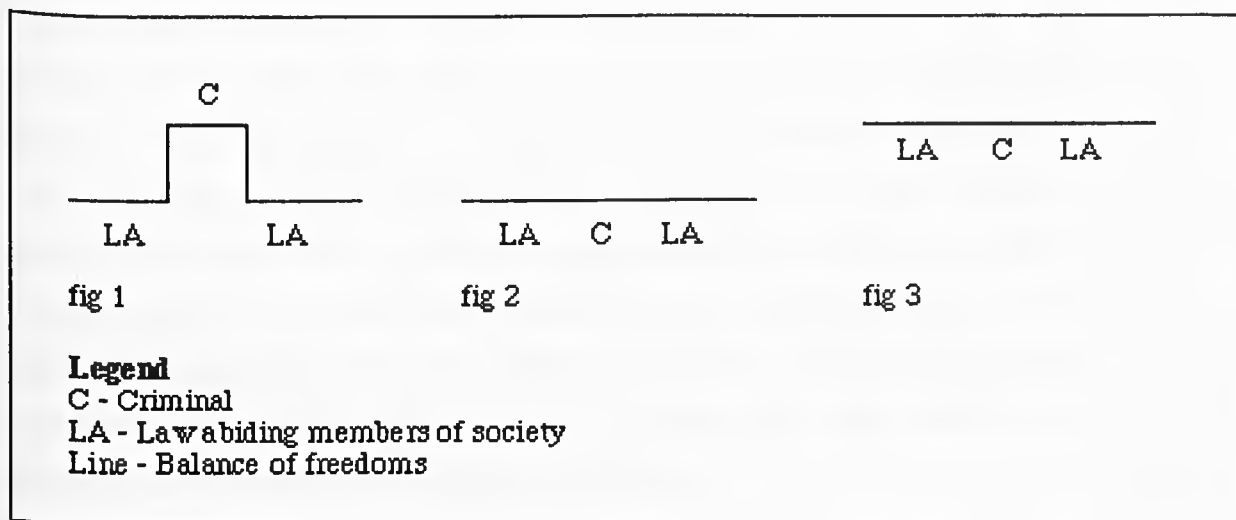
It must be remembered that it is freedom from constraint which the punishment is designed to rectify by imposing a balancing infringement of freedom to restore equilibrium. Finnis writes that

when someone, who really could have chosen otherwise, manifests in action a preference . . . for his own interests, his own freedom of choice and action, as against the common interests and the legally defined common way-of-action, then in and by that very action he gains a certain sort of advantage over those who have restrained themselves . . . Punishment, then, characteristically seeks to restore the distributively just balance of advantages between the criminal and the law-abiding, so that, over the span of time which extends from before the crime until after the punishment, no one should actually have been disadvantaged - in respect of *this* special but very real sort of advantage - by choosing to remain within the confines of the law. (1980, 262-263; see also Finnis 1972, 133; cf. von Hirsch 1986, 58 who is a critic of this line)

However, if *all* that is wrong with crime is that it gives the offender a greater number of freedoms (represented in fig 1, see overleaf) than the law abiding then it would make as much sense to let all others offend as it would to punish the individual offender. This would restore the balance just as much as punishing the offender. George Sher raises a similar objection when he asks the question "[i]f what is important is only the final balance, then why does it matter in what order the debits and credits are entered?" (1987, 83).⁵⁹ Indeed if the maximisation of freedom is in itself a good thing, then it would be *better* to address the situation in this manner. Thus in fig 3, where instead of punishing, everyone is allowed to offend (which as a political agenda leads to anarchy with the manifesto slogan "might is right") the equilibrium of freedoms is restored just as it is in fig 2, which represents the balance restored after punishment. But in fig 3 the equilibrium is restored, as it were, at a higher level, so that to use Finnis's own terminology "over the span of time which extends from before the crime until after the

⁵⁹ Sher's actual point here is against the Marxist critique (see section 1.3.4 of this chapter): if someone has been previously disadvantaged does this allow him a "free crime" to offset the disadvantage? However, the point holds for present purposes because it shows that balance theory is only concerned with the equilibrium of benefit and burden being maintained without especial reference to how this is achieved.

punishment, no one should actually have been disadvantaged” and everyone should actually have been *advantaged* (1980, 263).



A possible objection to this idea would be that the point is not the restoration of *any* equilibrium but the restoration of the equilibrium that existed before the crime. However, this fails because as we have seen the theorists are expressly not trying to restore equilibrium as concerns, for example, the victim of the offence (see also Sadurski 1986, 227). In this case it could be argued that the situation represented in [fig 3](#) is at least as desirable as that in [fig 2](#).

But of course the theorists are not willing to allow that the two situations, punishment and anarchy, are equivalent. Finnis for one is clear that human rights are important and so he can not countenance a situation where in the name of freedom all are free to offend (e.g. Finnis 1980, 226). It would not be a good thing if human rights were continually trampled. But this only goes to show that the wrong in an offence is the taking away of rights and not the relative diminution of freedom in the law abiding community. If we are going to hold that human rights are of ultimate importance then we cannot consistently hold that “[f]reedom itself is the only value which can be used to limit freedom” (Murphy 1979, 97). Either freedom or rights are the foundational concept and unless we are to hold them identical we cannot have it that both are foundational. To merely

talk in terms of right is insufficient if "right" is no more than a simile for "freedom".

It is here that the theorists fall back on the aphorism that "[w]hat is done cannot be undone" (Finnis 1980, 263). But this is an unsatisfactory way of disposing of the issue. The attitude seems to be that we must do something in response to crime, that nothing we can do will undo the crime against the victim and that, therefore, we should look to a "crime" done against the law-abiding, constituted in such a way (i.e. in relative terms) that we can do something about it. My argument is that within the terms of reference used by the theorists (i.e. the maximisation of freedom) it would make as much sense, and maybe more, to do nothing about the crime and let everyone get on with it.

It seems to me that the theorists turn to the law-abiding as the losers in crime because of the apparently irrefutable statement that with respect to the victim what is done cannot be undone. I now wish to scrutinize this assertion and suggest that it merits rather more thought than it is generally given in the literature. In the ultimate material offence - theft - what is done *in material terms* can be undone because the stolen goods can always be returned. So it is not necessarily true to say that "what is done cannot be undone" even in material terms. What *cannot* be undone is two fold. First, nothing can remove from the victim the experience of undergoing the offence. Second, nothing can make it so that the criminal did not actually offend; nothing can force him to go back and choose not to break into the house etc. But, and this is the crucial thing, these are both as true when considering the law-abiding/criminal relationship as when considering the victim/criminal relationship. Nothing can remove from the law-abiding the fact that they underwent a process of having their freedoms relatively diminished and (if this is what we are to view as the crime) nothing can make the criminal go back and choose not to diminish their freedoms. There remains one unconsidered factor; can rights that have been trampled be restored in any meaningful sense by the process of punishment? Although, as we have seen, Morris and others talk in terms of the benefit of having a right, I believe that the

issue is fudged when it comes to consideration of what punishment actually restores. Your experiencing the benefit of a right is necessarily linked with my observance of the burden of not infringing that right. But if I fail to observe that burden, the imposition of a *post facto* compensatory burden does not restore *to you* the benefit of having your right untrammelled. There is little consideration given to whether the violation of the victim's rights can be undone because it is assumed (wrongly as it turns out in some cases) that the material offence cannot be undone. For example, Murphy states that

undergoing punishment for (say) murder, unlike paying the final installment on a loan, can hardly be said to make things all right again, to make the world morally the same as it was before. (1979, 78)

But this raises the question *why* is the world not "morally the same as it was before"? If before a crime there are a certain set of rights and duties and after the crime there are a certain set of rights a duties then perhaps the world *is* morally restored. Certainly this argument is not refuted by stating the opposite case. In order to make a case that the world can be morally restored it must be argued that there is an ongoing infringement of right which is stopped by the act of punishment. In the conclusions to this thesis I will identify a rights-based interest vested in the victim and the law abiding (and, indeed, the criminal) which is restored by the act of punishment. For now I concur with Duff and with Murphy in his recent criticism of his own theory, that the benefits and burdens theorists are hasty in rejecting the direct victim of the offence as relevant to the punitive equation. This criticism is linked to the previous one. The theorists tend to misconstrue both the nature of the crime committed and the identity of the persons against whom it is committed.

To summarize I see four possible things that theorists may be talking about when they refer to "what is done...": the experience of the offence; the criminal's actions; the material offence; the violation of rights. With regard to the first two the phrase "...cannot be undone" is equally true of the law abiding and the direct victim of the offence. With regard to the third, the material offence against the

victim can sometimes be undone (theft, fraud, damage to property) and sometimes cannot (murder, rape, assault). There is no *material* offence against the law-abiding as such; however, the theorists imagine one in terms of having their relative freedoms diminished or in terms of their being wronged by the criminal experiencing a lesser burden of restraint than them. This is what the balance theorists claim can be undone. With regard to the fourth thing, the violation of rights, this seems to be an unconsidered possibility.

1.3.4 The Marxist critique

The Marxist critique is probably the objection most frequently levelled at balance theory. Indeed, as a criticism of that theory it originates from one of the theorist's himself, Jeffrie Murphy. It is not an immanent critique of the balance theory but a contention that the theory is inapplicable to actually constituted societies. It states that whatever else might be true about punishment, our present society (because of class inequalities) is not the sort of society in which punishment could be justified. In *Marxism and Retribution* Jeffrie Murphy offers a Kantian theory of punishment that he holds to be the best possible of all such theories. However, he then offers a Marxist analysis of society which seeks to show that in our present unfair system even Kantian punishment is unjustified (1979, 103, 104). This is because of "the empirical falsity of the factual presuppositions of the [Kantian] theory" (1979, 104). The central "empirical falsity" which Murphy identifies in the theory is that the rules of society benefit all members equally (1979, 107; Sadurski 1985, 232; Duff 1986, 209; von Hirsch 1986, 58). The objection proceeds thus. The benefits of living in a modern society can only be secured if people restrain themselves from trampling over the benefits when enjoyed by others. Therefore, in return for enjoying the benefits each must adopt the burden of respecting the benefits of others. But in practice some people do not enjoy the benefits of modern society; this "underclass" is the loser in a capitalist system. Consequently, they do not owe the debt of obedience which could otherwise be legitimately expected of them. Therefore, punishment is unjustified (in these cases, at least). Although the Marxist critique is first posited as a Marxist critique by Murphy, this

line of thought precedes the formulation of benefits and burdens theory. W.D.

Ross wrote in 1930 that

the criminals that a retributive theory of state punishment would call on us to punish . . . may well be persons who are more sinned against than sinning, and may be, quite apart from our intervention, already enjoying less happiness than a perfectly fair distribution would allow them. (1930, 59)

It is to be noted that according to this line of criticism punishment is justified *in principle* but is restricted to a society in which the distribution of benefits and burdens is *prima facie* fair. The Marxist critique is particularly instructive because it is levelled by Murphy at his own theory. When he says that the presuppositions upon which the balance theory is based are flawed we get an insight into what he considers the presuppositions of his own theory to be. Many criticisms may be dismissed as an inaccurate or unsympathetic reading of the original but not this particular example.

In its crudest form the Marxist critique would have to state that anyone who has ever been wronged is exempted from culpability (to the extent that they have been wronged). Thus, the issue of ultimate importance is that over all, and by whatever means, the benefits and burdens of the system should be balanced for each agent. This process is one of unsophisticated moral accountancy. George Sher defends the benefits and burdens theory against this crude form of the Marxist critique holding that the burden being offset against the benefit of crime is in fact of a different nature to the crime and therefore incommensurable with it. He says that

if the wrongdoer's extra benefit is measured by his act's degree of wrongness, whereas his previous burden is measured on the different scale of preference-(dis)satisfaction, then there is indeed a convincing reason why the later benefit cannot be balanced by the earlier burden. (Sher 1987, 84).

It is to be noted though that this defence only works for Sher's developed version of the theory. For Murphy, and for his theory as expanded by Finnis, the extra benefit lies precisely in the preference satisfaction of expressing freedom. It might be said then that the basic version of the Marxist critique only works against the

basic version of the original theory. I have already discussed whether it is correct to regard the expression of freedom as an unqualified good. It seems necessary to take the view that it is, if it is to follow that it is of ultimate importance that everyone have an exactly equal amount of freedom. When something distinct from freedom (such as right) is held up as the highest end it ceases to follow that whoever has been granted fewer freedoms is allowed to take unto himself enough freedom to restore equality with others.

Sher goes on to consider a more refined version of the Marxist critique in which the criminal may be excused his crime where it is commensurable with some identifiable wrong done to him in the past (1987, 85-86). The important notion here is that of commensurability central to which are the nature of the action taken, the direction of the action and the seriousness of the action. It is hardly difficult to imagine a regime that makes unjust demands of its subjects and I will take slavery as an example. It is an uncontentious conclusion of Kant's categorical imperative (and most other moral theories) that slavery is wrong. A state may nevertheless - talking here in purely positive terms - pass a law requiring certain people to become slaves and making it a crime to refuse. But any person refusing to render himself up for slavery ought not to be punished.⁶⁰ This is because the nature of his action is the resistance of a wrong done against him; and the direction of his action is against the one who is doing the wrong, i.e. the state; and his action in refusing to give himself up is not disproportionate to the wrong done against him. Insofar as the Marxist critique shows this it is acceptable. However, we do not actually require the Marxist critique in order to reach the conclusion that resistance to slavery is legitimate. All of this can be thought of in terms of self-defence. I have a right not to be made a slave and therefore anyone attempting to make me a slave commits a crime against me. I commit no wrong in defending myself against crime. The self-defence caveat of proportionality ensures that only relevant actions against relevant parties are taken (see also Sadurski 1985, 232). When it is put like this we can see why the actions ought not to be

⁶⁰ Which is to say that punishment would be morally unjustified but is not to ignore the fact that such an appeal is unlikely to persuade a state that has had the moral temerity to institute slavery in the first place.

referred to as crimes in the first place; self-defence is not a wrong. The positivism of many of the balance theories is unhelpful here. A proper notion of what constitutes a law (and hence what constitutes a crime) would show that posited rule R that X give himself up for slavery is no law hence a breach of it is no crime. In turn it follows that intervention against one who commits the posited crime cannot be justified as an intervention against *an offender for an offence*.⁶¹

Slavery is an easy example because it is an obvious wrong and it is such a serious wrong that almost any resistance of it can be justified. In other instances, for example, where the state wrongly withholds a monetary benefit to which an individual is entitled, it may be much more difficult to say exactly which actions are justified, but the principle holds. Thus, someone who is wrongly denied housing benefit does not have *carte blanche* to commit rape. Rape is not the sort of action which resists the wrongful withholding of housing benefit, nor, being aimed at an individual as opposed to the state, is it aimed at the relevant party. If we change the crime to murder, which may be aimed at an official of the state and may be in pursuit of the wrongly denied benefit, we still identify incommensurability between the crime and the burden that the criminal has suffered because murder is wholly disproportionate to the burden of poverty. Within the parlance of freedom that many of the theorists employ, this conclusion follows because the freedom-limitation of being killed is much greater than the freedom-limitation of being poor. This shows that some actions are patently unjustifiable even in the face of an illegitimately imposed burden but it does not show which actions are justifiable in a particular case. Ultimately a theory which holds that certain actions are justifiable but cannot specify which will be rightly condemned. However, it is not my current purpose to pursue this line any further. What I seek to show here is that the principle of justified resistance holds whatever the wrong done against the individual.

⁶¹ This is not an instance of the definitional stop. The definitional stop would contend that this is not a case of justified punishment because it cannot be defined as *punishment*. I am contending that it is not a case of justified punishment because the immorality of the purported law precludes that any action taken in defence of that law could be *justified*.

The Marxist critique lacks the subtlety of this self-defence based model. Focusing, as it does, on a net balance of benefits and burdens, it says that a person denied any right is excused his offences (to the extent that he is denied rights). The rough equation is as follows

$$iR - aR = OA$$

In this equation iR is the ideal balance of benefits and burdens, aR represents the actual balance of benefits and burdens experienced by the individual, and OA represents the resultant "offence allowance". Thus, the more rights an individual is denied the greater offence(s) he can get away with. Where the offence committed (OC) exceeds OA then the punishment is proportional to $OC - OA$.

The self-defence model does not really engage in such thinking because, as we have seen, it does not regard self-defence to be a wrong. However, for the sake of comparison we may write it thus

$$R1 - R2 = C$$

Where $R1$ is the magnitude of the rights which a person is wrongly denied, $R2$ is the magnitude of the "rights"⁶² which he infringes in self-defence, and C is his culpability. So long as C is greater than zero the person ought not to be punished. It is important to reiterate that $R2$ represents the magnitude of the "rights" which he infringes *in self-defence* because obviously certain actions under certain circumstances could not be typified as self-defence. The important criteria again are the type of action, the direction of the action and the seriousness of the action.

What Murphy in his Marxist critique has failed to see is that there are some things which would constitute crime no matter how unjust the society in question is.

⁶² I use inverted commas because in a developed theory of self-defence it is probably the case that no right is infringed. Thus, if I attack you, I cede my right, to a certain extent, to personal integrity, and you do not infringe any right of mine when you defend yourself against me. The term "right" can be used with the qualification of inverted commas though because there is a *prima facie* right, for instance, not to be assaulted.

Thus, where the oppressed person takes something or some freedom from another, and that thing or freedom is not one to which the person would be entitled in an ideally constituted society then, notwithstanding the fact that he is unjustly burdened by society, he has committed a crime and becomes the legitimate target of punishment. Some things are wrong irrespective of the underlying injustices in society. At this point, lest my theory sound too divorced from the realities of who commit crimes and for what reasons, I want to consider Murphy's point about alienation. He states that "[t]here is something perverse in applying principles that presuppose a sense of community in a society which is structured to destroy genuine community" (1979, 107). The import of my arguments above is to suggest that there is nothing necessarily perverse in such an application. However, it is true to say that one must be extremely careful (and perhaps even merciful) in applying such principles. I posited above the frivolous example of someone who is denied housing benefit and commits rape. What about someone who is wrongly forced to undergo some extremely degrading form of treatment such as torture and who commits a random rape? Obviously within the terms discussed above the rape could not be typified as self-defence; rape is not the sort of action which defends against torture and nor, in this particular example, is it aimed against the relevant party, *viz.* the torturer. Nevertheless, it would remain to be considered in settling the culpability, or the punishment, the extent to which oppression may have affected the will of the person involved. If someone is treated like an animal then it may be no surprise if his rational capacities are not fully functional. In some instances it may be possible to say that a person's will is so affected that he ought not to be held culpable (which is how we typically consider animals) and in other instances it may be possible to say that, while the person is culpable, his degree of responsibility has been affected because of the oppressive behaviour he has experienced, and that the amount of punishment should be reduced accordingly. In the instances of self-defence we held that no wrong was done. In these instances of maltreatment leading to diminished responsibility we say that a wrong is done but one for which the individual's culpability is reduced.⁶³

⁶³ In the case of an irrelevant action aimed against the *relevant* party (e.g. torturing the torturer)

2 Why the Equilibrium of Benefits and Burdens is Held to be Important

2.1 The equilibrium as a feature of distributive justice

These theories are mainly couched in terms of distributive justice. The premises are a society founded on principles of distributive justice and a criminal act which mal-distributes benefits and burdens within that society. Distributive justice exists where a certain commodity is in limited supply. It is a way of ensuring that the correct people gain access to that commodity, and in the correct proportions. These I take to be the formal requirements of a theory of distributive justice, and there are any number of substantive principles which may decide what is a limited commodity, who ought to have access to it, and how much access particular individuals ought to get (Rawls 1972, 4 who talks in essentially the same terms, although he refers to "social justice"; cf. Lucas 1980, 163; see also Nozick 1974, 184). Finnis identifies three formal elements to justice *per se*: other directedness; that which is the commodity of justice; proportionality (Finnis 1980, 161-163). The distinguishing features of distributive justice are thus to be found with respect to these three features. Principally, distributive justice concerns a commodity which is resultant upon the co-operation of a given society. Thus, for Finnis, all the welfare benefits which are located within the modern state are the subjects of distributive justice. The constituent group is all the members of the particular society involved in co-operation. The proportionality of distributive justice is dependent upon the particular principles of distributive justice employed. Rawls' difference principle, for example, will give results at variance to Marx's needs principle (Rawls 1972, 60-78; Marx 1977, 569).

we might reduce the punishment because the original treatment is so bad as to reduce the culpability of the individual. It could then be argued that the punishment should be reduced more in this case than in the case of an irrelevant action against an *irrelevant* party (e.g. torturing a random person) because the desire for revenge aimed at the perpetrator of the wrong is an understandable (and perhaps even partially legitimate) one. Alternatively we might hold that an irrelevant action has no relevant target: the torturer only becomes a relevant target when the action aimed at him is some sort of resistance to the torture.

Second, it is necessary to distinguish between distributive justice as a mechanism for justifying punishment and a mechanism for deciding *how much* punishment ought to be apportioned in any one case (cf. the two strands of retributivism discussed in chapter 2 above). This is not to say that if a particular formula for distributive justice is used to justify punishment that a different one should be used to apportion it. Rather it is to point out that at this stage when I refer to the theories relying on distributive justice I am referring to theories which seek to demonstrate *why we ought to punish in the first place* on the basis of principles of distributive justice. If any of those theories get off the ground then it would no doubt be appropriate to apply the distributive principle (whatever it may be) to apportioning punishment to individual offenders.⁶⁴

Filling in the blanks in this formal description of distributive justice would lead to a report along the following lines. A certain commodity - the benefit of having one's purposes respected - is supposed to be limited; i.e. not everyone can have their purposes respected all the time. This follows from the fact that some ends are incompatible. My end of a good night sleep, for example, is incompatible with my neighbour's end of practicing his saxophone in the wee hours. More seriously, my end of making myself very rich may come into conflict with my neighbour's end of keeping his possessions. In order to resolve these potential conflicts of interest individuals must assume the burden of not pursuing their own ends willy-nilly, so that they and others can enjoy the benefit discussed the rest of the time. The outlaw is said to have appropriated to himself a certain proportion of this benefit to which he is not entitled. Punishment removes this excess commodity from him.⁶⁵ Wojciech Sadurski having rebutted John Rawl's thesis that there is an

⁶⁴ cf. James Sterba who argues that while Rawl's arguments in justification of distributive justice could also be used to convince original contractors that they ought to adopt a system of punishment there would then be the secondary question of which *retributive* principle they would actually adopt (Sterba 1980, 75-78).

⁶⁵ However, it seems that Morris, for one, is not clear on this point. He holds that the debt talked of above can also be erased by forgiveness which can be viewed "as a gift after the fact, erasing a debt, which had the gift been given before the fact, would not have created a debt" (1971, 79). In the case of punishment the debt is owed to society but forgiveness, as Morris discusses it, only makes sense if we think of the debt as owed to the victim. Society cannot give, before the act, a gift to D of the opportunity to violate V's rights. Only V can waive his own rights. See Hestevold (1987:249-257) for the argument that state punishment and mercy can be reconciled.

essential asymmetry between distributive and retributive justice asserts that the way is "cleared for the application of the principles of distributive justice to the sphere of punishment" (1985, 221). Likewise Finnis is open that punishment is justified by analysis of the principles of distributive justice (1980, 263). Perhaps surprisingly one theorist who takes issue with the applicability of distributive justice to the issue of punishment is John Rawls. He states that

[t]o think of distributive and retributive justice as converses of one another is completely misleading and suggests a moral basis of distributive shares where none exists. (1972, 315)

Rawls objection is that, whereas retributive justice is premised on the notion of C *deserving* his punishment, distributive justice (as he conceives it at least) is not based on moral worth. The principle which Rawls is defending here is that in a system of distributive justice no one individual "deserves" his share of a given commodity any more than another. Hence, he believes, distributive justice is not an appropriate mechanism for justifying punishment because punishment is conceptually linked with some statement of moral worth or desert on the part of the offender. The objects of distributive justice must receive the subject irrespective of moral worth. If the balance theorists discussed above regarded *punishment* as the subject commodity then Rawls point would be good because for retributivists the notion that C deserves his *punishment* is dogma. In other words, if punishment was the subject of distributive justice then we would be saying precisely that it should be distributed on the basis of moral worth. But none of the theorists actually see punishment as the relevant commodity to be distributed. Rather "benefits/burdens" is the commodity that they have in mind. In exactly the same sense that basic education, health services etc. are undeserved the benefits and burdens of rights are undeserved; there is nothing one can do to merit them, the argument goes, because one is simply entitled to them by virtue of membership of society.⁶⁶ These benefits and burdens then, are the subject of

⁶⁶ Cf. Sadurski (1985, 221-225) who argues that there is no asymmetry between distributive and retributive justice. Having established this he holds that the way is open to apply the principles of distributive justice to punishment. This line of argument seems strange to me because *if* one is arguing that punishment can be justified simply by distributive justice then the status of retributive justice should be irrelevant. Certainly it should be unnecessary to show that

distributive justice and punishment is one of the mechanisms by which it is ensured that benefits and burdens are fairly distributed. Punishment is a step taken to remedy any failure of the primary distributive mechanisms. Where C has taken an unfair proportion of the commodity, punishment's function is to restore the distributively just balance. Whatever the other merits of Rawls' distinction between retributive and distributive justice, his argument does not invalidate the use that the above theorists make of distributive justice in their treatments of punishment.

In order for the justification of punishment to be compelling it must be shown why the just distribution of benefits and burdens takes moral priority over the claim that agents ought not to be harmed against their will. The importance of maintaining equilibrium must be such that it overrides the countervailing objections to punishment such as, it is cruel, unnecessary or vengeful. The imperative of maintaining equilibrium is itself something which stands in need of demonstration if we are to have anything more than a program of intuitive conservatism (cf. Lacey 1988, 24). I will now consider three bases for the claim that the maintenance of equilibrium is of ultimate importance developed in the literature. These are intuition, egalitarianism and the social contract. It will be my contention that none of these is sufficient to ground the claims which benefits and burdens theory must make.

2.2 Intuition

Sadurski's book *Giving Desert its Due: Social Justice and Legal Theory* is largely a polemic against Rawls' *A Theory of Justice*. Sadurski's moral position is to reject Rawls' contractarianism but to accept his notion of reflective equilibrium; that is, "a two-way deliberative movement between our considered convictions of justice and the principles derived from the original position" (Sadurski 1985, 69). As Sadurski says, this is a recognition of the role of intuition in moral theory and he justifies this by pointing out that intuition is (a) useful for

retributive justice is akin to distributive justice as a precursor to establishing that distributive justice justifies punishment.

bringing moral principles and moral judgments in to line and (b) in fact how many people decide hard moral cases. He goes on to acknowledge the vagaries of moral intuitionism *qua* moral theory but contends that his appeal to intuition is not necessarily the same as meta-ethical intuitionism (1985, 70-72).

Intuitionism is the theory that we have a moral sense which can be used to settle all moral questions; the answers are "self-evident" in the same way that the faculty of sight makes it self-evident that grass is green. Therefore, the foundation of any set of intuited prescriptions is precisely this faculty of moral intuition. The main problem with this as a foundation for moral theory occurs where two people claim to have contrary moral intuitions. Any attempt to solve the conflict will, for those holding that intuition is the basis of moral judgment, have to be made by intuition itself. This is a viciously circular process. If you intuit that you should do P whereas I intuit that I ought not to do P then, as intuitionists, we have nothing to appeal to in order to resolve our dispute other than intuition. But when you say that intuition shows you that you were right I will respond that, on the contrary, intuition shows *me* that I was right all along. By this process of gainsaying we will never get any closer to resolving whether or not we are to do P. Sadurski's response to this dilemma, the difference for him between meta-ethical intuitionism and intuition, is to appeal to what the majority reflectively (or the reflective majority) offer as the reasons for their intuitions. He specifically denies the claim, which he attributes to meta-ethical intuitionism, that moral prescriptions can be true or false. This is not merely to draw a distinction between moral claims and scientific claims. It goes somewhat further leading Sadurski to the conclusion that "to strive for a universal consensus on matters of moral judgments is to be doomed to failure" (1985, 76). While consensus may be unattainable, we can inquire into what people hold to be the right answer to a situation and the reasons which they give for this. We cannot have a right answer but we can have a better answer.

In reply to this three things can be said. First, on some issues there may be no majority, or certainly no clear cut majority, intuition. Is the failure to reach a

majority intuition a sign that there is no best answer on a particular issue? Or perhaps it is a sign of mass moral blindness. Even if moral blindness is the case, so long as we hold intuition to be the foundation of moral theory we will never know which side is morally blind and which can see. Debate will be limited to each side accusing the other of moral myopia (also Walker 1991, 72 for intuition and retribution). Second, even if everyone except me intuits that killing is wrong why should I be held responsible for acting on my sincerely held intuition that it is a good thing to do? There must be recourse to something other than intuition if we are to settle such issues, otherwise morality becomes that which Sadurski specifically denies it is, a matter of head-counting (Sadurski 1985, 65; cf. Aune 1970, 31). The solution which Sadurski offers to both of these problems is the possibility of rational discussion. This leads me to the third criticism which is that to insist something can be a good reason for holding an intuition, whereas something else can be a bad reason, is to write normative criteria into the process of intuiting. One of Sadurski's tests for discussion is that the reasons for holding a particular moral intuition should "pass the test of universalization" (1985, 75). This itself is a normative requirement, specifically eliminating, for example, egoistic intuitions.⁶⁷

Michael Moore also appeals to intuition to back retributivist theory.⁶⁸ He claims that

we are entitled to rely on the connection that generally (but not inevitably) holds between the virtue in possessing an emotion and the truth of the judgment that that emotion generates. (Moore 1987, 208)

This is to say that emotion can act as a guide in divining whether a particular moral judgment, which we hold to be true, is in fact true. This is a particularly appealing argument in the context of retributive theory because most of us do in fact have a gut reaction to at least the most heinous forms of crime that some

⁶⁷ cf. Hudson (1967, 7 *et seq.*) for a discussion of rationalist intuitionism. This suffers from a similar problem. Either rationality is reducible to intuition, which is viciously circular, or it is a separate requirement, which means we are no longer relying solely on intuition.

⁶⁸ He uses the term "emotion" which I shall use interchangeably with intuition in my treatment of his theory.

sort of punishment is merited. Moore offers intuition as a heuristic guide to back up (or contradict) what we otherwise hold to be the case. His argument is complex but seems to me to suffer from a defect that renders intuition an implausible source of moral judgment. Moore himself says there is

a rationality of the emotions that can make them trustworthy guides to moral insight. Emotions are rational when they are intelligibly proportionate in their intensity to their objects, when they are not inherently conflicted, when they are coherently orderable, and instantiate over time an intelligible character. (1987, 190)

This is true. However, if it doesn't make emotion opposed to rationality, which is the immediate context of the quote, then it certainly makes emotion secondary to rationality. Moore elsewhere rejects the Kantian argument that emotions are irrelevant to the truth of moral belief, yet the force of his argument is precisely this (Moore 1987, 199). Of course, it may be necessary to examine an intuition, its force, the reasons why it is held, its implications and so on. But even if we then conclude that the intuition is true it is not the intuition itself that has revealed this, but the rational faculty's consideration of the intuition. This does not render emotions irrelevant to moral belief but it does render emotions irrelevant to the *truth* of a moral belief. Against this Moore argues that

[i]f we look in this way for a purely cognitive test for when emotions are epistemically reliable, my suspicion is that we will always end up slighting the role of the emotions in generating moral insight. We will do this because a purely cognitive test will inevitably seek to derive a criterion of epistemic warrant that is independent of any theory generated from the emotions themselves. It is like attempting to set up a criterion of epistemic warrant for perceptual experiences without using any theory derived from such experiences. (1987, 207)

The response to this is that as a matter of fact epistemological theories are given after a discourse of truth/falsity is entered into. But, and this is crucial, an epistemological theory is logically prior to any such discourse. By extension, any theory of the emotions is tempered by our actual experience of emotion but at the same time the criteria by which we say that we know a particular emotion to be good/bad or expressive of a true/false moral judgment cannot themselves be emotional. To have it otherwise is to enter into contradiction. At the very least

we would be committed to the claim that “whatever we emotionally feel to be the case is the case” which is not itself an emotional appeal.

Ultimately, emotion cannot offer a moral justification for punishment. The best it can offer (which is in fact what many punishment theories do offer) is a psychologically satisfactory reason for punishment. That is, it can hold up as a reason for punishing something that the majority believes to be a good reason, such as protection, deterrence or, in some cases, ill concealed blood lust. If the reason being offered is held to be a good reason simply because people actually do hold it to be a good reason then we do not have a *justification* for punishment. Thus, while we may intuit that balance is a good thing this does not, of itself, offer a moral backing for the contention that the balance of benefits and burdens ought to be maintained.

2.3 Egalitarianism

One way of justifying the contention that balance ought to be maintained is a resort to egalitarianism as discussed by Hugo Bedau in *Justice and Equality*. He states that a *radically* egalitarian society is one

in which every other consideration yields before the demands of equality, a society in which everyone is as equal to everyone else in as many respects as possible. (1971, 168)

Certainly, by this definition, radical egalitarianism would be sufficient to ground balance theory’s claim that maintaining a balance of benefits and burdens is more important than not inflicting harm on an individual. But radical egalitarianism is untenable not least because, as Bedau himself is at pains to point out, some social inequalities are inevitable while others can be justified. This leads Bedau to formulate a less radical egalitarianism which states that “[a]ll social inequalities not necessary or justifiable should be eliminated” (1971, 179). It should be axiomatic that the social inequalities caused by crime are neither necessary nor justifiable. Hence, if shown to be true, Bedau’s less radical egalitarianism would

be sufficient to establish a justification for punishment along the lines drawn by balance theorists.

Bedau's formulation of non-radical egalitarianism is in one sense uncontentious. Equality is regarded as a good thing to be striven for where there are no countervailing considerations to justify inequality. Indeed in this sense it may be seen as no more than a restatement of the principle that unless there is good reason to the contrary, objects with similar properties ought to be treated similarly (see also Benn 1971, 154 and Williams 1971, 126 whom Bedau cites in support of his theory). The specific contentious issue that I wish to consider here is whether Bedau's statement of non-radical egalitarianism is relevant to the concept of punishment. As a starting point I have already stated that according to its orthodox definition, crime is neither necessary nor justifiable. However, I will now contend that for balance theories there is a residual problem in the application of Bedau's egalitarianism to punishment.

The social inequality produced by crime, *as viewed by the balance theories*, is the unfair advantage gained by the criminal in accepting benefits without correlative burdens. Thus, while each law abiding member of society has adopted the burden of self-restraint the criminal has failed in this. Following, Murphy's line the criminal gets the added intrinsic benefit of expressing greater freedom. The problem for balance theory, if it is to be founded on egalitarianism, is that the inequality identified, in and of itself, *would* be justifiable. This is due to the theories' failure to correctly describe the loss inherent in a crime. If a particular individual is able to create for himself a situation where he has less burdens or more benefits than others then so much the better for him. This is commonly acknowledged in the case of work meriting reward; those who work harder have earned the increased fruits of their labours. Similarly, those who strive to become more physically fit and consequently have fewer medical bills, fewer unproductive days, and live longer, are seen to have earned these privileges by their own industry. But if crime is merely regarded as an activity which creates for the criminal increased benefits or fewer burdens then crime fits within these

two models of justifiable inequalities. If crime is viewed this way then it is no longer axiomatic that it is an unjustifiable social inequality. The reason that crime is unjustifiable is that it represents the arrogation of benefits or the avoidance of burdens *at the expense of others*. Thus, if my increased work activity involves making another to be my slave the increased fruits of my labour are no longer justifiable. Similarly, if I increase my level of fitness by forcing another to donate his heart to me then this privilege is no longer uncontentious. It should be noted though that these problems are only realised as such because we have an independent notion of right which demands that individuals are not treated as means to another's end. Egalitarianism, as stated by Bedau, cannot by itself lead us to this notion of right. At the very least there is the need for a substantive concept of justifiable and necessary inequalities. This requires resort to a further principle not offered by Bedau. Balance theory fails to make the claim that others lose out in the commission of crime except in a relative sense. There is no *prima facie* reason for holding this to be an unjustifiable inequality. Indeed, by earning more money than another I make that other worse off than me in a relative sense. The egalitarian principle is therefore inapplicable to the notion of crime which balance theorists following Murphy's line actually have.

However, as Burgh points out, there is another line, that taken by Morris, which states that the unfair advantage obtained by the criminal is that he has the benefit of the law's protection without adopting the correlative burden of obedience. This is an inequality that, by application of Bedau's egalitarianism, is to be eliminated. Egalitarianism offers a formal principle which states that people ought to be treated equally where there is no good reason for treating them differently. Causing avoidable harm to a person is generally held to be a bad thing. Therefore, it can be argued that the criminal and his victim ought not to be treated equally in the sense of getting exactly equal benefits and burdens because this can only be achieved by inflicting a harm on the criminal which we could chose not to. Egalitarianism itself does not provide us with the substantive moral principle which states that this sort of balance is sufficiently important to outweigh the harm caused to the criminal in punishing him. Balance theory

purports to do this. However, without recourse to some moral support for the contention that equilibrium is of ultimate importance the theory becomes circular. It states that balance is to be maintained on the basis of egalitarianism which states that unjustifiable inequalities are to be eradicated. But the reason why this sort of balance outweighs the orthodox complaint that avoidable harm is unjustifiable is left undeveloped or else given circularly in terms of the fact that the harm is justifiable because it is necessary to restore equilibrium. What is required to unravel this knot is a principle which conceptually links the receipts of protection/benefit to the fulfillment of duty/assumption of burden. Under such a principle it would be clear why we may (although not necessarily must) take protection/benefits from those who offend. This principle is suggested by Morris when he states that “the benefits of noninterference are *conditional* upon the assumption of burdens” (Morris 1971, 78; my emphasis). This though is not itself an egalitarian principle. This theory relies on precisely the sort of social contract highlighted in the next section.

2.4 The Social Contract

It is to be remembered that the Marxist critique, levelled by Murphy at his own Kantian theory, states that those who have been denied the benefits of social organisation should not be expected to assume the burdens of social organisation. Logically this entails a relationship of necessity between the enjoyment of benefit and the assumption of burden. It is not held to be coincidence that those who do not enjoy benefits need not assume burdens. Rather it is held that *because* individuals do not enjoy benefits they need not assume burdens. To say that an individual *need* not assume a burden is to make the assumption of that burden prescriptively *optional* to him. Given that the burdens we are discussing are the burdens of complying with laws in the form of “you must do X” or “you must not do X” the effect of this is to make the laws inapplicable to the individual. To say that the law “you must do X” is optional for an individual is to render that law a nullity for him because “must” is incompatible with optionality. This is by way of demonstrating that to say an individual need not assume a burden is the same as saying the individual does not

have that burden. Thus, the absence of benefits *entails* the absence of burden which is written

$$-\text{[benefit]} > -\text{[burden]}$$

Logically this implies that burden entails benefit which is written

$$\text{[burden]} > \text{[benefit]}$$

This no more than the reverse of the common logical progression

$$P > Q, -Q > -P$$

The logical statement that burden entails benefit, expanded, leads us to the conclusion that whoever is under the burden of compliance with the law must also be enjoying the benefits of the law. The Marxist critique is hinged upon the "empirical falsity" of the claim that a particular individual is actually enjoying the benefits of the law. But the crucial question for my current purposes is why is it that someone who is obliged to comply with the law must be in receipt of the benefits of the law. And the answer offered by many of the theorists is in terms of the social contract. This line is traced back to Murphy who (in his original theory) states that Kant

offers a theory of punishment which is based on his general view that political obligation is to be analyzed, quasi-contractually, in terms of *reciprocity*. (Murphy 1979, 83)

As developed this implies somewhat more than the Marxist critique alone would suggest. What we now have is a situation in which not only does burden imply benefit but also benefit implies burden, which is written

$$\text{[burden]} > \text{[benefit]} \text{ and } \text{[benefit]} > \text{[burden]}$$

In support of the second statement Murphy, discussing the citizen, says that

since he derives benefit from them [the laws], he owes obedience as a *debt* to his fellow-citizens for their sacrifices in maintaining them. (1979, 83)

We now have a situation of mutual sufficiency whereby simply enjoying the benefits of law is enough to place one under the burden of compliance with law, and conversely being under the burden of compliance with law means that one must actually be enjoying the benefit of law.

Murphy expressly denies that this is an intuited state of affairs, holding, as we have seen, that it arises on the back of a social contract. It is supposed that rational contractors, stripped of all contingent knowledge and beliefs, would adopt rules that would work to their own individual advantages when obeyed by others. To take a basic example, given that no rational contractor will want his life arbitrarily cut short, he will insist on a rule, to which all other rational contractors will agree, that he is not to be killed against his will. As a law this will proscribe murder. Similarly, although less obviously, the contractors will not want to have their possession of chattels in a constant state of flux, which will give rise to *inter alia* a law against theft. These are offered by way of examples. What is important is not the particular rules which the social contract is supposed to birth but the form of the contract and the nature of the obligations which it generates. Taking the example of a law against theft, the Marxist critique might be claiming that because an individual has had his property stolen he is not under an obligation to respect the property rights of others. But in fact it makes the wider claim that because the criminal has not had some particular benefit he is not under a *general* burden of compliance. The social contract has been broken thus exempting the individual from his part of the bargain. He says to (the individuals comprising) society "you have not granted me the benefits which you promised me, therefore I am no longer under an obligation to grant you the benefits which I promised you".⁶⁹

⁶⁹ cf. Duff (1986, 209) and Sher as discussed in section 1.3.4 of the present chapter, both of whom allow for proportionality between the degree to which an individual has been disadvantaged and the degree to which he is free from the constraints of the law. This is more intuitively appealing, but, given that the social contract has been broken, hard to sustain within

A *moral* social contract seeks to ground rights and duties in an original agreement between individuals. Balance theory and the Marxist critique depend upon such a contract. The moral social contract would establish that an individual gets rights as a signatory to the contract and that in return he agrees to abide by the rules (that is, fulfill his duties). It is precisely this that balance theory must demonstrate in order to establish that where a contractor takes more than his fair share of the advantages, the other contractors have been wronged. They have been wronged because of the metaphorical breach of contract. Similarly, the Marxist critique must rely on such a claim in order for its assertion that "I have not received the benefits of protection therefore I need not comply with the rules" to be proved true. If, as in my interpretation of Kant, rights and duties precede any contract, then the fact that I have not fulfilled my duties would not necessarily mean that the rest of society has been wronged. This would cast doubt on the justification of punishment offered by the balance theorists. Similarly, the fact that I have not received benefits would not necessarily have any implications for my obligation to fulfill my duties. This would cast doubt on the force of the Marxist critique.

The question now is whether the *moral* social contract can be established. Although this is an issue that has occupied volumes I will offer only one thought, albeit a thought which I believe is destructive to the moral contractarian's enterprise. Put simply, if my rights and duties are birthed *out* of a contract then *do I have a duty to fulfill the contract?* Under this model all my duties (including my duty to keep promises) are post-contractual. In other words duties begin at the moment that the contract is "signed". My agreement is itself a promise to be bound by the contract but one which is anterior to the completion of the contract. If before the completion of the contract I have no duties then I cannot be bound by that pre-contractual promise. Given then that balance theory and the Marxist critique clearly rely on some sort of moral social contract in order to

the moral framework offered by Murphy. This particular get out is therefore not open to Murphy.

establish their claims the theory and the critique must be found wanting. Again I must reiterate that the Marxist critique is a critique of the theory presented by one of the theorists himself and so any suppositions of the theory which the critique attributes to the theory can in fact be taken as suppositions of the theory.

In defence of Kant I contend that the above criticism is destructive of the sort of contract which Murphy needs to maintain in order for his theory to follow but that this contract is wrongly attributed to Kant. Crucially, in the quote given by Murphy in support of his claim to be a Kantian contractarian, he states that "political obligation is to be analyzed, quasi-contractually, in terms of *reciprocity*" (Murphy 1979, 83). I think that this is a correct interpretation of Kant, but that in making it Murphy misses the importance of what he has said, viz. that political obligation is to be *analyzed* in terms of reciprocity. Kant himself says that

[t]he act by means of which the people constitute themselves a state is the original contract. *More properly, it is the Idea of that act that alone enables us to conceive of the legitimacy of the state.* (1965, 80; my emphasis)

Stressing that it is not the contract itself but a rational ideal, expressed in terms of a contract, which gives rise to law, he states that

[t]he necessity of public lawful coercion does not rest on a fact, but on an a priori Idea of reason. . . (1965, 76)

Thus, for Kant the social contract is a heuristic device through which the state and its laws are to be better understood. The social contract though, is not the source of the state's or the law's legitimacy. These are legitimate because the arguments in support of them are rationally compelling for all agents. The arguments used may be, for purposes of clarity, expressed in terms of the social contract. Murphy then is right in what he says but wrong in what he infers from it. Crucially, it does not follow from a Kantian notion of right and duty as pertaining to all rational beings that where the benefits of society are absent an individual owes no (or fewer) duties. This only follows from a social contract which grounds right and duty in the contract itself, i.e. one which states that I

owe my duties *because* others have guaranteed me rights. In his more recent work Murphy concedes to Bruce Aune that his (Murphy's) interpretation of the social contract is at best Kant heavily filtered through Rawls (Murphy 1992, 48; Aune 1979, 167-168). Kant then is spared from the criticisms levelled at Murphy's version of the social contract in this section of the thesis.

In conclusion balance theory relies on the notion that burdens are to be assumed where benefits are enjoyed. This allows the proposition that those who shirk the burdens are unfairly advantaged. Given the ultimacy of the state of affairs in which benefit and burden are balanced, unfair advantage justifies punishment as the imposition of a burden necessary to restore equilibrium. The ultimacy of equilibrium is justified by reference to the social contract in which all beings agree to apportion the limited commodities of benefit and burden fairly. The force of my current argument is to deny that the derivation of rights and duties from a social contract is tenable. This means that any balance theory relying on the social contract to back its key contention (that maintaining equilibrium is more important than avoiding harm) is flawed. In fact, none of the bases considered so far is sufficient to ground this key contention. When I develop my own justification of punishment in the next chapter I will look to an equilibrium construed in terms of agents' interests in rights. For the time being I will turn to the criticisms of this stage of the argument developed in the works of others. These criticisms suggest that the equilibrium of benefits and burdens is not of paramount importance either by holding that some other moral base is needed in order to justify punishment (2.5.1 and 2.5.3); or by implying that there can be no meaningful discourse of justification (2.5.2); or by holding that punishment infringes moral injunctions more important than the maintenance of freedom (2.5.4).

2.5 Criticisms

2.5.1 Intuitive anti-retributivism

An argument raised by Andrew von Hirsch is that “the arcaneness of the benefits and burdens theory . . . troubles” him (1986, 59). This is presumably a reference to the outmoded contractarian thought which is explicit in Murphy’s construction of Kant, as well as the fact that retributivism is generally associated with reactionary vengeance. If von Hirsch’s objection is to retributivism *per se* then all he is saying is that he is troubled by a theory which pre-dates and conflicts with twentieth century liberalism (see fn. to 1986, 59 which suggests that this might be what von Hirsch is contending). This is no argument at all. This sort of uneasy gut feeling (which is really moral intuitionism with the virtue of honesty) is also present when he says that

[a]ccording to a wholly deontological rationale, punishment would have to be preserved even if it were found to have no preventive utility. To avoid this conclusion, I was compelled to argue in *Doing Justice* that while the benefits-and-burdens argument was the *pro tanto* reason for having a penal system, the absence of preventive utility could be a countervailing reason for its abolition in order to reduce suffering. (1986, 59-60)

The first sentence confirms what von Hirsch objects to in retributive theories of punishment. But as Jeffrie Murphy correctly notes of common objections to Kant, “[o]ne cannot refute a retributive theory merely by noting that it is a retributive theory and not a utilitarian theory” (1979, 79). All that can be left is a gut feeling that retributivism is too vicious to be a correct moral theory and this is not especially convincing against a theory that has been argued in such depth as Kant’s. Moreover, to argue that something is morally justified is to hold that it is compelling irrespective of one’s countervailing sensibilities.⁷⁰

⁷⁰ For the argument that retributive instincts are not primitive or barbarous in any case see Moore (1987, 185-217).

2.5.2 The incompatibility between determinism and retribution

Murphy develops what he calls a Marxist critique of his own balance theory. We have already encountered the major variation of this as developed in the literature. However, in the original presentation there is a second level (which is not so developed) at which the critique operates. According to this secondary level the process of alienation, identified by Marx, renders criminals determined beings.⁷¹ A prerequisite of justified punishment (certainly within the sort of Kantian theory espoused by Murphy) is that the individual criminal be autonomous so that he can be held responsible for his crime. But because of the way capitalist society is structured, individuals cannot be held responsible for crime: “[w]ithin bourgeois society, . . . crimes are to be regarded as normal, and not as psychopathological, acts” (1979, 105). This is because bourgeois society emphasises and encourages exactly those egoistic and selfish calculations that cause crime. It is to be noted, as was the case when we previously encountered the Marxist critique, that this critique holds punishment to be justifiable in principle but not within a capitalist society (see also Honderich 1976, 169-170).⁷² The implementation of communist principles will mean that all are equally advantaged by the rules of society and that individuals are liberated from the structural determinism of the bourgeois state.⁷³

⁷¹ Within classical Marxist theory individuals are determined because they are alienated from the means of production.

⁷² In *Doing Justice* von Hirsch makes substantially the same criticism as Murphy. However, his is open to wider interpretation that must be considered. He states that

[a]s long as a substantial segment of the population is denied adequate opportunities for a livelihood, any scheme for punishment must be morally flawed. (1976, 149)

Insofar as this refers to actually existent institutions for punishment it may well be correct and certainly contends no more than Murphy has contended. However, insofar as von Hirsch's assertion concerns *theories* of punishment it is incorrect. The fact that society is badly organised does nothing to detract from the moral integrity of desert theory (or utilitarianism or intuitionism or any other moral theory for that matter). It may mean that the application of that theory will lead to injustice but that theory *qua* theory may still be morally compelling (cf. von Hirsch 1986, 58 fn. where he appears to recognise this point).

⁷³ Ironically, perhaps, it seems to be an implication of Marxism that within a communist society *there will be no crime*. This means that punishment is justified in principle but unjustified in practice in a capitalist society and irrelevant in a communist society. This renders the project of justifying punishment an interesting academic exercise at best.

If true, this level of the Marxist critique is fatal to the retributivist account or, at least, to any attempt to apply it in a non-communist society. Alienation applies to the bourgeoisie as well as the proletariat; neither class, according to Marx, is in a proper relationship with the means of production and so both suffer alienation in the Marxist sense. If alienation renders individuals determined then, following this line, both the proletariat and the bourgeoisie are determined. In this case punishment could never be justified within a capitalist society. However, following the introduction to this thesis, it is not possible for the individual to act under the *idea* that he is determined. Thus, even if structural determinism is an accurate world account we must act as if each agent is *prima facie* the author of his own actions. Certainly Murphy is correct to identify free-will as a presupposition of retributivism. But the counter-presupposition, determinism, is not one on which any theory involving agency could be based. Any *justificatory* account must therefore assume that agents are not determined. Indeed, as Michael Moore notes, the claim that beings are determined does not merely impinge upon retributivism: it is damaging to any theory of justice "since no one could act in a way (*viz.*, freely) so as to deserve anything" (1987, 188).

2.5.3 Desert is a necessary but not a sufficient condition of punishment

Von Hirsch argues that desert is a necessary, but not a sufficient reason for punishment. He states that

[a] system of punishment is justified, the argument runs, simply because it is deserved. However, there are countervailing moral considerations. (1976, 53)⁷⁴

The only one of these raised by von Hirsch is the consideration that it is generally wrong to inflict human suffering. This observation is precisely the reason why the question about punishment arises in the first place. Punishment, a common sociological phenomenon, inflicts suffering. Suffering is *prima facie* to be avoided. Therefore, what, if anything, justifies punishment? But having

⁷⁴ He continues his thread in a footnote; "[t]o say someone deserves to be treated in a certain manner is to claim a reason for so treating him - but still allows one to conclude that he should not be treated as he deserves if there are sufficient countervailing reasons" (1976, 53 fn.).

established the case for a desert based theory of punishment we have already answered this question; what justifies punishment is that the criminal deserves it. However, von Hirsch claims that utilitarianism is needed to answer the question completely and it is here that his eclectic approach to moral theory seems particularly unhelpful. His argument runs thus

Step 1: Those who violate others' rights deserve punishment. That, of itself, constitutes a prima-facie justification for maintaining a system of criminal sanctions.

Step 2: There is, however, a countervailing moral obligation of not deliberately adding to the amount of human suffering. Punishment necessarily makes those punished suffer. In the absence of additional argument, that overrides the case for punishment in step 1.

Step 3: The notion of deterrence, at this point, suggests that punishment may prevent more misery than it inflicts - thus disposing of the countervailing argument in step 2. With it out of the way, the prima-facie case for punishment described in step 1 - based on desert - stands again. (1976, 54)

The reason why this argument fails is that it seeks to rest on two incompatible foundations. Step 1, that is desert theory, purports, in its own right, to answer the "countervailing moral obligation" raised in step 2. If it does not successfully do so, then the second step does not override the first but rather shows that the first step is not established. That is, desert theory sets out to answer the question "why can we inflict suffering in the case of punishment?" If it does not answer this question it is a flawed theory. Indeed it shows nothing. The answer that it offers is "we can inflict suffering because the individual deserves it". To then say that question "why can we inflict suffering" remains unanswered is to deny that the answer offered is true (cf. Moore 1987, 185-186). In this case desert theory shows nothing. But then von Hirsch is wrong to rely on it at all. Either, desert theory answers the question as to why we inflict suffering or it doesn't. If it does, then utilitarian considerations are entirely superfluous.⁷⁵ If it doesn't then there is nothing to be said for it and it should be disregarded entirely rather than miraculously resurrected on the third step. Utilitarianism offers us an internally

⁷⁵ It might also be added that the sort of arguments which Kant and Murphy use to establish a desert based theory would specifically deny the moral legitimacy of utilitarian logic in any event.

consistent theory of punishment but one which must rely on something external if it is to overcome objections that it would legitimate punishment of the innocent. Desert theory offers us an internally consistent theory of punishment and not one which needs recourse to any other theory as to what legitimates suffering. The objections that von Hirsch raises to desert theory would show that it is in fact not an internally consistent theory.

This objection of von Hirsch's is closely related to his intuitive anti-retributivism. Both objections are rooted in a moral hunch that balance theory ought not to be enough, by itself, to justify punishment.

2.5.4 Punishment fails to respect the criminal as an end in himself

In his essay *Does Kant have a Theory of Punishment?*, Jeffrie Murphy comes to doubt the accuracy of his previous construction of Kant as a balance theorist. Further, he questions whether balance theory can justify punishment in the first place. In particular he holds that

it is politically implausible to think liberty-loving rational beings would choose liberty-limiting punishment, not primarily to protect themselves *in advance* against rights violations, but rather to make sure that - *after the fact* - any unjust distributions resulting from rights violations would be rectified. (1992, 48)

Although there is an initial sense in this claim, it is, I think, flawed. Significantly, punishment cannot possibly protect in advance against rights violations. On the one hand, only the threat of punishment can stop someone from pursuing a contemplated course of action. By the time we come to consider the issue of punishment it is definitionally too late to prevent the crime. On the other hand, we *could* use preventive detention to stop certain contemplated crimes from occurring but this would not constitute punishment in the sense of a response to an actual offence. Certainly such an instance of "punishment" could not be justified by reference to Kant's moral theory. Although Murphy is now questioning Kant's theory of punishment there is nothing to suggest that he is giving up on Kant's moral theory.

A further response to this particular criticism is that the implausibility of punishment being a *post facto* rectification of “unjust distributions” is only apparent when those “unjust distributions” are conceived in terms of unfair advantages gained relatively over the rest of society. Given that Murphy has just (indeed in the same sentence (see 1992, 47-48)) rejected that this is the correct way to conceptualise crime, it is disingenuous of him to use it as the basis of a second criticism of his original theory. Alternatively, if Murphy is now going to hold that the primary evil in a crime is the overriding of the *victim*'s rights then he ought to consider whether punishment is suited, after the fact, to rectifying this particular wrong

However, even if it is so suited there remains the objection that punishment treats the criminal as a means to an end rather than as an end in himself. Anthony Duff, for instance, objects to Murphy's original thesis that while it may be consistent with an individual's autonomy to require him not to offend it is not so consistent to punish him in the event that he does offend. Duff then proposes his own solution to this problem which is the solution he supposes Murphy and Kant, *qua* Rawlsian contractarian, would propose; namely, that the criminal himself has willed his own punishment (Duff 1986, 208).⁷⁶ He bases this on a social contract under which the criminal consents in advance to his punishment if he should commit an offence. However, having considered this possibility, Duff debunks the contractarian scheme and leaves standing the claim that to punish is inconsistent with treating the criminal as an autonomous subject. He says of a hypothetical criminal

[w]e may try to persuade him to obey the law, by appealing to the relevant reasons which justify its requirements; if he still breaks the law we may condemn and criticise him; we may even try to persuade him that he ought to seek or accept punishment, as a debt which he owes to his fellow-citizens; but we cannot impose that punishment on him, against his present and avowed will, and claim that in doing so we are treating or respecting him as an autonomous agent. (1986, 223)

⁷⁶ This is despite the fact that Kant for one is explicit that it is irrelevant whether the criminal has willed his own punishment (see chapter three, section two).

But this assumes that punishment is forward looking rather than (as is the case with retributive punishment) backward looking. The issue is not whether the criminal consents to the future state of his punishment, but whether he was morally responsible for his past crime. Of course this is a considerable problem for any theory of punishment but retributive punishment at least need not concern itself with the criminal's consent to punishment.⁷⁷ All that needs to be shown for a retributive theory of punishment is that in acting as he did the criminal was culpable and that the correct response to culpability is punishment. Which is of course not to say that balance theory has established these claims. But for the time being it must be noted that we do not treat the criminal as *merely* a means if punishment is something to which he is rationally bound to consent. It may be that we drag him to his punishment screaming and kicking but if, *qua* rational being, he must consent to it, we treat his rationality as an end and not nearly as a means. I maintain that this is all a justification of punishment need show, not that the particular criminal has signed some pre-dated agreement consenting to punishment under certain circumstances, nor that at the instant of punishment's infliction he is actually willing to undergo the experience. What this entails is a distinction between treating someone as a means and treating him as an end notwithstanding that he does not realise this. Duff himself concedes that matters would be different for his objection if it could shown that the individual is deprived of his rational faculties - for examples, the cases of a drunk or someone who is mentally disordered (1986, 223). The force of my reply is to argue that *if* punishment is rationally justified then it follows that a criminal refusing to consent to his punishment is not acting rationally. Hence it follows that to override his irrational will is permissible and not inconsistent with his autonomy for the reason that autonomy is a feature of rationality rather than irrationality (see further Kant 1991, 144 and Falls 1987, 36-37).

⁷⁷ The reason why Duff raises his objection is the line of thinking in contractarian thought which says that a rational being would will the ideal state of affairs and that therefore all beings are bound by this state of affairs. In fact, in this sense, I would allow that Kant is a contractarian; that what is morally right is what the rational being with a will, stripped of all contingent factors, would say is morally right. However, it is still irrelevant whether or not the criminal consents to his punishment; all that must be shown is that it is rationally compelling that we punish this criminal. Duff's response to this line of argument is to say that there is no one rationality and no one correct moral theory (1986, 219, 230).

3 Punishment's Ability to Restore Equilibrium

Given the above arguments this is the easy stage in the process. Whatever it is that the criminal is supposed to have gained unfairly, to have maldistributed in his own favour, punishment takes away from him. For instance, as we have seen, Morris talks of punishment restoring equilibrium by "exacting the debt" which the criminal owes to society (Morris 1979, 79). The criminal owes restraint to society and this debt is exacted by imposing restraints upon him when he fails to adopt them voluntarily. It should be noted that this scheme is self-justifying. A certain balance of benefits and burdens is held out as being the ideal balance. Crime upsets this balance and punishment restores it. Punishment then is perceived as being simply another mechanism for ensuring that the just distribution of benefits and burdens is maintained. There are two main criticisms of this stage of the argument developed in the literature. The first is targeted at Punishment (capital P) in suggesting that balance theory does not justify Punishment but some other intervention. The second is targeted at punishment (small p) in suggesting that balance theory can not prescribe which punishments are actually justified in the instance of a given crime.

3.1 Criticisms

3.1.1 Balance theory allows some response but not Punishment

A consistent advocate of this line is von Hirsch. Throughout his many incarnations - he has proposed at least three different theories of punishment since 1976 - he has held that balance theory, in and of itself, would not allow punishment. In *Doing Justice* von Hirsch comes closest to actually advocating a balance theory but still does not consider it to be a sufficient justification of punishment; in order to be fully justified, he claims, punishment has to look also to deterrence theory (1976, 54). Von Hirsch considers Kant's desert theory and says of it

Kant's theory . . . accounts only for the imposition of *some* kind of deprivation on the offender to offset the "advantage" he obtained in violating other's rights. It does not explain why that

deprivation should take the peculiar form of punishment. (1976, 48)

The crucial question must be as to the difference between punishment and 'other kinds of deprivation'. Von Hirsch goes on to say that "[p]unishment differs from other purposefully inflicted deprivations in the moral disapproval it expresses" (1976, 48). This seems to me to involve flawed reasoning. On the one hand we might contend that whatever is done to the criminal by way of imposing a deprivation constitutes at least the attempt to punish. But let us, for the sake of argument, posit a deprivation which is *prima facie* not a punishment. According to von Hirsch, the thing which distinguishes it from a punishment is that it doesn't express moral disapproval. Supposing then that in response to the taking of an unfair advantage over all others we impose this deprivation on the criminal. Simply by virtue of the fact that we are imposing it *in response* to something that is held out as being wrong the deprivation becomes an expression of moral disapproval. Hence, by von Hirsch's own definition, the deprivation becomes a punishment. If, as von Hirsch concedes, Kant's theory demands some sort of deprivation in response to the taking of an unfair advantage then it seems to me to be impossible to conceive of a deprivation which would not take the "peculiar form of punishment". Punishment expresses moral disapproval precisely because it is inflicted in response to something that is wrong and is held to be administered because of the wrong. Von Hirsch claims that desert theory is an improvement over general deterrence because it only allows punishment of the guilty (1976, 44). In other words punishment is a response which is only ever appropriate where someone has committed an act which is held out as being wrong. It is flawed then to contend that desert theory could allow a reaction which is anything other than an expression of moral disapproval.

In *Past and Future Crimes* (where von Hirsch has expressly moved away from balance theory and towards a theory of censure) he offers a similar criticism:

[e]ven if the benefits-and-burdens theory were correct - that *something* needs to befall the offender who has gained an unfair advantage over law-abiding persons by infringing upon mutually beneficial rules - it is not obvious that the "something" need be the hard treatment involved in punishment. Suppose the

offender were formally censured and also required to compensate the victim. Why is this "disadvantage" not sufficient to satisfy the benefits-and-burdens theory? (1986, 59)

This criticism is slightly more to the point than the formulation offered in *Doing Justice* but is still flawed. Von Hirsch is right to suggest that not every crime merits "hard" treatment, however that is to be defined, but he would be wrong to draw a distinction between "soft" treatment and punishment. Whatever we impose upon the criminal in response to his crime constitutes his punishment.⁷⁸ It would be a case of the definitional stop to say that giving someone a stiff talking to or formally censuring him is not punishment simply because it does not come within a preconceived notion of what is meant by that term. So the answer to von Hirsch's question is that censure plus compensation may well be sufficient to satisfy the benefits and burdens theory. Whether or not it is will depend upon the nature of the offence. Within balance theory censure plus compensation will be sufficient where they are jointly equivalent to the advantage that the criminal has taken over the rest of society.⁷⁹

Von Hirsch is not alone in this criticism of balance theory. Anthony Duff in *Trials and Punishments* has it that the state

must perhaps as a matter of consistency be prepared to do *something* about those who break its laws; it cannot simply ignore breaches of the law as if they did not matter: but this argument requires at most that those who break the law should be judged and condemned through a system of criminal trials, not that they must be *punished*. (1986, 215-216; also Ten 1987, 59, 65)

Within the context that Duff raises this objection there are two interpretations. The first, which goes to the core of this thesis, is that Duff cannot see why, as a matter of fairness, the state ought to institute punishment in the first place. As

⁷⁸ Which is not to say that whatever we impose (brutal torture, a medal, a safari) is an *appropriate* or justified punishment. The point is that the object of our inquiry is the response that is justified in the singular case of an offender committing an offence.

⁷⁹ See further *Censures and Sanctions* where von Hirsch develops this idea of censure as the correct response to crime (1993, 14). Again he draws a distinction between the visitation of censure and the hard treatment involved in punishing. Again it is argued that there is no reason in principle why a censure delivered to a criminal in response to a particular crime should not constitute punishment.

long as the state is consistent in its treatment of offenders (e.g. by letting them all off) it does not act unfairly. The response to this is that the aim of the state in this area is not to act consistently towards all offenders but to act consistently towards all of its subjects so that each has an equal share of the benefits and burdens of living in organised society. The second interpretation of Duff's criticism is substantially the same as the criticisms we have already seen raised in von Hirsch. That is, why must the state actually punish, why not impose some "lesser" form of treatment? Again, the flaw in this thinking is to suppose that there is a spectrum of responses to crime on which punishment only starts at some sufficiently "hard" point. At the outset of this thesis I argued that the problem with punishment is what, if anything, justifies the state's imposition of something non-consensual on an offender in response to his wrong actions. This is a question which demands an answer whether the non-consensual imposition is conceived of as being punishment, censure, compensation, judgment or whatever.

Galligan raises a similar objection but with a particular slant. He argues that balance theory does not allow us to specify which sort of disadvantage is to be imposed on the offender, whether it be imprisonment, fine, community service etc. (Galligan 1981, 158-159). It is true to say that balance theory itself cannot settle this matter. What the theory can tell us is whether a proposed disadvantage is fair or not. Thus, it might be that a three week education course is not sufficient to counterbalance the unfair advantage gained by murder. However, it might also be that ten years in prison or five years in prison *plus* community service are equally able to offset the unfair advantage. This is no more surprising than the revelation that a ton of bricks weighs the same as a ton of feathers. It would be a matter of policy whether the lengthier sentence or the shorter term with other constraints is imposed on the offender. Balance theory cannot set the policy for us and it need not purport to. What it can do is tell us why a proposed penalty is fair.

3.1.2 Unfairness can't be measured

While Galligan objects that balance theory cannot decide between X and Y where $X = Y$, this objection argues that balance theory cannot decide between X and Y

where X does not = Y . That is, balance theory can't justify one particular amount of punishment as opposed to another. Murphy notes of his own theory that "unfairness is unfairness, murder being no more *unfair* than robbery" (1979, 85). For Murphy, this is so because the unfairness which he emphasises is the exercise of freedom in a case where the law-abiding have been restrained. Murphy's solution to this problem is pragmatic. He suggests that the best a legal system can hope for is to grade both crimes and punishments in order of seriousness and then match them, the most serious punishment for the most serious crime and so on down to the least serious punishment and the least serious crime. This achieves a system of proportionality. Burgh insists that this solution is unsatisfactory because, as he rightly points out, the theory itself does not allow us to ground even the most basic of intuitions, such as, that murder is more serious than robbery (1982, 207). It is disingenuous of Murphy to say that murder is no more unfair than robbery and then intuit that murder is more serious than robbery for the purposes of developing a hierarchy.

Sadurski, whose considered moral position is expressly one of competing intuitions, takes this line of Murphy's that punishments can only be agreed as a matter of convention. He asks

[w]ould anyone consider as 'arbitrary' the proposition that the punishment for rape should be more severe than for stealing a raincoat? . . . This judgment relies upon a certain hierarchy of values which helps us compare the severity of punishment with the gravity of a crime. . . As we can never be certain about an absolute validity of such judgments, it follows that we can never be sure that any particular system of punishments is absolutely just. (1985, 238-239)

This constitutes acceptance of Burgh's claim that the benefits and burdens theory fails to offer a justification for the quanta of particular punishments. Not only does Sadurski concede this point in the final sentence of the above quote but further the residual *intuitive* strength of his theory is undermined by anyone who believes that the punishment for simple theft should be more severe than that for rape. There is little response to such a person if intuitions are the only justificatory reasons given for action. As soon as reasons are given for intuitions (for example,

that rape is a more severe intrusion of agency than theft) the argument from intuition is given up.

A similar response to Murphy's (and Sadurski's) intuitive pragmatism is offered by von Hirsch. In *Past and Future Crimes* he holds that both the interests abused by crime, and the response which society makes are culturally relative. For example,

[i]t is sometimes said that there is no objective way of gauging the importance of the interests infringed by criminal conduct, because those interests depend upon culturally conditioned beliefs about what is most worth preserving in life. (1986, 66)

It should be noted that for von Hirsch this does not present a problem. He goes on to say

[o]f course they depend on such beliefs, why not? . . . A free society . . . is one in which individuals *should* be entitled to decide for themselves what interests are most important in their lives. (1986, 66-67)

I would question whether this leads to a morally justified as opposed to psychologically satisfactory account of punishment. In his most recent work, *Censures and Sanctions* von Hirsch stands by his claim that the quantity of punishment is culturally relative, stating that "the amount of disapproval conveyed by penal sanctions is a convention" (1993, 19).⁸⁰

If these contentions are true then punishment theory is doomed in my reckoning because it can never offer a practical account of justified punishment. It might tell us what justifies punishment but it is impotent to tell us how much punishment is justified in a particular case. Apart from intuition the best backing that could be given to any claim about how much punishment a criminal deserves is either a type of moral point of view theory or a version of legal positivism. In the first case we might claim that if it is accepted that M is the correct moral point of view then

⁸⁰ cf. Day (1983, 744-745) who agrees that the correct amount of punishment is a convention but claims that we could make this convention fair by the theoretical auctioning of licenses to commit offences. The amount of deprivation that can justly be imposed for a particular offence would then be equal to its license value on the open market.

it would be true (stipulatively by reference to M) that R1 is a more important right than R2 and that therefore infringement of R1 merits more punishment than infringement of R2. There would, however, be no way of establishing that M is the correct moral point of view. This would be merely a matter of belief or intuition. In the second case one would claim that given a "legal" system (which is not justified on moral grounds) the definition of crime and the correct amount of punishment are to be found within the rules of that legal system. Again, as in Burgh's response to Murphy, this is unsatisfactory because while benefits and burdens theory purports to offer a morally grounded reason for why we punish it is failing to offer a morally grounded theory of how much we should punish in a particular case. This rather strips the theory of practical significance.

Sher defends against this criticism of the theory. In his example he considers the *prima facie* intuitive claim that a murderer deserves more severe punishment than a tax evader. Grounding this claim in the benefits and burdens calculus he maintains that

the reason he [the murderer] has benefited more is not that he has indulged a stronger inclination, nor yet that he has received greater financial or psychic rewards. It is, instead, that he has violated a moral prohibition of far greater seriousness. (Sher 1987, 81)

It is to be remembered that the original problem with Murphy's theory in this respect is that the benefit identified in each case of crime, and to be eradicated in each case of punishment, is the (unfair) expression of freedom. Because unfairness cannot be quantified neither can the amount of punishment appropriate. But Sher, in comparing the benefit with the violation of a quantifiable moral prohibitions, sidesteps this problem. What is required now is (as Sadurski and von Hirsch note) a hierarchy of moral goods (and by extension moral prohibitions against infringing those goods) grounded not (as Sadurski and von Hirsch contend) in intuition or convention, but in a developed moral theory. As Sher notes elsewhere "our justification of punishment construes fairness as operating upon *preexisting moral obligations*" (1987, 89; emphasis added). It is by reference to such a moral theory that my own justification of punishment will proceed.

4 Conclusions

4.1 The theories contend that either (a) freedom or (b) the benefits of rights are the fitting subjects of (re-)distributive justice. In the first case the theories fail to take us beyond the point reached by Kant as discussed in chapter two above. Specifically, in terms of freedom, punishment is a gratuitous further diminishment of freedom (section 1.3.3 of this chapter). In the second case I believe that the theorists are on to something. There does seem to be little point in saying A has right R if having R makes no practical difference to A. This insight will form a major step in my own justification of punishment. This particular branch of the theory takes us beyond the point reached by Hegel because rather than merely positing right as that which is diminished by crime it indicates a specific aspect of right which is diminished by crime (and which by inference is to be restored by punishment).

4.2 However, as presented, the theories do not prove the vital step that punishment can restore to the law abiding the benefit of right. Without this the theories do not provide a justification of punishment. The theories do not show us why punishment is preferable to anarchy. To demonstrate this the theories would have to show at least that the maintenance of equilibrium is more important than the avoidance of suffering. The proofs of this offered in the literature, whether in terms of intuition, egalitarianism or social contract theory are unsatisfactory. While there is a conceptual link between the adoption of burden and the maintenance of benefit the theories fail to demonstrate that the *post facto* imposition of burden restores the wrongly assumed benefit. In this case, the imposition of burden (in the form of punishment) is no less gratuitous than the deprivation of freedom; it constitutes a further diminishment of the sum total of benefits of rights enjoyed by the community of agents.

4.3 Concerning the criticism that the theories fail to take account of the victim I believe Hegel had it right. That is to say, I believe that punishment focuses on an interest vested in both the immediate victim *and* a wider community. This is seen

in Hegel's discussion of punishment restoring a dual aspect of *Recht*. By protecting the realm of right punishment guarantees the particular right within that realm and by protecting the particular right punishment guarantees the integrity of the realm. In the final chapter of this thesis I will develop this idea by employing the Gewirthian notion of generic rights, i.e. rights that are necessarily possessed by each and by all.

4.4 Concerning the Marxist critique of the theory, I believe that the basic form of the critique is too crude and that the form developed by Sher can be re-explained in terms of self-defence. The basic form of the critique is too crude because it fails to see that some things are morally (and legally) wrong irrespective of any deprivation experienced by the criminal. To have it otherwise is to fail to acknowledge that rights are important. There is no reason why *my* right to bodily integrity should necessarily lapse just because (through no fault of mine) your right to monetary assistance in the face of poverty has not been fully observed. The developed form of the critique can be re-explained in terms of self-defence because in the situation where another agent is directly threatening my interest in my right I have a concomitant right to defend my threatened interest. This holds irrespective of any commitment to Marxism.

Chapter Five: The Justification of Retributive Punishment Within a Gewirthian Framework.

I believe that none of the philosophers considered so far offers a convincing answer to one crucial question. That question is this:

Whatever the equilibrium disrupted by crime is held to be, and however that equilibrium is supposedly justified, how does punishment restore equilibrium?

If a certain equilibrium can be shown to be right then it follows by consideration of what it means to say that something is right that that equilibrium ought to be maintained, and where necessary restored. What has not yet been shown is how *punishment* is supposed to function within this equation. Kant's theory of punishment was constructed in terms of freedoms. He failed to show how punishment restored freedom in any meaningful sense; on the face of it punishment constitutes the diminution of freedom. While it is logically true to say that $-(-X) = X$ and that therefore $-(-\text{freedom}) = \text{freedom}$, this is merely formal. When it comes to adding substance to this insight we were given no good reason why $\text{punishment} = -(-\text{freedom})$ and hence $\text{punishment} = \text{freedom}$. This is critical because in order for Kant's theory to follow it must be shown why $\text{punishment} = -(-\text{freedom})$. From the point of view of the criminal $\text{punishment} = -\text{freedom}$. From any other point of view it is not clear how punishment has any bearing on freedom. Specifically, from the victim's point of view it is too late to restore his wrongly denied freedom.

Hegel's notion of equilibrium was in terms of an equilibrium of rights. However, in many respects it suffered from similar deficiencies to Kant's equilibrium. In particular the crucial part of the demonstration is not to prove that $-(-\text{right}) = \text{right}$ because this is a simple and obvious logical statement. What must be proved is that $\text{punishment} = -(-\text{right})$ and Hegel failed to achieve this.

The twentieth century proponents of the theory draw up an equilibrium in terms of the benefits and burdens of living in community. They claim that there is a just distribution of these benefits and burdens and that punishment is a secondary mechanism designed to rectify any failure in the primary distributive mechanisms. Their attention to the benefits and burdens of social living is a step forward in two ways. First, it looks to a wider system than the binary criminal-victim relationship. This offers us the potential to sidestep the common objection that as concerns the victim of the offence the crime cannot be undone. Second, it focuses on something other than the abstract concept of right showing what has been lost in the commission of a crime. However, I do not believe that the theories focus clearly enough on what is at stake in crime (a point which I will deal with in furthering my own justification of punishment in this chapter). Consequently they are unable to demonstrate how that which they contend is lost in the commission of crime is restored in the act of punishment.

This present chapter comprises two main sections. First a consideration of Gewirth's tentative justification of punishment as offered in *Reason and Morality*. I will contend that he fails to answer the question posited at the start of this chapter. This is so, I believe, because as with the twentieth century proponents of benefits and burdens theory, he fails to accurately describe that which punishment is supposed to restore. In the second section of this chapter I will develop my own justification of retributive punishment utilising Gewirth's own concept of rights. I will propose what I believe to be a unique subject matter for equilibrium, the dispositional value of right, and I will contend that this is restored by punishment thus answering the question as to how punishment restores equilibrium. In combination with Gewirth's derivation of right and Beyleveld and Brownsword's derivation of law from the Gewirthian category of right, this offers a justification of retributive punishment drawn exclusively from the concepts of offender and offence.

1 Gewirth on Punishment

The first thing to be said here is that Gewirth does not have, and does not claim to have, a consistent and developed theory of punishment. In the final chapter of *Reason and Morality*, entitled *Indirect Applications of the Principle* (i.e. the PGC), he discusses punishment under the PGC but this is more in the fashion of 'interesting thoughts' than rigorous argument (1978, 294-304). The purpose of his book is to argue for a rationally justifiable moral principle and in this he is, I believe, successful. It is for other works, including those later writings by Gewirth himself, to fully explain the application of the principle to morally relevant situations.

Gewirth's consideration of punishment resorts to analogy with a game. Just as a baseball hitter cannot complain if he is given out by the umpire so the defendant in a trial cannot complain if he is imprisoned by the judge (1978, 274). The first issue to consider is how Gewirth gets round the major *disanalogy* between sport and the judicial process *viz.* that sportsmen voluntarily partake in their chosen fields in a way which citizens in a legal system do not. Sportsmen choose to play a particular game and by virtue of this choice subject themselves to the rules of the game. The same cannot be said of the criminal who, in many cases by the very nature of his act, is demonstrating that he does not volunteer to partake in society and does not subject himself to the rules of that society. It is open to the hitter to give up baseball if he decides that he no longer agrees with the laws of the game. Unlike the times of Socrates and the city state, it is not generally open to the modern citizen to renounce his citizenship and leave society if he decides that he no longer agrees with the laws of the land.⁸¹ To get round this disanalogy Gewirth develops a fourfold typology of justification for social rules: the optional-procedural, the static-instrumental, the necessary-procedural and the dynamic-instrumental (1978, 282-327). For present purposes the important distinction

⁸¹ Furthermore, even if he were to emigrate, the agent would not escape the core dictates of the criminal law because these are no more than expressions of the central duties under the PGC and therefore non-optional from the point of view of agency.

within this typology is as between those rules which are procedurally justified and those which are instrumentally justified.

A social rule is procedurally justified where it is created in line with a procedure which doesn't breach the PGC and its substance is not in breach of the PGC. Thus, sportsmen may come together and decide that, for example, the duration of the game will be ninety as opposed to eighty-nine or ninety-one minutes. Nothing in this rule violates rights to freedom and well-being and given a satisfactory method of decision, one which is voluntary, the selection method does not contradict the PGC. If the governing body of rugby union makes the decision that rugby union matches should be ninety rather than eighty minutes long, this is legitimate even if all the teams affected by the decision disagree with it because at some stage in the constitution of the rugby union it was voluntarily agreed that the governing body would have the mandate to set the rules of the game. If on the other hand the governing body of association football dictates that rugby union matches will be ninety and not eighty minutes long it is an illegitimate rule because it is non-consensually imposed. The passing of the rule does not conform with the voluntarily adopted procedure set by the rugby teams when constituting themselves as a union and thus, being a breach of those teams' voluntariness, the rule is in breach of the PGC. If the World Boxing Council decides that a contest will only be judged as complete when one contender has killed the other then this rule is unjustified even though the WBC has a legitimate mandate to set the rules of its sport. The rule would be illegitimate because the substance of the rule itself violates the PGC.

A rule is instrumentally justified where it is necessary to ensure that the PGC is upheld. Agents live in relation with one another and the projects of any particular agent will not necessarily be compatible with the rights of others. Rules or laws (as instrumentally justified rules are more usually known) are justified when they exist to ensure that the rights to freedom and well-being of each agent are upheld. From the point of view of the agent there is no optionality concerning the upkeep of the PGC: it is an imperative. Because of this it is not an objection to a

particular law that the individual agent did not consent to it. This argument is in parallel to one which claims that it does not matter whether an agent agrees with the PGC or whether he has ever given a moment's thought to it. He is nevertheless, as an agent, bound by the principle. As Gewirth puts it, the

difference between procedurally and instrumentally justified social rules is that the latter, unlike the former, are required for restoring and maintaining the equality of generic rights. Since this equality is obligatory, the rules that protect it are also obligatory. (1978, 300-301)

The distinction between voluntary associations (including sports) and the minimal state is that the former govern morally optional activities while the latter governs morally imperative activities. An agent does not have to be bound by the rules of a voluntary association unless he freely chooses to place himself within that association. He therefore only becomes subject to the rules when he decides to partake in the optional activity governed by them. An agent is bound by the rules of the (legitimate) minimal state because the activities governed by that state are not optional for the agent. He necessarily interacts with other agents and this interaction is governed by the laws set down by the minimal state.⁸² It should also be noted that, in line with Kant, Gewirth makes the entirely valid point that even if a particular agent does not agree to be bound by the law this is a merely empirical contingency. As a *rational* agent he is bound to concur with the PGC and with laws derived from that principle (1978, 291).

The distinction between procedurally justified rules and instrumentally justified rules also gets round another objection to the analogy with sport. It could be said that the rules of a game are in one sense arbitrary and that agents should not be bound by arbitrary laws. However, as we have seen above, in order for the law to be binding on the agent, and from the point of view of legal idealism, for it to be law, it cannot be arbitrary. It must not contradict the PGC and it must be passed

⁸² Even in the purely negative case of cutting himself off from other agents and living as an ascetic the agent's actions impinge upon others. With regards to the criminal law he will generally fulfill his duties to others simply by not acting positively towards them. Nevertheless the PGC imposes duties to help other agents in need and it may be that the hermit fails in meeting some of these duties.

and enforced by an institution which is not itself in contradiction with the PGC (1978, 279). In fact, with regard to that section of the law which is specifically the criminal law, rather more than non-contradiction of the PGC is required. What is required is active observance of the most basic injunctions of the principle; agents must not murder, assault, steal from, defraud or endanger each other.

Gewirth's theory of retributive punishment is akin to that of all the writers considered so far in that it sees the proper function of punishment as being to restore some pre-crime equilibrium. Unlike the benefits and burdens theorists discussed in the previous chapter, Gewirth has, in the PGC, a principle which resolves the potential conflict between retributive and distributive justice. For Gewirth

[b]oth kinds of justice have the function of maintaining or protecting the supreme principle of moral rightness consisting in the equality of generic rights that the *PGC* makes mandatory for actions and institutions; thus both kinds of justice are instrumental to securing this equality. (1978, 293)

Given that the PGC is a grounded concept we are closer to a grounded justification of retributive punishment.

The action of crime is seen as artificially setting up an "occurrent inequality" between the criminal and the victim (1978, 294). As the criminal is the one guilty of bringing this situation about he is the one who can properly be expected to restore it. The remaining question, that which none of the retributivists from Kant through Hegel to the twentieth century proponents of the tradition have properly dealt with, is how punishment operates as a mechanism to restore the pre-crime equilibrium.

1.1 Punishment's ability to offset an occurrent inequality of noninterference between agents

Gewirth accepts that punishment is instrumentally justified but defends it against the vagaries of utilitarianism. This is because the end towards which Gewirthian

punishment is instrumental is the PGC itself which specifically prohibits that individuals are treated as a mere means. The object of the principle as well as the secondary mechanism of punishment are one and the same, namely upholding of the generic rights of each agent. The criminal law and punishment are instrumental to the end of "occurrent equality of generic rights" (1978, 296). Thus, the theory is deontological in the sense of being justified by reference to an antecedent principle and simultaneously instrumental in expression. As to how punishment achieves its end of upholding the occurrent equality of rights, Gewirth is very close to Jeffrie Murphy when he states that

A not only removes X units from B; he also adds a comparable number of units to his own stock of well-being. For while A benefits from other persons' observing toward himself the aforementioned restraints on their conduct, he disrupts the equality of mutual restraint by lifting the prescribed restraints on his own conduct; he hence derives a certain degree of additional satisfaction. The amount of the latter is roughly proportional to the amount A has removed from B. (Gewirth 1978, 297)

As against Hegel, Gewirth claims that punishment is itself an evil, necessary for the greater good of restoring the rightful equilibrium under the PGC. In line with Hegel, Gewirth acknowledges that while it is good for the threat of punishment to deter potential criminals, the good of this end is subordinate to the justified good of restoring the rightful equilibrium (Gewirth 1978, 299).

Gewirth's theory is stronger than those of other modern thinkers discussed in the previous chapter in at least one sense. Because he has a developed hierarchy of rights he is able to say how serious an infringement of right is. Any retributive theory, to have practical import, must be able to say how serious a crime is. Otherwise there can be no proportionality between the crime and the punishment. It will be recalled that one criticism levelled at Murphy's version of retributive punishment was that preference satisfaction is not measurable. Either a preference is satisfied or it is not. Gewirth, however, links preference satisfaction to the right actually infringed. This is good moral philosophy because in ignoring V's right R, C acts as if he has a right at least equivalent to R (i.e. a right of sufficient size to be weighed favourably against V's right in a calculus that would allow C to act as

he did). Thus, C's satisfaction in acting as he did is proportional to the right overridden. It is important to note that the satisfaction discussed here is not the psychological one of enjoyment. In the nature of many crimes a considerable amount of self-loathing and regret may be involved. The satisfaction is rather in terms of the occurrent (i.e. at the time of the crime) ability of C to pursue a goal he holds to be at least sufficiently good enough to be worth acting upon. Whatever the undesirable consequences of this for C he acts in the instant of the crime so as to add to the stock of his own realised agency. As a matter of fact, it is to be hoped that the consequences of crime for C are undesirable because this would tend to be effective in deterring C and other contemplative criminals from crime. The point is that this hope is not in itself sufficient to justify bringing those bad consequences about. That would be utilitarianism writ large.

It is necessary to consider exactly what Gewirth means by an "occurrent inequality" between the criminal and his victim. In committing crime the agent forces the victim to lose at least two things. First, crime, being definitionally non-consensual, takes away freedom and freedom is something which the victim and the criminal both necessarily hold to be good as a precondition of agency. Second, the specific crime will take away some measure of well-being and/or freedom from the victim. Assault takes away health and thus full ability to pursue one's goals. Theft takes away the non-subtractive good of property. In either case the specific loss involved in the individual crime will involve an element of freedom and well-being and thus something to which both the victim and the criminal are committed, on pain of self-contradiction, to holding to be good. However, Gewirth looks to a third good as being the subject of inequality as between the criminal and victim. He says

[t]he equality of rights with which the criminal law is concerned is strictly limited in its scope: it extends not to all goods but only to the . . . *mutuality of occurrent noninterference*. (1978, 298; my emphasis)

All agents are equally entitled to the basic and other important goods. All agents are consequently equally entitled to have their basic and other important goods

left in tact. The former entitlement is generally the project of government which passes laws and sets up institutions to ensure that agents within its authority live in a state where rights under the PGC can be enjoyed. It is the infringement of the latter entitlement which the criminal law is designed to prevent and which punishment is designed to rectify. The occurrent inequality which crime sets up is that the agent-criminal has his basic and other important goods left in tact while acting so that the agent-victim has his basic and other important goods diminished. The equality is not perceived as one between the actual goods which particular agents enjoy but as between the mutual restraint enjoyed by each agent. This is important because it renders Gewirth's theory immune from the Marxist critique discussed in the previous chapter. It is not important that the victim is wealthier or stronger than the criminal. What is important is that he has restrained himself from infringing the criminal's basic and other important goods.

The occurrent good which the criminal is observed to have gained is that of setting and (at least partially⁸³) achieving an end which wrongly thwarts the agency of another. As with all such theories the argument to the justification of punishment proceeds in three stages: an equilibrium, an infringement and a restoration.

Agents possess a variety of different things, capabilities and goals which they hold to be good. The "basic and other important goods" protected by the criminal law are those goods which taken together form the generic conditions of agency. Society ought to operate such that each agent has these goods in equal measure because each agent necessarily claims them as rights under the PGC. However, the equality which the criminal law protects is not as to individual agents actually possessing these goods but equality as to individual agents not taking these goods from others. The equality which generally exists between two agents (but which is disturbed in the case of crime) is

⁸³ This is important because attempted crimes are punished as well as successful crimes. In many cases an attempted crime will thwart the agency of the victim to some lesser extent than the intended crime. Thus, an unsuccessful strangling is still an assault and attempted burglary will involve property damage.

in respect of basic well-being in the specific sense that neither is actually depriving the other or any third parties of basic or other important goods such as life, physical integrity, or good reputation. This means, in particular, that there is an occurrent equality of mutual restraint . . . (1978, 297)

Equilibrium exists between two agents where each forbears from infringing the basic and other important goods of the other. The second stage of the equilibrium-crime-restoration equation occurs when crime is committed. Crime occurs, and by definition the equilibrium is broken, where one of the agents infringes the basic or other important goods of the other. Punishment is justified by reference to its ability to restore equality as to occurrent non-interference. When addressing the problematic issue of how punishment achieves this Gewirth, as we have already seen, in line with Murphy states that punishment operates to remove "additional satisfaction" and "restores the original equality and prevents him [the criminal] from profiting through his violation of the mutual restraints prescribed by the *PGC*" (1978, 297).

There are, I believe two issues here, and the confusion of them has troubled many of the modern attempts to justify retributive punishment so far considered. The issues are on the one hand, the removal of preference satisfaction from the criminal and, on the other hand, the restoration of a pre-crime equilibrium concerning the victim's enjoyment of his rights. The confusion is acute because while retributive punishment seems well suited to the first issue it is not so obviously suited to the second issue. However, even though punishment is suited to removal of preference satisfaction it is its ability to restore the victim's enjoyment of his rights which intuitively, at least, makes it a justified practice. Certainly from the point of view of rights as supremely important, the restoration of the enjoyment of right is a good reason for harming people but the removal of an illicitly gained enjoyment is not so obviously a good reason for causing harm.

First, retributive punishment does seem well suited to the project of removing the additional satisfaction gained by commission of crime. This additional satisfaction takes a specific form in Gewirth's theory. Two agents mutually forbear from

infringing each others' basic and other important goods. One then breaches this forbearance and infringes the other's basic or other important goods while at the same time enjoying the forbearance of that other from infringing his basic and other important goods. This infringement constitutes the criminal's decision to lift a right restraint on his conduct. By imposing an equivalent restraint on him, punishment removes the "additional satisfaction" involved in crime. As we have already seen, Gewirth's theory gives us a clear notion of equivalence because it involves a strictly derived hierarchy of goods. The satisfaction involved in ignoring forbearance is equivalent to the right infringed and punishment thus removes an equivalent to the satisfaction by removing an equivalent right from the criminal. However, the problem with this, taken as a justification of punishment, is that it falls foul of Duff's objection considered in the previous chapter: it constitutes a begrudging attitude towards the criminal whereby we are jealous of his 'luck' in having more preference satisfaction than us. This objection holds so long as the equilibrium we are restoring does not refer to any considerations extraneous to the criminal. At the level of preference satisfaction the units of equilibrium are the agent's forbearance from satisfaction, the agent's taking of satisfaction and the imposition of 'preference frustration' to negate the taken satisfaction. As Kant states in his treatment of the spiritual dimension of punishment, it *may* be good if the evil suffer such that considerations internal to the criminal are sufficient to justify the imposition of harm on him. However, as Kant was equally strong to contend, the factors necessary for *us* to make such a judgment are not available. It was this realisation that caused Murphy to doubt whether punishment could be justified in retributive terms at all. Thus, we look to factors extraneous to the criminal. I contend that focus on mere preference satisfaction does not offer us any such factors.

This issue can be considered in another way, by asking why we should frustrate the preferences of the criminal. The PGC states that the freedom and well-being of each agent is to be equally respected. Punishment is a diminishment of a particular agent's freedom and well-being and as the PGC is the *supreme* principle for all action we can only look to the principle itself to justify such a

diminishment. This rules out appeal to the self-evidence of the statement that it is good for the wicked to suffer. Further, Gewirth, as we have already noted, states that any deterrent effect of punishment is strictly secondary to the justification of punishment in terms of its ability to restore the pre-crime equilibrium. Thus, it is not sufficient to state that the criminal law, carrying with it the threat of preference frustration, deters potential criminals and that punishment is therefore justified. From the perspective of the PGC deterrence is a good thing because every crime not committed is an instant of the principle being upheld. However, while the threat of punishment may legitimately be directed at every agent, to punish every agent for the crime of a particular agent would breach the PGC's applied injunction against treating agents as mere means to an end. Gewirth therefore turns to consideration of the fact that it is the criminal as opposed to any other agent who has broken the law against infringing the rights of his recipients. This factor is sufficient to tell us why, if anyone is to be punished, it ought to be the criminal who is punished. But on its own it does not tell us how punishment serves to uphold the PGC. Specifically, the deterrent moment of the criminal law has passed: the threat of enforcement has failed to deter the criminal. Thus, *in terms of deterrence*, it is impossible to restore any antecedent status quo. If all we do in the case of punishment is diminish the freedom and well-being of the criminal then we ought not to punish.

The type of punishment considered above aims only at maintaining the criminal's antecedent level of preference *dis*-satisfaction. In terms of the PGC this is an illegitimate use of state power. Perhaps it is here that the analogy with sport is helpful to Gewirth's argument but ultimately unhelpful in thinking clearly about the issues involved. Within sport 'punishments' or penalties exist to delimit the nature of the sport. In rugby union as opposed to Gaelic football, the ball may not be thrown forward. For this reason the former game includes in its rules certain retrograde steps to be taken against any player/team performing the forbidden action. By stating that agents have no option about partaking in a society which is governed by rationally justified rules and that therefore, agents are bound by those rules, Gewirth creates a similarity between society and sport. Having reached this

point it seems only natural if the governing body (i.e. the state) authorises umpires to impose penalties on those who don't play fair. However, in sport, as opposed to society, one side or participant is *supposed* to win and in this sense one side/participant is *supposed* to have its preferences satisfied and achieves this at the expense of frustrating the aims of the other side/participant. The structure of sport is, by nature, a series of opportunities for preference satisfaction/frustration. All that penalties do is define the ways in which one side/participant may go about its project of beating the other. The ultimate end of sport is not equilibrium, it is supremacy. This is not the case within society under the PGC because that principle, *inter alia*, imposes duties on agents to help others achieve a minimum level of basic and other goods.

If we look only at the criminal then the model of punishment is exclusively one of preference frustration. This is a legitimate end of such prophylactic mechanisms as the police force, burglar alarms or neighbourhood watch schemes, but this is so because in each of these cases the preference being frustrated is the preference towards a particular contemplated crime. In the case of punishment it is too late to frustrate *that* preference. All we can do is reduce the criminal's overall level of preference satisfaction to its pre-crime level. But all other things being equal we ought not to reduce preference satisfaction. Preference satisfaction is necessarily held to be good by each agent and is therefore to be maximized rather than diminished. The reason we frustrate the criminal's end by arresting him in the act, or putting insurmountable objects in his path, or frightening him, is not simply to prevent him achieving his goals but to prevent another agent (the potential victim) having his goals thwarted. Because this other agent's goals are legitimate the prophylactic mechanisms discussed are not problematic. They require no further justification. But unless punishment can be construed in terms of the preference satisfaction of some agent other than the criminal, all we have is the gratuitous imposition of harm so that from the perspective of the criminal's preference satisfaction he is no better off. It is at this point that the confusion between imposing a preference dis-satisfaction on the criminal and restoring the enjoyment of a right to the victim arises. If the former can be construed in terms of the latter

then we have a sound justification for punishment because the PGC necessitates that if V is entitled to something then he is given that thing even at the expense of an (equivalent or lesser) entitlement in another. It is my contention that Gewirth's notion of "occurrent equality of noninterference" fails to make the required conceptual link between the preference satisfactions of the criminal and the victim's right. It will also be my contention that Gewirth's theory contains all the necessary ideas to bridge the gap. First though, having shown where Gewirth's theory fails, it is important to demonstrate why I think this is.

The idea upon which Gewirth's justification of punishment rests is the occurrent equality of noninterference. I believe that he is at fault in confusing this with the occurrent equality of preference satisfaction. Although it is true to say that one can be *measured* in terms of the other it is not true to say that one is the same as the other. The relationship is one of equivalence rather than identity. Preference satisfaction for C can be measured in terms of the balance of noninterference between C and V for any action by C which affects V. This is so because the magnitude of the 'good' which C pursues in his action is only morally relevant in as far as it affects V. Thus, the moral weight of C's preference satisfaction is equivalent to the sum total of the goods vested in V which his action infringes. In terms of a Venn diagram the circle of C's aspirations and actions is morally judgworthy only in that section where it overlaps with the circle of V's aspirations and actions. Thus, the culpable portion of C's action (his illicit preference satisfaction) is equivalent in area to V's overridden goods. This area of overridden goods is equivalent to the balance of non-interference because V has not interfered with C's goods but C has interfered with V's. Therefore, any overlap is a balance of interference created by C and is observed in terms of the goods which C has ignored while relying on V's observation of his own goods. The balance of interference is the same as the overridden goods belonging to V which is equivalent to the morally culpable preference satisfaction of C.

This shows how the balance of (non)interference can be *measured* in terms of C's preference satisfaction. However, the point is that as they are not the same, the

undoing of one does not necessarily constitute the undoing of the other. Specifically, in terms of equilibrium the former can be undone and the latter cannot. This is because, conceptually the former does not depend upon the involvement of other agents. I can satisfy my preferences without involving any recipients of my action. In a hypothetical case, if I live alone on a desert island then none of my actions impinge upon the agency of others. In a more prosaic case, if I go for a jog then, all other things being equal, I can fulfill a purpose of mine without affecting any other agents and certainly without affecting them adversely. The preference satisfaction involved in, for examples, picking some fruit on a desert island and going for a jog can be offset, respectively, by starving me for an afternoon and, say, forcing me to eat an unhealthy meal or forcing me to do an extra half hour's desk work. The preference satisfaction of the original act and the preference frustration of the response are both measured in terms of the purposes of the individual agent. In these terms the balance of preference satisfaction can be rectified from the criminal's point of view, but as has already been argued this alone is a gratuitous diminishment of rights. In terms of an occurrent balance of noninterference and from the point of view of the victim, the dis-equilibrium cannot be undone by punishment as normally conceptualised. Certainly, given the starting point of this thesis, that punishment is something done against an offender, the disequilibrium from the victim's perspective is not undone by punishment. If anything it would be undone by allowing the victim a preference satisfaction equivalent to the goods denied by the original crime.⁸⁴ This is a problem we have already encountered considering the theorists in the previous chapter. Punishment restores the criminal to his pre-crime situation but not the victim. Within Gewirth's own terms of reference punishment can take from the criminal the preference satisfaction of crime but cannot take away from the victim the interference with his basic goods. Thus, what Gewirth is actually dealing with when he talks of the occurrent balance of noninterference is in fact an occurrent balance of the criminal's preference satisfaction. From the point of view of the

⁸⁴ Within a capitalist system it is possible to conceptualise punishment in this way. Every crime would be met by a fine, payable to the victim, which, insofar as is ever possible, remunerates him for the preference frustration of the crime. While this may work well for certain offences against property perpetrated by an offender of means, it is obviously problematic for other instances of crime.

victim this is unsatisfactory. In the previous chapter that conclusion was largely based on intuition but we now have principled grounds for saying why this is so. The PGC demands that the generic rights of all agents to freedom and well-being be upheld. Punishment justified by reference to the overall preference satisfaction of the criminal diminishes rights overall which I term gratuitous punishment. What the PGC demands, something which Hegel saw but failed to formulate properly, is that from the perspective of the victim within an ideal realm of rights (which is really no more than the state under the PGC) the imbalance must be undone. The wrong in crime is not that the criminal has enjoyed himself, it is that his enjoyment has come at the expense of another's right. Punishment, to undo the wrong, is not aimed at taking back the criminal's enjoyment but at restoring the position of the wronged agent. I now turn to a justification of punishment within a Gewirthian framework which does not rely on confusion between the occurrent equality of noninterference and the occurrent equality of preference satisfaction.

2 A Gewirthian Justification of Punishment

Having criticised Gewirth's own theory of punishment it now remains for me to offer a justification of retributive punishment which utilises Gewirthian notions of right. To achieve this I will suggest a different equilibrium to the one employed by Gewirth himself. I will then go on to synthesize this Gewirthian analysis of right with the insight made by the theorists considered in the previous chapter that the relevant parties to crime are the criminal and the other members of society. Finally I will show that in these terms it makes sense to say that punishment undoes disequilibrium.

2.1 An equilibrium of dispositional interest in the value of right

I hold that Gewirth's mistake is to focus on an equilibrium which cannot be maintained by punishment: the equilibrium of occurrent noninterference. The equilibrium which I believe is the proper subject of punishment and which I believe punishment is suited to restore, is an equilibrium of dispositional interest in the value of right.

Before going any further it is necessary to re-consider the Gewirthian concepts of *good* and *right* because much of the following discussion will be couched in these terms and it is important to demonstrate the relationship between the two. It will be recalled that Gewirth's derivation of the PGC moves from the notion of goods to the notion of necessary goods to the notion of rights. A good is whatever the particular agent values as being good. There are no specific criteria beyond this by which good is to be judged. This entails that from the point of view of an agent something can be good albeit that from every other point of view it might be bad. Necessary goods are those things which each and every agent must hold to be good for the reason that they are at least instrumentally necessary for the achievement of any other goal which he holds to be good. Rights are these necessary goods when expressed as a demand against all other agents that the goods be left in tact. Thus, a right is a way of expressing a claim to the content of a necessary good. It should be noted that the purpose of a right is to protect something which the agent necessarily holds to be good as instrumental to agency. I contend that the equilibrium to be maintained by punishment concerns the *value of a right* to the agent. That is, its value as a claim to protect the generic features of agency.

2.1.1 The notion of "value"

Consider the following three situations: an anarchic state in which gangs of brain-washed children roam the land killing anyone who does not belong to their group; a liberal state in which gangs of opportunist delinquents hang about waiting to mug passers-by for their money; an ideal state in which citizens respect the rights of all other citizens. For the purposes of argument, in the anarchic state the right to physical integrity is not given as law; in the liberal state it is given as law but not actually recognised in every case; and in the ideal state the right is given as law and actually recognised in each morally relevant situation. In each instance a particular agent has the right to physical integrity: the rational necessity of this is absolute and trumps the empirical contingency that in two of the hypotheses the right is not recognised. Now suppose that an individual is placed in each state in

turn and told to walk from one point to another, ten miles away. The relevant question does not concern whether he has the right to physical integrity but concerns *the value of that right to him*. The value, in this instance, refers to the extent to which the agent can rely on his right to affect the way others act towards him. It is more valuable the more he can rely on it to protect his agency related interests. It ought to be the case that regardless of any other consideration the right is sufficient to protect the physical integrity of the agent. The right is progressively less valuable to him the more he is forced to rely on other considerations such as his guile, speed, physical strength, powers of persuasion, friends etc. In the first hypothesis the right is of no value to him whatsoever.⁸⁵ The fact that he has the right will make no difference to his success in walking ten miles through such an anomic country. Given that risk is posed by brain-washed minors even the chance of persuading them that rationally they ought to respect his right is nil. In the second situation the right is of limited value to the agent. He will be able to rely on it to the limited extent that some people will be deterred from offending by the fact that the right is recognised in law. Further, secondary mechanisms such as the police, security cameras, education programmes or the presence of law-abiding citizens may operate to ensure that individuals who would otherwise mug the agent choose not to.⁸⁶ In the third hypothesis the right is of full value to the agent because he need rely on nothing else to ensure that his walk is uninterrupted by assault. The equilibrium interrupted by crime is the equilibrium of the value of a right as shown in these hypotheses. This is a two fold concept of value, consisting of both efficacy and substance. The right is of value by being effective in protecting the agent's good. It is also of value insofar as it contains a certain substance which will place it within a hierarchy of rights. A

⁸⁵ Even if he has the mental comfort of the moral high-ground (the knowledge that whatever harm is done to him is, in the greater scheme of things, *wrong*) the right is of no value to him *in protecting his physical integrity* which is the purpose for which he should be able to rely on it.

⁸⁶ In practice the value of rights are tied up with the secondary mechanisms which are put in place to protect them. It is therefore no objection to the example just cited to say that the agent is not relying on the right but on the police or the law-abiding etc. The point, as my argument unfolds, is that in order to have *bona fides* concerning right, the state must set up such mechanisms to protect the right. The mechanisms are not ends in themselves: the justification of the police force, for example, is that it in fact serves the end of protecting right.

right of a particular worth is of value because of its ability to trump any right of a lesser worth and this is what I refer to as the value of the substance of the right.⁸⁷

2.1.2 Crime as an attack on the dispositional interest in the value of right

Crime may be regarded as attacking right in an occurrent sense. The occurrent infringement is the particular good which the crime attacks, whether this be the victim's physical integrity, mental welfare, reputation or property. It is true to say, as Finnis and many others say, that this harm often cannot be undone. One cannot be un-beaten and even in the case of theft, where stolen goods can be returned, this is hardly to make the situation as if the crime did not occur. The victim is still deprived of his chattels for a while and is still put through the mental harm of experiencing the theft and worrying about how he will cope if his things are not returned to him.⁸⁸ It is at this point that Gewirth turns to the notion of the occurrent good of noninterference, but as I have argued above this cannot be restored in a meaningful sense either. What is restored by punishing C is not noninterference but the criminals antecedent level of preference satisfaction. While the victim of an offence has a very real interest in not having his rights interfered with it is not clear why his interest in seeing that the criminal's overall level of preference satisfaction is unaltered, should be relevant.

Gewirth himself makes the distinction necessary to give a justification for punishment. This is the distinction between a good/right in its particular-occurrent and generic-dispositional modes (1978, 58-59). The particular-occurrent mode of good/right is the specific end to which the individual aims his agency.⁸⁹ The generic-dispositional mode of good/right is the background level of freedom and

⁸⁷ In practice the efficacy and substance of a right are very closely related because particular right R is only fully effective where it is given its full substance i.e. where it can outweigh rights of a lesser value.

⁸⁸ It is nevertheless true to say that insofar as the occurrent harm of a crime can be undone it ought to be undone and that the criminal is the one who can be properly expected to do this. If a particular antecedent state wherein V does have his car and is not beaten up is right, then it must also be right to return, as far as is possible, to this state once it is wrongly disrupted.

⁸⁹ Nb not every specific end which the agent regards as good (subjectively) can be claimed as a right. However, talking in ideal terms the particular legitimate end of an agent is regarded by him as good and defended by a right to pursue particular ends so far as is in accord with the generic rights of other agents.

well-being necessary to pursue any specific end. The generic-dispositional good of property gives rise to an individual's particular-occurrent good of owning a car and the concomitant right not to have it stolen. For my own theory, the value of a right is its ongoing or dispositional ability to inform the way in which others act towards the right holder. The equilibrium to be restored involves not the occurrent right of V to his car but his dispositional interest in controlling his own property by his unforced choice.

Discussing the generic-dispositional aspect of goods Gewirth states that goods viewed in this manner

consist in the general conditions and abilities required for fulfilling any . . . particular purposes. In this generic-dispositional view, the emphasis falls not on particular goods for which the agent occurrently acts, but rather on the general necessary conditions that enable him to act for any purposes he might have and that he would hence regard as generically good. (1978, 58)

What we have here is a distinction between the good of particular ends and the good of being able to set and achieve ends generally speaking. The former is an expression of the latter; it is the dynamic mode of the underlying capability. As we have already seen, the derivation of the PGC proceeds from the true claim that without the generic goods of freedom and well-being no other goods can be secured. It is my contention that crime not only attacks these other goods but also strikes at a portion of the generic good. Crime does this by setting a precedent. It takes the empirical world of the victim from the antecedent security of rights to the post-crime insecurity of those rights having been violated. The right to ownership over a piece of property not only guarantees that the owner, and no other, has control over the enjoyment of that property, it sets up a series of legitimate expectations which the owner can take into account when calculating his future actions. For example, he might take out a bank loan secured by the property, or he might decide not to pay into a pension scheme because he can sell the property in his old age, or he might decide that he will dedicate next Tuesday to enjoying the piece of property, whatever it is. Similarly the right to physical integrity means not only that at any given instant an agent can expect to be free

from attack it means that he can plan ahead to do other things. If he decides to go for a walk he aims not at being physically unharmed but his aim, of getting healthy, or enjoying the scenery, or getting from A to B, relies on it being the case that he remains physically unharmed. The right to physical integrity gives him legitimate expectations of what will happen (or not happen) if he goes for a walk. This right is therefore taken for granted in much of the agent's everyday activity. The precedent set by crime is that the expectations the agent held as a result of his particular right are no longer *empirically* legitimate. As Hegel claims, crime causes a disjunction between the empirical and moral realms. This is because right entails the expectation (backed by a legitimate claim) that from now until I give up the right the world will operate in a particular way towards me. This expectation, even tangentially, informs other decisions which I make as to my future. It is almost impossible to conceive of everyday transactions in modern society without the background assumption of respect for property rights or rights to physical integrity. Crime throws this background assumption into doubt. It sets a precedent in the mind of the victim for the claim that his rights will not be respected and that therefore he must not act on the assumption that they will. This not only introduces a high level of insecurity into the actions of the victim but also limits the goals that he can pursue. In an extreme case he may feel unable to leave his house if he believes that leaving his property unattended will invite interference with it. To adapt Gewirth's own phrasing crime viewed in this way poses a grave threat to any stable transactions (1978, 300).

The equilibrium to be maintained can also be regarded as a balance of risk and certainty in the mind of the victim. In any given state there will be a number of contingent factors which weigh on the mind of an agent in assessing the likelihood of success in undertaking a project, even if this project is no more than going about his everyday life. Not all of these contingent factors are "bad" in a morally relevant way. It might be the agent's own weakness which prevents his success, or his greed, or over-ambition, or it might be that his project is itself wrong and that others will therefore justifiably seek to thwart it. However, crime, and the fear of crime, *wrongly* add to the unlikelihood of a project's success. They detract

from the likelihood of a right serving to protect some of the interests of the agent as he goes about his business. But, and this is an important link in the argument, the security of knowing that a right will operate in this way is itself a generic good. It is part of the background psychological well-being necessary for the pursuit of any of the particular goals of an agent. It is necessarily good, from the point of view of agency, that the agent operate in a context where the risk of not achieving his (legitimate) goals is minimised (and conversely the certainty of achieving those goals is maximised). Crime not only thwarts a particular-occurrent goal of the victim, it sets a precedent which has the tendency, left unchecked, to threaten the successful achievement of goals in an ongoing or dispositional sense. It diminishes the agent's future chances of pursuing his own ends by introducing an element of insecurity concerning the value of the particular right infringed. Where his right is ineffective, the agent is forced to fall back on other means of protecting those things which he holds to be good and this enforced activity, in turn, detracts from other projects to which he may wish to dedicate himself. Gewirth specifically includes in the generic goods "mental equilibrium and a feeling of confidence as to the general possibility of attaining one's goals" (1978, 54). This feeling of confidence is kept at a certain level where the agent considers his rights as having an effect on the way others act towards him. It is diminished in a state where the agent's rights are ignored because reliance on rights is itself constitutive of successful action.

In summary, on the basis of the rights claim to the content of a necessary good an agent has legitimate expectations of the ways in which others will act towards him. Crime constitutes a failure to act towards the agent in the expected fashion. This has the consequence that a particular occurrent good of the agent-victim's is thwarted. This thwarting cannot necessarily be undone but insofar as it can be the state ought to act to bring this about. Crime also has the consequence that it throws the agent-victim's right into doubt in an ongoing or dispositional sense. It is an action predicated upon a maxim which is incompatible with each agent's necessarily made claim "I have rights to the generic features of agency". At each point in time after the commission of the offence the agent-victim is dedicated (on

pain of contradicting that he is an agent) to the claim "I have right R", R being the right ignored by the criminal. The criminal's activity is predicated on a maxim which entails "the agent-victim does not have right R". So long as the situation is one in which these two incompatible claims are met with indifference the agent-victim's ongoing or dispositional interest in R is insecure. It will be my contention that the role of the state is to restore the agent-victim's confidence in the value of his right. First though, I will show that the equilibrium (of dispositional interest in the value of right) is to be measured in terms of the overridden right.

2.1.3 In what "units" is equilibrium to be measured?

I have already contended, in considering Gewirth's own justification of punishment, that the magnitude of a crime is to be measured in terms of the harm done to the victim's good. This is because the action of the criminal is only morally judgworthy (i.e. any business of ours) insofar as it tends to affect the freedom and well-being of another agent. However, there is an initial distinction between my own theory and Gewirth's because the subject of the equilibrium is different. The disturbance which crime seeks to rectify in my theory is a disturbance in the dispositional value of right. It must now be shown how this disturbance is to be measured. Of course, it must also be acknowledged that to talk in terms of "units" is artificial. What we have are a series of increasingly accurate proximations. Gewirth's theory is well suited to this because it does provide us with a developed hierarchy of goods.

The important issue here concerns the insecurity introduced by a particular crime. Take as an example, a criminal who snatches someone's bag which turns out to contain minor valuables such as petty cash, address book, credit card, cheque book and other such items. The victim is probably traumatised to a slight extent, worse off to a slight extent and inconvenienced to the extent of having to cancel his cards and wait for replacements. The magnitude of these harms is not definite: the trauma, the hardship and the inconvenience will be relative to the mental fortitude, wealth and ability to cope of the particular victim. However, within relative and general parameters these hardships can be measured. The parameters

are relative to other harms. So the trauma is much less than seeing one's family slaughtered, less than being held up at knife-point but more than having "Liverpool F.C." graffitied on one's wall. The parameters are general because it is impossible to measure precisely, in this case, trauma. We may know that trauma 1 is less than trauma 2 and greater than trauma 3, without knowing the exact "distance" between two traumas on a scale. This issue is one for the power of empathy on the part of those agents acting as judges. If the victim suffers an unusual reaction to an offence, such as a heart attack, or malnourishment because he was in fact very poor, then the criminal can be held responsible for these additional harms insofar as they were foreseeable, bearing in mind that to a slight but definite extent it is always foreseeable that a reckless exploit will go awry. The point for the moment though, is that to a general extent, and relative to other harms, the magnitude of harm done by a particular crime can be measured.

It is my contention that the unbalance in the value of right is equivalent to the magnitude of the harm done by crime. This is because in making calculations about future projects the agent will have to take into account the additional factor of crimes like the one experienced. By taking away the security of an occurrent right of a particular weight the criminal forces the victim to question the ongoing value of such rights. If the victim has his hubcaps stolen and nothing is done about it he will thereafter find the protection offered by his rights to minor property less valuable in pursuing his agency. Those rights have been devalued. The imbalance, from the point of view of the victim, in the value of his right is equivalent to the amount of his right(s) which has been breached. This can also be seen in terms of what I have called the value of the substance of a right. Particular right R ought to be effective against all rights less important than it within Gewirth's hierarchy. By calling it into question the criminal has, as it were, removed from the victim a trump card of a particular value.⁹⁰ To extend this metaphor, the victim is put in a position where it is as if he can no longer rely on the eight of hearts to beat lower cards. Thus, the loss which the victim has experienced (that which punishment is

⁹⁰ In fact, as I have already pointed out, he has not removed the "card" (i.e. the right) but he has made the victim believe that it has been removed, which, from the victim's agency related point of view, amounts to the same thing.

to rectify) is equivalent in value to the particular right which the criminal has overridden.

2.1.4 The move from the particular victim to Potential Future Victims (PFVs)

It could very well be objected to the theory so far presented that the one crime which is rendered immune from punishment is the most serious of all, namely murder. In this case there is no surviving victim with an ongoing dispositional interest in the infringed right (not to be killed against one's will). Therefore, as there is no meaningful sense in which the equilibrium of the victim's dispositional interest in the right not to be murdered can be restored, the murderer ought to go unpunished. It might also be objected that in measuring the disequilibrium as set out above, the units offered are too dependent on the subjective make-up of the particular victim. However, the theory is safe from such objections. So far I have presented it in terms of the binary state of a criminal and his victim. This is for ease of discussion as much as anything else. It will now be shown that each and every agent, as a potential future victim (PFV), has a dispositional interest in the criminal's punishment. If this is the case it does not matter (for the justification of punishment) that as an empirical contingency murder does not have a surviving victim.⁹¹ Further, the peculiar circumstances of the victim are less important in weighing up the harm done. The relevant measurements are the harms as viewed from the perspective of agents generally.

That which punishment defends against is a state of uncertainty as to the practical importance of rights and the correlative diminishment in the agent's ability to plan ahead and pursue his agency. The criminal creates a situation in which the background security necessary for the voluntary pursuit of self-set goals is

⁹¹ This also has the consequence that it does not matter, for punishment to be *justified*, if the immediate victim of the offence forgives the criminal. This was a problem identified in Morris (see ch. 4 fn. 65). Of course, in practice it may be hard to successfully prosecute a criminal where the victim does not wish to offer evidence against him. Nevertheless, it remains true that all PFVs have a dispositional interest in the criminal being punished. This in turn may make sense of the practice in the English legal system of prosecuting victims who refuse to give evidence when the case comes to trial. This is not the rationale offered within the English legal system and so I will leave this as no more than a suggestive pointer.

reduced. A particular crime will reduce the background security to a greater or lesser degree. The knowledge that one's washing may be taken from the line will be less pressing than the knowledge that one stands to be brutally beaten. The state of knowledge in the former example will therefore affect a narrower and less fundamental range of goals that an agent might wish to pursue. The point though, is that the range of goals affected is not confined to the particular right infringed; nor is it confined to the particular victim. In the first example just given, the theft of V's washing will not merely affect V's decisions relating to his washing, nor will it merely affect *his* decisions. The fear raised in his mind will be as to his wider (albeit minor) property interests and any other agent with similar property interests will have similar doubts as to the efficacy of his rights. Even if the crime is one of the personally directed victimisation of V as opposed to anyone else, any other agent can put himself in V's position and think that, if he offends someone's sensibilities, he might end up the casualty of a hate campaign. This would then define its own set of beliefs with a correlative diminishment of a range of goals. In the second example, where V is brutally beaten up he will be insecure concerning his physical integrity in general and anyone hearing of his misfortune will be similarly concerned. The thing to which V is opposed is a diminishment of his agency because of a diminishment in the range of goals he can successfully pursue. This is an agency specific situation and is therefore shared by all other agents by application of the principle of universalizability. The good of pursuing ends voluntarily and purposefully is held to be a good and therefore a right by each agent. A reducing of the good in one particular agent is a reducing of the good in all agents. In this case, the interest of V in C's punishment is shared by all PFVs.

However, it might still be objected that if other people do not *know* about the crime then they will not be subjected to the fear of suffering like experiences. But this is not the crucial issue. *As a matter of fact*, in a state which does nothing in response to crime an agent is less able to rely on his right as a protector of his agency. It might be that through luck, or strength, or some other personal quality an agent manages to walk ten miles through the anomic territory discussed in one

of the examples above. It might even be that throughout his walk he is under the misapprehension that anyone he encounters will respect his right to physical integrity. But as a matter of fact, in this hypothetical state, his right offers him no protection from attack. And the state *ought* to operate so that rights can be relied upon to protect the generic features of agency: it ought to move towards becoming the ideal state governed under the PGC in which rights work in this manner.

It could be objected now that the potential criminal is not actually respecting PFV's right under this model. All that happens is that through fear (of punishment) he acts as if he actually respects PFV's right. This issue is not important, however. The use of the word "respect" in this context does not necessarily import a conative attitude on the part of the agent. For the purposes of right and hence agency, it is sufficient that the potential criminal despises the right of the criminal so long as none of his actions manifest this attitude. This is sufficient because it entails that the PFV's right is respected in the passive sense albeit that for the potential criminal to be rational he must also respect PFV's right in the conative sense. The state under the PGC fulfills its primary functions if it ensures that the rights of each agent are respected in this passive sense of non-interference. This is really no more than a modern reworking of Kant's distinction between *recht* and virtue. *Recht* is the state of affairs where the external actions of agents conform to morality. Virtue is the state of affairs where the external actions of agents conform to morality for the correct reasons. While it is true to say that under the model so far presented the potential criminal doesn't necessarily respect the victim's right in the conative sense, he does respect it in at least the passive sense of refraining from infringing it and this is sufficient for right to be upheld under the PGC. Moreover, we are concerned with the perspective of the PFV and from his point of view it is as if his right were respected in the conative sense. The mental state of the criminal or any potential criminal is noumenal to us (and to PFV) and therefore, in this particular regard, of no concern.

Finally, I must consider how this development in my justification of punishment avoids Duff's objection that the wrong involved in crime is not any relative diminution experienced by the law abiding but the very real violation experienced by the immediate victim (Duff 1986, 212 cited above in chapter 4 section 1.3.3). I concede that the wrong involved in crime *is* the violation of right experienced by the victim. However, this violation has dispositional implications not only for V but for all other agents, i.e. all PFVs. The loss experienced by PFVs in my theory (as opposed to the loss experienced by the law abiding in the benefits and burdens theories considered in the previous chapter) is not relative; they actually lose a dispositional interest in right because a particular right becomes less reliable for them. To employ the example which Duff uses, the wrong involved in rape *is* that it constitutes a gross violation of the victim's personal integrity. But the wrong with which we are concerned is that this vile act throws into doubt interests shared by all agents; the interest which punishment seeks to restore is one vested not only in V but also in each and every PFV.⁹² To employ tabloid speak it is not only V who feels unsafe walking the streets at night after an attack. This is not to say that each PFV undergoes the same infringement of right as V himself. No matter how wrong that infringement is, it cannot always be undone, though where it can be it ought to be. But there is still the issue of the dispositional interest in the value of the right (i.e. its ongoing reliability in protecting the actions of the agent) and this is shared by all PFVs. This is really no more than Hegel's point that the particular right and the realm within which it resides are inextricably linked. V's dispositional interest in the right is one and the same as PFV's dispositional interest in the same right.

The point being made in this section as to PFVs is really no more than an extension of Gewirth's point that some interests are generic to agency *per se* rather than to particular agents. The objection that murder should not be punished

⁹² In a different but related context Burgh claims that the specific right not to be raped is not one shared by all agents: only women have this right (Burgh 1982, 205). Even if this were the case (which it is not), the point is that each and every agent has the generic right not to undergo violent and degrading acts against his will. Although it *may* be that a particular application of this right is only appropriate to a sub-group of the class "agents", the generic right from which the application is hewn holds for the entire class. Therefore, it is correct to say that the dispositional interest in this right is shared by all PFVs.

holds if we focus only on the occurrent harm to the victim. This occurrent harm is done to the victim and no other. The dispositional harm is not so limited. Each agent has an interest in punishment because each agent has a like right to the victim. Certainly, each agent does not have the victim's right to own the victim's car, for example. But each agent does have the generic right to own property. And it is the right viewed in this dispositional sense that punishment seeks to defend.

2.2 How does punishment function to maintain the equilibrium?

Punishment is concerned with the disjunction between claims which the PFV necessarily makes about himself and the way the world acts towards him. Central to this disjunction is the fact that there can be no reconciliation between an agent saying "you must not do X" and a society of agents which continually does X. This disjunction is seen even more acutely when it is realised that with regard to rights, each agent is bound, on pain of contradicting that he is an agent, to concur with the claim "I must not do X": there can be no reconciliation between an agent claiming "I must not do X" and his doing X. The claim is prescriptive and evaluative regarding the doing of X. It binds the agent to a particular frame of mind concerning X. Inconsistency lies not in any *logical* impossibility of doing X and saying that it is impermissible; of course it is logically possible to say one thing and do the opposite. Rather the inconsistency lies in the *logical incompatibility* of simultaneously doing X and saying that it is impermissible to do X. The doing of X is not some random occurrence. From the point of view of agency it is an action chosen by the particular agent and manifesting a conative attitude towards the doing of X. It is not possible for the agent to do X without at least valuing X as good enough to be worth doing. This cannot be squared with the claim that "I must not do X" because that claim entails the statement that "from the point of view of agency it is necessarily good that I do not do X". We thus have two contrary mind sets. One, from the point of view of a contingent desire to achieve an end, states "it is good that I do X"; the other, from the point of view of non-contingent agency, states that "it is not good that I do X". These two claims are obviously inconsistent with each other. Inconsistency does not in

itself imply immorality but given that the point of view of agency is epistemologically prior to the point of view of contingent desire the second statement must take priority over the first. When I talk of epistemological priority, I mean that desire is a feature of agency rather than the other way round. Therefore, any claim which is shown to be true from the point of view of agency must also be true from the point of view of desire. This doesn't mean that the agent actually desires to do "not X"; in the hypothesis he explicitly desires to do X. However, it does mean that he must assent to the truth of the claim "it is not good that I do X". The agent, as an agent, is necessarily committed this statement in any event. (On this issue see also Beyleveld 1991, 106-108.)

Punishment is the defence of the antecedent and ongoing legitimate expectations of the PFV. As has already been argued the PFV's expectations remain morally legitimate throughout. What punishment seeks to rectify is an attack on the empirical or practical legitimacy of the PFV's expectations. The PFV demands that the criminal be punished by way of defence of his ongoing dispositional interest in the particular right. PFV's right to physical integrity does not lapse with an attack: he has a continuing expectation that he will not be assaulted. It is C who has thrown into doubt how realistic this continuing expectation is, and it is therefore against C that the right (dispositionally viewed) is defended. Punishment restores to PFV, insofar as is possible, the antecedent security that his right gave rise to. This security is founded in legitimate expectations as to how the world will act towards him. Indeed, once punishment is shown to be justified, punishment for infringement *becomes* one of those legitimate expectations. Punishment does not make the situation as if C had not infringed his right but it does make the situation one in which the right is treated with a degree of seriousness and in which a precedent is set that any attack on the right will be met with defence. This allows PFV to plan ahead with a renewed degree of confidence in the practical relevance of his right which would not be the case if every breach of the right was met with indifference. It is recognition of the fact that rights do have this double aspect both, occurrent and dispositional, that allows punishment to be a proper defence of the right rather than a gratuitous *post hoc* revenge. As

the right has ongoing importance to the PFV, once diminished it is diminished in an ongoing sense. In these terms punishment can be viewed as analogous with self-defence: it is a reaction to an ongoing wrong (initiated by a past crime) rather than a purely backwards looking process.

So far I have talked generally of PFV's dispositional interest in right because I believe that this is the correct focus for restoration by means of punishment. Now though it is important to delimit precisely what it is that is at stake in a crime in order to demonstrate that punishment is suited to restoring the pre-crime equilibrium.

2.2.1 Strict derivation of the State's duty to punish

1 On pain of contradicting that he is an agent PFV makes statement S: "C must not attack R".

This much follows from Gewirth's derivation of the PGC. Each agent is committed to the claim that he has rights which necessarily entails that he is committed to the claim that those rights be left unmolested.

2 R is a claim to necessary good E.

This follows from my earlier consideration of the relationship between goods and rights. A right is the correlative of a strict ought claim that the all other agents refrain from interfering with PFV's necessary good.

3 E is of magnitude G where G is a relative measure of the importance of a good within a hierarchy and where importance is measured in terms of how instrumentally necessary E is to secure the generic features of agency.

The rights to freedom and well-being are absolutely necessary to secure the generic features of agency because freedom and well-being *are* the generic

features of agency. However, when it comes to deriving a practical hierarchy of legal-moral rights some rights will be seen to be of more importance than others. Thus, while the right not to be assaulted is an absolute human right, it is not as important as the right not to be killed. Being killed is a bar to agency in a way which an assault is not. Certainly, while the agent is being assaulted his freedom and well-being are being denied but he has the potential to regain his freedom and well-being in a way that the murder victim does not. For the same reason a brutal beating infringes a more important good than a slap in the face albeit that both assaults deny freedom and well-being. Similarly, within the category of well-being there is a hierarchy of goods from the basic to the non-subtractive to the additive. This entails, for example, that the goods of shelter and food are more important than the good of owning a CD player, even though you would deny me a good (and infringe a legal-moral right) if you stole my CD player.

*4 Harm (the denial of good) of magnitude H may be perpetrated
in defence of E, so long as $H \leq G$*

The criminal sets up (i.e. is responsible for) a situation in which V's right is not observed. In this situation it may be the case that all V can do to defend his good is attack C. He may only be able to stop him burgling his house by fighting him off or he may only be able to stop C killing him by killing C. This is permissible because it is C who has brought about this situation where only one agent's goods can prevail: either V lets his rights be trampled or he hits back. In this situation it is not arbitrary for us to prefer the interests of V to those of C because it is C's fault that any incompatibility arose in the first place: he ought to have chosen not to attack V's goods. However, because of the hierarchy of goods and the consequent hierarchy of rights it would be wrong for V to impose on C a harm greater than the harm with which C threatens him. Thus, he cannot kill C to protect his car.

5 In committing crime, C acts on the statement -S: "It is not the case that C must not attack R"

This much follows from the argument offered above, that doing an action necessarily entails a conative attitude towards that action whereby it is at least good enough to be worth doing. C cannot have it that his action is worth doing and that it is good that he does not do it.

6 In attacking R (V's interests) he removes E of magnitude G

That which constitutes crime is the non-consensual removal from the victim of something which he necessarily holds to be good.

7 E has two aspects: the occurrent and dispositional; the occurrent aspect of E is removed in that during the attack PFV does not enjoy E; the dispositional aspect of E is removed in that during and after the attack -S implies that it is okay to attack R and hence remove E.

When I claim that E is a necessary good and that I have a right to it, this entails certain expectations of how others will treat that good. Most pertinently, if I have a right to it, then it is the case that others ought not to interfere with it. My possession of E has implications for the state of affairs 'right now' and also for my ability to pursue my agency henceforth. E, as a necessary good, is instrumental to agency *per se*. In altering the state of affairs 'right now' the criminal takes away from the victim a certain extent to which he can rely on that good in pursuing his agency in the future. The victim is faced with the prospect that even though he necessarily makes the claim S others act on the claim -S. Therefore, the claim itself (the right) has no practical force.

8 The state is that mechanism which ensures that the dictates of the PGC (including S) are given effect.

This much follows from the definition of the legitimate state as considered in the introduction to this thesis.

9 The term "may" in 4 exists because PFV may waive S but it is definitionally not the case in the instance of crime that PFV has waived S.

It is always open to an agent not to defend his good or indeed to grant another agent permission to ignore it. In both cases he waives his right (in the first instance implicitly, in the second explicitly⁹³). However in the instance of crime the victim cannot have waived his claim to the trampled good because otherwise a crime would not have been committed. This much follows from the definition of crime as the deliberate breach of a legal-moral right. Where a right is waived it cannot be breached.

10 From the perspective of the state defence of R is not optional but mandatory because the state must not allow -S.

Because V has not waived his right the state is charged with doing something about the breach (Hegel earlier brought us to this same point). Furthermore, the aspect of the right which we are now considering (the generic-dispositional aspect of the right) is shared by all PFVs and it is certainly not the case that all PFVs waive their right in perpetuity. The state must not allow that "it is not the case that C must not attack R". It must restore the force of R and hence of the claim S. It achieves this by restoring to PFV the generic-dispositional aspect of his good. This is the ability to rely on his claim to that good (his right) to a certain extent to inform the way that others behave. This behaviour includes acting towards the

⁹³ NB this is not to say that just because V offers no resistance to attack he is waiving his right not to be assaulted. All I am contending is that an agent can waive his right either by manifesting a certain attitude towards that right or by explicitly saying "I waive my right".

right so that it is given a certain position within a hierarchy of rights. For the state to fail to give the right this importance would be a direct contravention of the PGC and this is impermissible because the state derives its legitimacy from that principle.

11 *State defends R by inflicting H (= -G) on C; anything less concurs with -S by failing to give E its proper worth, G.*

The harm inflicted on C is first *permissible* because of the argument under point 4 above and thence *mandatory* because the state (unlike V) cannot grant a permission that -S. The infliction of harm defends R by giving R its due worth and by making it effective for PFV. Punishment gives R its due worth by defending it to the extent that it has value against other goods. This is not merely some symbolic display of how important the right is. *Without* defending it to the correct amount the state effectively concurs with -S and by its actions states that PFV does not have the right.

From the point of view of the state punishment is a necessity. I have already argued that from the point of view of both PFV and the criminal the statements involved in claiming a right and in the commission of crime are incompatible. The state, which is charged with bringing the relationships between individual agents into line with the PGC, cannot tolerate this situation. Just as the criminal in committing his offence makes the claim that "PFV does not have R" so the state that fails to respond to crime makes the claim "PFV does not have R". In punishing C, the state acts decisively between PFV's claim "I have R" and the ongoing force of C's claim "PFV does not have R". That claim has ongoing force precisely because so long as the state does nothing, and therefore acts with complicity regarding the claim, PFV cannot act with any reliability on his own necessarily made counter claim. Punishment is a defence of PFV because it manifests an evaluative attitude on the part of the state towards PFV's right. It gives it its due importance in defending it to an exact extent. Furthermore, it is itself a statement to the effect that, given that it cannot be true that "PFV has R"

and that “PFV does not have R”, the state will prove that “PFV has R” is the case by defending the claim against the actions manifesting the counter-claim. R is restored to PFV because its value to him (i.e. its reliability as a protector of agency) is restored when he can plan ahead in the knowledge that R will be treated seriously by the state.

It might be objected that punishment viewed in this way does not do enough because it cannot restore full certainty to the PFV and there are aspects of the psychological harm, particularly in the more vile crimes, which will never be undone. The correct response to this is to acknowledge the truth of the objection but to hold that nevertheless punishment is as much as the state can do to undo that harm and that the state must do whatever is within its legitimate power to restore the equilibrium as discussed. The psychological trauma caused by crime, while serious, is not the aspect of harm with which we are directly concerned. It could be that psychological trauma prevents the immediate victim (and even his friends) from ever operating with the same degree of confidence in his right. However, from the perspective of PFVs, what is important is that the right is seen to be worth protecting to a set extent. This at least allows that in making everyday calculations the right has a certain dependability. We have here a distinction between the contingent characteristics of the particular victim (*viz.* that he is affected in a certain detrimental way by the crime) and the necessary characteristics of the PFV (*viz.* that he is committed to his right and demands that it be given protection). Of course, it would be wrong to belittle the contingent harms done to the immediate victim of an offence but their undoing is not the purpose of punishment. Much confusion (and much frustration with the attempt to justify punishment) is caused by the supposition that this is punishment’s function. It is human nature, which is a contingency, which prevents a right ever having full dependability. It must be incorrect to conclude that because the state cannot fully undo the harms of crime it ought not to undo them at all. A common objection to punishment is that it doesn’t actually undo the harm of crime because that harm is in the past. The theory I have presented here shows in what way it might make sense to say that punishment has a legitimate function in undoing the harm of

crime, by drawing attention to an aspect of the crime which is *not* in the past. The reason that punishment is justified is that it restores to PFV the value of his right as a protector of the generic features of agency. Punishment is instrumental to the efficacy of right and under the PGC it is imperative that rights be upheld.

2.2.2 The requirement of proportionality between punishment and offence

The reason proportionality between the punishment and the offence is necessary lies in consideration of what it is that is being defended by punishment. This is the value of right. As I have already argued, Gewirth's theory is good at offering a substantive hierarchy of rights. Within this hierarchy, rights have relative values from the most important rights not to be murdered or committed to slavery, to the less important rights to have one's reputation kept in tact and have one's property respected. An offence attacks one or more of the rights situated within this hierarchy, thereby diminishing its value from the point of view of agency. Punishment defends the right and restores its value. It is for this reason that punishment must be proportional to the offence. To punish a serious offence with a trivial punishment does not restore the value of the right. It increases the value of the right from its diminished post-crime situation - it allows PFV to rely on it to a certain extent - but it does not increase it sufficiently. Punishment in this case leads to a right which has less value than it did *before* the offence. Similarly, to punish a minor offence with a very serious punishment would not restore the pre-crime equilibrium. It would *over-value* the right offended against. Furthermore, the act of punishment would itself become a breach of right, this time the right of the criminal. It would make the greater right of the criminal give way to the lesser right of the PFV. For example, if petty theft were met with capital punishment, this would distort the hierarchy of rights by effectively holding that minor property rights are equivalent to the right not to be killed. It is an intrinsic part of Gewirth's theory that a given right can only be made to give way to an equivalent or greater right. In the instance of punishment we make a given right (vested in the criminal) give way to an equivalent right (vested in the PFV). This is not arbitrary because in the specific instance of punishment the criminal and the PFV are not equal in our considerations. Significantly it is the criminal who has brought

about the state of affairs in which only one right can be respected. It is therefore just that we make it his right which gives way.

It is tautological but nevertheless an insight to say that a right is worth defending to the extent that that right has worth. To defend it to a lesser extent would be to take the right less than seriously: to ascribe to the right less worth than it actually has. What PFV demands is not that the fact that he possesses his right be reiterated because this was never in any doubt: no action of the criminal's changes the truth of the statement "PFV has R". What PFV demands is that his right be given worth by being defended to the extent that it has worth.

2.2.3 Is this theory of punishment deterrent in justification?

The ultimate test of whether a theory is deterrent in its justification of punishment is whether it matters (to the justification) that criminals are actually deterred by the punishment or threat thereof. If punishment can be justified irrespective of deterrent considerations then it is not correct to typify it as a deterrent theory.

It is my contention that the justification of punishment offered above is not deterrent. The purpose of punishment is not to deter the crime: it is already too late for this. The purpose of punishment is to restore the value of the right. If each breach is met with indifference then the right is not given its worth and hence has no reliability as a safeguard to action. Punishment ensures that, relative to a background of society which will have more or less respect for rights, the right maintains its value as an agency protector.⁹⁴ It does this by defending the right against those who would render it less reliable as a protector of agency. Deterrence is relevant to such a defence but does not determine its success. Other factors, such as something being seen to be done, the punishment reflecting the importance of the right, the preventative effect of certain punishments, or the sense of closure, will also be of significance. It might be objected that unless punishment somehow deters agents from infringing the PFV's right then it does

⁹⁴ It is the role of other state institutions to increase the background respect for rights and ensure that the state comes ever closer to the ideal model of society governed under the PGC.

not alter that right's value (i.e. its ability to protect the agent's necessary goods). But this objection only holds where the right-value in need of restoration is conceived in terms of an occurrent interest. The force of my argument is to focus on a dispositional interest in the value of right which is ongoing and once breached is breached in an ongoing sense until restored. Consider the following analogy with self-defense. Individual B is the second weakest member of society. Individual A is the weakest. A attacks B with a knife and in the course of defending himself B kills A. B is now the weakest member of society. The fact that B successfully defended himself against A will not now deter other members of the society from attacking B because B has become the weakest and it is to be supposed for the purposes of argument that each other individual is aware that he will easily overcome any defence mustered by B. Does this fact mean that B was unjustified in defending himself against A? Of course not. The self-defence exercised against A was justified solely by considering that B had a right which was in danger of being thwarted if not defended against a wrongful attack. The potential of the defence to alter the way others act towards B is irrelevant to the justification of self-defence. Similarly, punishment is justified simply by the fact that there is an ongoing and wrongful infringement of PFV's right which unless defended against renders that right value-less to the PFV. The purpose of punishment is not to alter the way that others act towards PFV but to alter the way that C *is* acting towards PFV. The unpunished C is responsible for the present situation in which PFV's right is of reduced value to him as an agency protector.

3 Conclusions

I stated at the outset that this thesis was part of a tradition which conceives of the justification of punishment in terms of its ability to restore a pre-crime equilibrium. Many claims have been made in the name of that tradition and I now wish to examine some of these in the context of the argument developed above. This will serve the dual purpose of establishing my position within the tradition and offering a modern interpretation of some dicta which are, on the face of it, highly problematic.

3.1 Hegel claims that the criminal has a right to be punished. I have claimed that the argument presented by Hegel fails to establish this and that in any event to say that punishment is the criminal's right tends to make justification of punishment superfluous. However, within my own justification of punishment there is a residual sense in which it could be said that the criminal has the right to be punished. This is an acknowledgment of the fact that the criminal, like every other agent, is a PFV and, as such, has an agency related interest in seeing that punishment is carried out. Crime throws his dispositional interest in the value of right into doubt as much as any other agent's; the very irrationality of his crime lies in the fact that it reduces the value of right to all agents, including himself. This is little more than a reworking of Kant's distinction between *homo phaenomenon* and *homo noumenon*. The rational man must assent to the truth of the claim that his punishment is justified notwithstanding that the actual man is very much opposed to being punished. The fact that punishment is not in his interests *qua* criminal is irrelevant because his status as an agent is epistemologically prior to his status as a criminal.

3.2 Hegel talks of punishment as justified from the perspective of the realm of rights. This concept is mirrored in my theory by the fact that punishment is justified from the point of view of a community of agents (PFVs) who have equal interests in crime being met with punishment. Similarly the benefits and burdens theorists write the immediate victim of the offence out of the punitive equation. This is correct insofar as it acknowledges that the material harm done to the victim cannot always be undone. However, the correct way to view the victim is, as mentioned above, as existing within a community of PFVs who have equal and mutual interests in crime being punished. Thus, the punishment of the offence committed against V secures the dispositional interests of all PFVs while simultaneously the defence of an interest vested in those PFVs guarantees the ongoing security of V's right.

3.3 Findlay states that Hegel's theory of punishment is an impressive arch from which the keystone is missing. I contend that my theory offers the keystone which Findlay identified as lacking in Hegel's original attempt to construct a justification of punishment. The secret to this is recognising that agents have an ongoing interest in a right once breached; this interest is in the dispositional value of right. This interest is restored where the state acts with a certain degree of seriousness towards the right by defending it. This defence takes the form of punishment of the criminal. Thus punishment restores equilibrium in the sense of returning PFV to a pre-crime position where he was sure that his right had a certain value against the aspirations and actions of other agents.

3.4 In the shipwreck example Kant states that there could be no punishment in the case of two shipwreck victims fighting for one place on a plank where one effectively kills the other by saving himself. I have already offered an explanation of this within Kantian terms. I now wish to suggest that whatever reasons Kant had for making the claim, the conclusions which he drew are sound within my own justification of punishment. Principally, the hierarchy of rights which Gewirth offers is of rights to the generic features of agency, which all agents necessary hold to be goods. As between A's possession of the good of life and B's possession of the same good there is no good ground for preferring one to the other. In a punishment situation we prefer PFV's interest in right to C's because C is responsible for bringing about the situation in which only one right can prevail. *Ex hypothesi* this is not the case in the shipwreck example. The PFV has no dispositional interest in the value of right breached in the shipwreck case because, if through misfortune the situation were to happen to him, he is just as likely to end up on the plank and safe as off it and drowning. PFV only has a concern in the victor being punished if that victor has caused him to doubt the value of a right which he possesses. PFV does not possess a right that other agents give up their lives just to save him.

3.5 Perhaps *the* most problematic dictum within retributive theory is Kant's insistence that even if we were to disband civil society tomorrow, we must

execute the last known murderer in prison and that unless we do this we share in his blood-guilt. Indeed, I suspect that this dictum has caused many theorists to shy away from any justification of punishment which calls itself retributive. I also believe that what Kant says is true.⁹⁵ PFV's right does not stand or fall with civil society; it is a feature of agency rather than a feature of any social structure. Nor does PFV's dispositional interest in the value of right stand or fall with civil society. To revisit the anomic territory employed already as an example in this chapter, the agent has an interest in seeing his right given its due weight notwithstanding that in such a country this is not going to happen. The empirical vicissitudes of the world do not impinge upon the rational force of the agent's rights based claims. Thus, the fact that civil society is on the verge of break up does not lessen PFV's insistence that the substance of his right be recognised. This also goes to show that the theory of punishment that I have offered is not deterrent in justification. C must be punished in this example notwithstanding that no one will be deterred by his punishment. The purpose of his punishment is to establish, and give effect to, PFV's right. So long as there are agents there will be an interest in punishing the infringement of a particular agent's right.

3.6 Finally, having staked a place within the tradition which seeks to justify punishment in terms of balance, I wish to argue for my position within the rights theory of Alan Gewirth. After Gewirth, on the basis of rational agency, we can conclude that agents have categorical rights to their freedom and well-being. After Beyleveld and Brownsword and on the basis of rationally based rights, we can derive a substantive and grounded theory of law. Logically, from the notion of law we can derive a notion of crime. On the basis of a limited notion of free-will

⁹⁵ NB I believe it is true in the formal sense that in such a hypothesis punishment is still an obligation from the point of view of the state. I do not believe that my own justification of punishment necessarily justifies the particular punishment of execution. It has not been my purpose to build on the justification of punishment in principle, to show exactly which punishments are necessary in each case, although it is a necessary feature of my justification of punishment that proportionality is required. However, I believe that there are punishments other than the death penalty which are proportional to the offence of murder (such as life imprisonment). I further believe that for contingent empirical reasons (such as the possibility that in judging someone guilty we may be wrong) we ought not to impose the death penalty. My purpose in considering this point of Kant's is to show that what he says is not some atavistic nonsense. Of course in the actual situation of civil society's imminent dissolution, we might have more pressing concerns than properly performing all outstanding punishments.

we can give meaning to the claim that a particular agent is responsible for the actions which he performs. On the basis of the notion that rights have a dispositional value we can conclude that, where possible, this dispositional value ought to be maintained and where diminished, ought to be restored. The interest in the dispositional value of a right is restored where that right is defended to the extent of its value against the agent who puts its value in doubt. Therefore, there is something intrinsic to the case of an offender who has committed an offence which justifies state intervention.

The End.

References

- AUNE, B. (1970) *Rationalism, Empiricism, and Pragmatism*, Random House, New York.
- _____. (1979) *Kant's Theory of Morals*, Princeton University Press, New Jersey.
- BECK, L.W. (1960) *A Commentary on Kant's Critique of Practical Reason*, The University of Chicago Press, Chicago.
- BEDAU, H.A. (1971) *Radical Egalitarianism*, in H.A. Bedau (ed.), *Justice and Equality*, Prentice-Hall, Englewood Cliffs, New Jersey.
- BENN, S.I. and PETERS, R.S. (1959) *Social Principles and the Democratic State*, George Allen and Unwin, London.
- BENN, S.I. (1971) *Egalitarianism and the Equal Consideration of Interests*, in H.A. Bedau (ed.), *Justice and Equality*, Prentice-Hall, Englewood Cliffs, New Jersey.
- BEYLEVELD, D. and WILES, P. (1979) *How to Retain Your Soul and be a Political Deviant*, in D. Downes and P. Rock (eds.), *Deviant Interpretations*, Martin Robertson, Oxford.
- BEYLEVELD, D. and BROWNSWORD, R. (1986) *Law as a Moral Judgment*, Sweet and Maxwell, London.
- BEYLEVELD, D. (1991) *The Dialectical Necessity of Morality: An Analysis and Defense of Alan Gewirth's Argument to the Principle of Generic Consistency*, The University of Chicago Press, Chicago.
- BOSANQUET, B. (1951) *The Philosophical Theory of the State*, Macmillan, London.
- BRADLEY, F.H. (1962) *Ethical Studies*, Clarendon Press, Oxford.
- BURGH, R.W. (1982) *Do the Guilty Deserve Punishment*, in The Journal of Philosophy, vol. 79, 193-210.
- _____. (1987) *Guilt, Punishment, and Desert*, in F. Schoeman (ed.), *Responsibility, Character, and the Emotions*, Cambridge University Press, Cambridge.
- BYRD, B.S. (1989) *Kant's Theory of Punishment: Deterrence in its Threat, Retribution in its Execution*, in Law and Philosophy, vol. 8, 151-200.
- CAMUS, A. (1963) *The Fall*, Penguin Books, London.

- COOPER, D.E. (1971) *Hegel's Theory of Punishment*, in Z.A. Pelczynski (ed.), *Hegel's Political Philosophy: Problems and Perspectives*, Cambridge University Press, Cambridge.
- DAY, J.P. (1978) *Retributive Punishment*, in *Mind*, vol. 87, 498-516.
- DOYLE, J.F. (1969) *Justice and Legal Punishment*, in H.B. Acton (ed.), *The Philosophy of Punishment*, Macmillan and Co. Ltd., London.
- DUFF, R.A. (1986) *Trials and Punishments*, Cambridge University Press, Cambridge.
- EWING, A.C. (1972) *On Retributivism*, in G. Ezorsky (ed.), *Philosophical Perspectives on Punishment*, State University of New York Press, Albany.
- FALLS, M.M. (1987) *Retribution, Reciprocity, and Respect for Persons*, in *Law and Philosophy*, vol. 6, 25-51.
- FINDLAY, J.N. (1958) *Hegel: A Re-examination*, Gregg Revivals, Aldershot.
- _____. (1981) *Kant and the Transcendental Object: A Hermeneutic Study*, Clarendon Press, Oxford.
- FINNIS, J. (1972) *The Restoration of Retribution*, in *Analysis*, vol. 32, 131-135.
- _____. (1980) *Natural Law and Natural Rights*, Clarendon Press, Oxford.
- FLETCHER, G.P. (1987) *Law and Morality: A Kantian Perspective* in *Columbia Law Review*, vol. 87 (1), 533-558.
- FLEW, A. (1969) *The Justification of Punishment*, in H.B. Acton (ed.), *The Philosophy of Punishment*, Macmillan and Co. Ltd., London.
- GALLIGAN, D.J. (1981) *The Return to Retribution in Penal Theory*, in C.F.H. Tapper (ed.), *Crime, Proof and Punishment: Essays in Memory of Sir Rupert Cross*, Butterworths, London.
- GARCIA, J.L.A. (1989) *Deserved Punishment*, in *Law and Philosophy*, vol. 8, 263-277
- GEWIRTH, A. (1978) *Reason and Morality*, The University of Chicago Press, Chicago.
- _____. (1991) *Foreword*, in D. Beyleveld, *The Dialectical Necessity of Morality: An Analysis and Defense of Alan Gewirth's Argument to the Principle of Generic Consistency*, The University of Chicago Press, Chicago.
- _____. (1996) *The Community of Rights*, The University of Chicago Press, Chicago.

- GREGOR, M.J. (1963) *Laws of Freedom: A Study of Kant's Method of Applying the Categorical Imperative in the "Metaphysik der Sitten"*, Basil Blackwell, Oxford.
- HART, H.L.A. (1961) *The Concept of Law*, Clarendon Press, Oxford.
- _____. (1968) *Punishment and Responsibility: Essays in the Philosophy of Law*, Clarendon Press, Oxford.
- HAWKINS, D.J.B. (1971) *Punishment and Moral Responsibility*, in S.E. Grupp (ed.), *Theories of Punishment*, Indiana University Press, Bloomington.
- HEGEL, G.W.F. (1967) *Hegel's Philosophy of Right*, translated with notes by T.M. Knox, Oxford University Press, Oxford.
- _____. (1991) *Elements of the Philosophy of Right*, translated by H.B. Nisbet, (ed.) A.W. Wood, Cambridge University Press, Cambridge.
- HESTEVOLD, H.S. (1987) *On the Moral Status of Punishment*, in Law and Philosophy, vol. 6, 249-257.
- HONDERICH, T. (1969) *Punishment: The Supposed Justifications*, Hutchinson, London.
- _____. (1976) *Punishment: The Supposed Justifications*, (Revised edition), Penguin Books, Harmondsworth.
- HUDSON, W. (1967) *Ethical Intuitionism*, Macmillan, London.
- INWOOD, M. (1992) *A Hegel Dictionary*, Blackwell, Oxford.
- KANT, I. (1929) *Critique of Pure Reason*, MacMillan and Co., London.
- _____. (1948) *The Moral Law: Kant's Groundwork of the Metaphysic of Morals*, Hutchinson, London.
- _____. (1956) *Critique of Practical Reason*, The Bobbs-Merrill Co. Inc., Indianapolis.
- _____. (1965) *The Metaphysical Elements of Justice: Part I of the Metaphysics of Morals*, The Bobbs-Merrill Co. Inc., Indianapolis.
- _____. (1960) *Religion Within the Limits of Reason Alone*, Harper and Row, New York.
- _____. (1983) *Perpetual Peace and Other Essays on Politics, History and Morals*, Hackett Publishing Company, Indianapolis.

- _____. (1987) *Foundations of the Metaphysics of Morals*, Prometheus Books, New York.
- _____. (1991) *The Metaphysics of Morals*, Cambridge University Press, Cambridge.
- KLEINIG, J. (1973) *Punishment and Desert*, Martinus Nijhoff, The Hague.
- LACEY, N. (1988) *State Punishment: Political Principles and Community Values*, Routledge, London.
- LADD, J. (1965) *Introduction to I. Kant, The Metaphysical Elements of Justice: Part I of the Metaphysics of Morals*, The Bobbs-Merrill Co. Inc., Indianapolis.
- LUCAS, J.R. (1980) *On Justice*, Clarendon Press, Oxford.
- MABBOTT, J.P. (1971) *Punishment*, in S.E. Grupp (ed.), *Theories of Punishment*, Indiana University Press, Bloomington.
- MARX, K. (1977) *Critique of The Gotha Programme*, in D. McLellan (ed.), *Karl Marx: Selected Writings*, Oxford University Press, Oxford.
- MOORE, M.S. (1987) *The Moral Worth of Retribution*, in F. Schoeman (ed.), *Responsibility, Character, and the Emotions*, Cambridge University Press, Cambridge.
- MORRIS, H. (1971) *Persons and Punishment*, in S.E. Grupp (ed.), *Theories of Punishment*, Indiana University Press, Bloomington.
- MURPHY, J.G. (1969) *Kant's Concept of a Right Action* in L.W. Beck (ed.), *Kant Studies Today*, Open Court, La Salle.
- _____. (1979) *Retribution, Justice, and Therapy: Essays in the Philosophy of Law*, D. Reidel Publishing Co., Dordrecht.
- _____. (1992) *Retribution Reconsidered: More Essays in the Philosophy of Law*, Kluwer Academic Publishers, Dordrecht.
- NOZICK, R. (1974) *Anarchy, State, and Utopia*, Basil Blackwell, Oxford.
- _____. (1981) *Philosophical Explanations*, Clarendon Press, Oxford
- PATON, H.J. (1947) *The Categorical Imperative: A Study in Kant's Moral Philosophy*, University of Pennsylvania Press, Philadelphia.
- _____. (translation and analysis) (1948) *The Moral Law: Kant's Groundwork of the Metaphysic of Morals*, Hutchinson, London.

- PHILIPS, M. (1986) *The Justification of Punishment and the Justification of Political Authority*, in Law and Philosophy, vol. 5, 393-416.
- PLATO, (1974) *The Republic*, Penguin Books, London.
- POSNER, R.A. (1980) *Retribution and Related Concepts of Punishment*, in Journal of Legal Studies, vol. 9, no. 1, 71-92.
- QUINTON, A.M. (1969) *On Punishment*, in H.B. Acton (ed.), *The Philosophy of Punishment*, Macmillan and Co. Ltd., London.
- RAWLS, J. (1969) *Two Concepts of Rules*, in H.B. Acton (ed.), *The Philosophy of Punishment*, Macmillan and Co. Ltd., London.
- _____. (1972) *A Theory of Justice*, Oxford University Press, Oxford.
- REYBURN, H.A. (1921) *The Ethical Theory of Hegel: A Study of the Philosophy of Right*, Clarendon Press, Oxford.
- RIEDEL, M. (1971) *Nature and Freedom in Hegel's "Philosophy of Right"*, in Z.A. Pelczynski (ed.), *Hegel's Political Philosophy: Problems and Perspectives*, Cambridge University Press, Cambridge.
- ROSS, W.D. (1930) *The Right and The Good*, Clarendon Press, Oxford.
- SADURSKI, W. (1985) *Giving Desert its Due: Social Justice and Legal Theory*, D. Reidel Publishing Co., Dordrecht.
- SCHEID, D.E. (1983) *Kant's Retributivism*, in Ethics, vol. 93, 262-282.
- SHER, G. (1987) *Desert*, Princeton University Press, New Jersey.
- SIBLER, J.R. (1960) *Introduction to I. Kant, Religion Within the Limits of Reason Alone*, Harper and Row, New York.
- STERBA, J.P. (1980) *The Demands of Justice*, University of Notre Dame Press, Notre Dame.
- SULLIVAN, R.J. (1989) *Immanuel Kant's Moral Theory*, Cambridge University Press, Cambridge.
- TEN, C.L. (1987) *Crime, Guilt and Punishment: A Philosophical Introduction*, Clarendon Press, Oxford.
- TODDINGTON, S.W. (1989) *Towards an Integrated Theory of Natural Law*, unpublished Ph.D. thesis, University of Sheffield, Sheffield.
- VON HIRSCH, A. (1976) *Doing Justice: The Choice of Punishments*, Hill and Wang, New York.

- _____. (1986) *Past or Future Crimes: Deservedness and Dangerousness in the Sentencing of Criminals*, Manchester University Press, Manchester.
- _____. (1993) *Censure and Sanctions*, Clarendon Press, Oxford.
- WALKER, N. (1991) *Why Punish?*, Oxford University Press, Oxford.
- WASSERSTROM, R.A. (1980) *Philosophy and Social Issues*, University of Notre Dame Press, Notre Dame.
- WEINRIB, E.J. (1987) *Law as a Kantian Idea of Reason*, in Columbia Law Review, vol. 87(1), 472-508.
- WILLIAMS, B.A.O. (1971) *The Idea of Equality*, in H.A. Bedau (ed.), *Justice and Equality*, Prentice-Hall, Englewood Cliffs, New Jersey.