A Nietzschean critique of Kant's ethics.

I Introduction

In the *Antichrist* (A), Nietzsche openly rejects central concepts in Kant's ethics: "Nothing works more profound ruin than any 'impersonal' duty (...) Kant's categorical imperative should have been felt as *mortally dangerous*" (A 11). Although Nietzsche clearly did not accept the notion of duty and the Categorical Imperative (CI), he often presents arguments against "slave morality" in general and does not criticise specific aspects of different value systems. This might get confusing because slave morality is a broad term that encompasses not just Kant's ethics but also utilitarianism and Christian ethics. Thus, it is important to investigate which of Nietzsche's arguments against slave morality apply particularly to Kant. For this reason, in this essay I aim to extract Nietzsche's critique against Kant specifically by relying mostly on the first essay of *On the Genealogy of Morals* (GM I). I conclude that Kant cannot defend his ethics against the objections in GM I.

I start with a brief exposition of Kant's ethics as presented in the *Groundwork of the Metaphysics of Morals* (Section II). Next, I present Nietzsche's critique of slave morality in GM I (Section III). I then discuss two objections raised by Nietzsche against Kant: 1) the CI assumes the validity of certain principles of slave morality without justifying them (Section IV.1) and 2) Kant does not prove the existence of a transcendental subject (Section IV.2). I also argue why Kantian replies to these objections fail.

II Kant's ethics

In the *Groundwork* (G), Kant starts from a key distinction that is explicated in the third section (G 4:452), albeit hinted in the preface (G 4:389): we belong to both the natural world and the intelligible world via our empirical self and our transcendental self, respectively. Kant assumes the existence of moral laws and natural laws which govern each world. Given this framework, Kant's main objective in the *Groundwork* (G 4:392) is to derive "the supreme principle of morality" (the CI) which grounds moral laws (G 4:416,429).

To find this principle, we need to approach morality *a priori* as transcendental subjects: "All moral concepts have their seat and origin completely a priori in reason" (G 4:411). One way to motivate this approach is that when we pronounce moral judgements, we take them to be valid irrespectively of our situation in the natural world (Johnson and Cureton, 2021, Sec. 1). However, if we appeal to experiences in our derivation, we will be unable to derive an unconditionally binding and objective moral system: its moral principles will hold conditional on having these experiences (G 4:411). Only *a priori* reasoning could give rise to unconditional and objective duties. Using such reasoning, Kant shows how we can derive the different formulations of the CI. Due to space constraints, I focus on two formulations: the Universal Law Formulation (FUL) and the Humanity Formulation (FH).

Firstly, FUL states: "*Act only according to that maxim through which you can at the same time will that it become a universal law*" (G 4:421). FUL allows us to evaluate the morality of a maxim by imagining a world, in which everybody acts according to that maxim. We can turn FUL into a testing procedure (Johnson and Cureton, 2021):

(i) Formulate one's desired course of action as a maxim.

(ii) Imagine a world, in which everybody acts according to that maxim.

(iii) Check whether or not following the original maxim in that world is logically contradictory.

(iv) If it is not logically contradictory, examine if one can rationally will to follow one's maxim in that world.

A maxim can fail stage (iii) or (iv). In this essay, I take stage (iv) as requiring rational coherence in the sense that following our maxim in the imagined world is coherent with practical rationality. I understand practical rationality as our ability to act in accordance with certain (moral) principles (Callanan, 2013, p.68).[1] If all stages are passed, the maxim is morally permissible.

Kant offers various applications to illustrate how FUL works (G 4:421-3). In the promise example, an individual who needs money contemplates whether to make a false promise to repay a lender later. Using FUL, this action fails (iii): in a world where everybody is making false promises the concept of promise is logically contradictory (since nobody would trust a promise). Thus, false promises are not morally permissible. As shown in Section IV, such applications of FUL are crucial for understanding Nietzsche's critique.

Secondly, FH postulates: *"[A]ct that you use humanity, in your own person as well as in the person of any other, always at the same time as an end, never merely as a means"* (G 4:429). Humanity is understood as a set of characteristics that make us human, e.g., our capacity to engage in rational thought. On FH each person has some intrinsic human dignity which we can violate if we treat them as means towards an end. We can directly see why FH rules out the promise example: by making a false promise the borrower disrespects the lender's human dignity by treating her as a means towards obtaining money.

Some passages of the *Groundwork* suggest that the different formulations might be equivalent statements of CI (G 4:436). Certain commentators have interpreted this to mean that FH and FUL give rise to the same type of duties in all situations (Johnson and Cureton, 2021). However, in my view Wood (2007, p.79-82) has argued convincingly that Kant does not make this claim in the *Groundwork*: FH and FUL can give raise to

[1]Defining practical rationality more precisely is beyond the scope of my essay. Broadly, I assume O'Neil's understanding of practical rationality (2015, pp.39-41).

different duties. In this essay, I assume his interpretation, although I cannot defend it for reasons of space.

This may seem contradictory at first since it implies that Kant is making the following three claims: (a) there is a supreme principle of morality, (b) FUL and FH are expressions of that unique principle and (c) FUL and FH do not entail equivalent duties. It is not obvious how FUL and FH are expressing the unique principle of morality but also giving raise to (potentially) different duties. Wood (2007) reconciles these claims by arguing that a third formulation of the CI (the Autonomy Formulation) combines both FUL and FH and so implies the combined set of duties prescribed by the other two formulations. There is some disagreement as to how the Autonomy Formulation should be stated, e.g., compare Callanan (2013, p.99) and Wood (2007, p.86), but in my view Johnson and Cureton (2021, Sec. 7) present it most clearly: "Act so that through your maxims you could be a legislator of universal laws." This formulation combines the idea of universal law with the idea of a person capable of rational thought (and so possessing human dignity) as a legislator (Wood, 2007, p.76). So, the Autonomy Formulation can be seen as Kant's intended expression of the CI. For our purposes, I will consider CI to be made of the separate (sub)principles FUL and FH since it illustrates more clearly how Nietzsche's critique applies.

III Nietzsche's critique of slave morality

In GM I Nietzsche provides a historical account of how European morality has arisen while simultaneously criticising its origin and its content. On the surface, GM I is a story about how in Ancient Rome a morality advocated by the priestly caste replaced the value system of aristocratic warriors (GM I,16). Importantly, Nietzsche envisages Ancient Rome as divided into common men and a noble class which consists of warriors (or knights) and priests. At first, the warrior-nobles win political power and impose their values which characterizes "good" as their own traits such as "powerful physicality" and "bad" as commoners' features (GM I,7). Unhappy with their lack of power, the priests

invent a new morality which would appeal to the common men. This "slave" morality considers "evil" the traits of the warriors such as violence, but it praises as "good" various characteristics of the commoners, e.g., being harmless, humble and patient (GM I,13-14). In this sense, slave morality inverts the "aristocratic value equation" (GM I,7). This new morality eventually replaces noble values as the established moral code in Rome and subsequently in Europe (GM I,16).

There are two related points that need clarifying. Firstly, it is important to realise what motivates the priests to replace aristocratic values, namely *ressentiment*. After their initial loss, the priests realise that their goal of political supremacy is unachievable. While they have still not given up this (unachievable) goal, they publicly commit to slave morality which promotes political equality and which devalues their original goal (GM I,7). This is an example of repressed vengefulness: the priests' endorsement of slave morality is motivated by their (repressed) desire to take revenge on the warriors by gaining political power (GM I,14).

According to Reginster (1997, p.281), Nietzsche objects to such *ressentiment* valuation because it undermines priests' integrity. They simultaneously hold two sets of values which are incompatible with each other. Moreover, such a strategy of *ressentiment* valuation is self-defeating (Reginster, 1997, p.303). If the priests achieve political equality, they would not be satisfied because it conflicts with their original goal of political supremacy. If they achieve power, they will be unable to enjoy it wholeheartedly, since it clashes with their professed values. Thus, the adoption of slave morality on the basis of *ressentiment* is problematic. Admittedly, other commentators have stressed different reasons for Nietzsche's objections to slave morality, e.g., it thwarts the flourishing of higher man (Leiter, 2021, Sec. 1.3). In my view, Reginster's interpretation has stronger textual support at least in GM I which explains why I focus primarily on his reading. While Nietzsche hints to the effects of slave morality on higher man once in the essay (GM I,12), the problematic nature of *ressentiment* is discussed in much greater depth (GM I,10-1,13-4,16).

Secondly, there are various differences between noble values and slave morality, e.g., the latter prescribes the same set of duties for everybody whereas the former postulates

different duties for nobles and common men. I would like to emphasise one difference in particular. With warriors values arise spontaneously and their value system "seeks its opposite only so as to affirm itself more gratefully and triumphantly" (GM I,10). Noble values are the honest expression of what warriors believe to be good. In contrast, given that slave morality arises out of *ressentiment*, the primary driving force is the condemnation of noble values: "slave morality from the outset says No to what is 'outside,' what is 'different' (...) and *this* No is its creative deed" (GM I,10). The creation of new values is secondary to the task of negating noble values. So, it is not surprising that many commands issued by slave morality condemn various actions: "he is good who does not outrage, who harms nobody, who does not attack, who does not requite, who leaves revenge to God, who keeps himself hidden as we do, who avoids evil and desires little from life, like us, the patient, humble and just" (GM I,13). Many of these actions would clearly be allowed on noble valuation which recognises that "happiness should not be sundered from action" and praises "enthusiastic impulsiveness in anger, love, reverence, gratitude and revenge" (GM I,10).

We can now see why Nietzsche considers Kant's ethics as slave morality: CI condemns various actions allowed on noble valuation, e.g., extramarital sex or suicide. Regarding the former, Kant can argue that it involves treating another person as a means towards the end of sexual gratification. Additionally, Kant's ethics implies the same set of duties for all individuals, irrespectively of whether they are nobles or common men (most obviously in FUL). As noted, Nietzsche disapproves of such moral systems giving raise to universal duties (A 11).

IV Nietzsche's critique applied to Kant

In my view, the most important objections raised by GM I are: 1) Kant's system already presupposes the values of slave morality that it seeks to establish and 2) it rejects the existence of a transcendental subject.

IV.1 CI assumes slave morality

To understand the first objection, let us consider how Nietzsche's critique applies to FH first and then to FUL. In *Beyond Good and Evil* (BGE), Nietzsche rejects the idea of intrinsic human dignity and thus Kant's requirement for treating others as an end rather than 'merely as a means'. He can support this claim by arguing that recognising the differences in agents' worth is fundamental for the flourishing of a society: "Every elevation of the type 'man,' has hitherto been the work of an aristocratic society and so it will always be – a society believing in a long scale of gradations of rank and *differences of worth among human beings*, and requiring slavery in some form or other" (BGE 257, my emphasis). While this quote can be read in light of Nietzsche's politics, it also shows his belief in the different worth of individuals. Moreover, Nietzsche can argue that Kant's belief in intrinsic moral worth originates in Christianity (A 10-12) and so is a prime example of slave morality.[2] Accepting FH already commits us to certain principles of slave morality.

Nevertheless, even if FH assumes slave morality, FUL might be sufficient to establish (most of) our moral duties without relying on slave morality. Nietzsche can directly argue against FUL's assumption that universalizability is a useful principle when deciding on the moral permissibility of a given action. Due to the intrinsic differences in human's worth, different values should apply to higher and common men, thus rejecting FUL (BGE 257).

Furthermore, Nietzsche can attack stage (iv) in FUL where a person examines if she can rationally will to follow her maxim in a world where it is universalised. As an example, let us consider extramarital sex. While Kant did not explicitly discuss sex in the *Groundwork*, he argued in the *Metaphysics of Morals* (MM) that outside of heterosexual marriage it is morally impermissible and within marriage it is permissible only if aimed at procreation (MM 6:277), i.e., Kant has a teleological view of sex. Let us examine if FUL can yield this result by formulating the maxim:[3] "One need not marry to have sex." If we plug in our maxim via FUL, it is clearly not logically contradictory: we can

[2]Below I discuss if this reasoning suffers from the Genetic Fallacy.

[3]We focus on a justification via FUL because we have shown why Nietzsche rejects FH above.

imagine a world where people engage in extramarital coitus but still marry. Kant can argue that engaging in extramarital sex is not rationally coherent because it will involve having sex for non-teleological reasons. For a Kantian, we are losing self-respect by doing this because we are driven by inclinations (MM 6:425). So, we cannot rationally will to live in a world where extramarital sex is permitted. However, this reasoning rests on various assumptions that can plausibly be connected back to Nietzsche's understanding of slave morality (GM I,10), e.g., one should not be driven by instincts or one should only have sex for teleological reasons. Thus, Nietzsche can argue that Kant's CI is not a helpful guide unless one also assumes the truthfulness of certain additional principles of slave morality in the content of practical rationality in (iv). Such a reasoning, however, seems cyclical: while Kant claims that on its own CI can establish all of our duties *a priori* (4:429), he needs to appeal to moral judgements outside the scope of CI to justify its application (specifically stage (iv)). In this sense, Kant's ethics already presupposes what it seeks to establish.

Nevertheless, a Kantian may object to Nietzsche's reasoning for rejecting slave morality. On Reginster's reading (1997), Nietzsche rejects it because it is grounded in the psychological state of *ressentiment* as argued in Section III. However, a Kantian can argue that this implies the Genetic Fallacy: the origin of a moral judgement has no bearing on its truthfulness. To illustrate this fallacy, suppose that in late 1910s I start believing that the universe is static because Einstein says so and I generally find the theory of relativity convincing.[4] While in general Einstein's works are a good source for learning about physics, his belief in a static universe is incorrect as shown by Hubble in 1929. So, whether my belief (in static universe) is correct is independent from my judgement about how good its origin is (Einstein's theory). When applied to Nietzsche, this objection suggests that even though one's belief in human dignity might have a problematic origin, this does not mean that FH is wrong. There is some textual evidence in the *Groundwork* that Kant might be sympathetic to this reasoning. Since he believes that ethics should be approached *a priori* (G 4:411), he would object to seeking the origin of a moral belief

[4]Einstein eventually revised his views on a static universe but this was only after Hubble's findings in the 1920s.

in experience and using this origin to judge its truthfulness.

It is not, however, always true that the origin of a belief is independent from its justificatory status. Consider the Explanatory Argument (Sinhababu, 2007): we are justified in believing a hypothesis if it features in our best explanation of a particular observation.[5] Let us apply this argument to the hypothesis that each human has intrinsic dignity. Kant can observe that he believes in FH and explain it in light of the hypothesis that each person actually has intrinsic dignity. However, Nietzsche can respond that Kant holds his belief in FH not because that hypothesis is correct but because of his Pietist upbringing which left a lasting mark on his moral outlook. Crucially, Pietist doctrines indeed postulate intrinsic human dignity.[6] Nietzsche can then argue that his explanation for Kant's belief is better than Kant's own explanation because it does not have to posit some additional feature (human dignity) but appeals to an empirical fact (Kant's upbringing). This is the principle behind Occam's razor: when faced with competing explanations for a observation, we should choose the explanation which makes fewer assumptions. In this case, Nietzsche's explanation makes fewer assumption. In *Daybreak* (10), Nietzsche expresses sympathy for the Explanatory Argument (Sinhababu, 2007, p.273). Thus, Kant's belief in FH is not a sensible assumption since he has no right to assume people actually possess intrinsic human dignity.

IV.2 Natural Substratum

A Kantian may not be entirely satisfied with the last argument, e.g., she might reject the use of Occam's razor in that case. Due to space constraints, I cannot address this response. Even so, GM I offers another argument against Kant. It rejects the possibility of assuming a transcendental subject and so the distinction that Kant draws between an empirical and a transcendental self: "[Slave] morality also separates strength from expressions of strength, as if there were a neutral substratum behind the strong man

[5]I focus on Sinhababu's statement of the argument because he discusses it explicitly in a Nietzschean context.

[6]See Crosby (1994, p.127-129) for the connections between Pietism and Kant's FH.

which was free to express strength or not to do so" (GM I,13). Slave morality assumes that behind the appearance of a person there is a subject possessing free will who chooses whether to express their strength. However, Nietzsche disagrees that such a subject exists: "But there is no such substratum; there is no 'being' behind doing, effecting, becoming; 'the doer' is merely a fiction added to the deed – the deed is everything" (GM I,13). In explaining deeds, there is no need to appeal to a transcendental subject. Since "a neutral substratum" is necessary for voluntary actions and moral responsibility, Nietzsche also rejects free will (Leiter, 2021, Sec. 1.2). Grammar is one reason for our mistaken belief in a transcendental subject. Our sentences often require a grammatical subject that does a particular action which some have taken as evidence for a metaphysical subject (BGE 17). Another reason is that assuming an agent responsible for her action serves the goal of slave morality (GM I,13). The priest can then make people feel "accountable" and so guilty of performing actions condemned by slave morality.

Kant not only assumes such a transcendental subject, but its existence is necessary for making moral judgements. Unless we have certain agency in deciding how to act, it does not make sense to discuss whether we choose to act from duty. If there is no *I* that decides whether to make a false promise, there is nobody who has a duty not to make false promises. Furthermore, it is doubtful whether we can derive the CI *a priori* without a transcendental subject. There is no *I* to do the reasoning required for Kant's transcendental argument for the deduction of the possibility of CI (Callanan, 2013, p.115-119). In that argument, Kant tries to deduce the possibility of CI from the process we use to make moral judgements and the concepts we utilise such as duty (G 4:454). However, a subject is necessary for the deduction to work: the *I* is needed to observe how moral reasoning works and how we use moral concepts. Notably, since Kant thinks the derivation of CI is *a priori*, that *I* could only be a transcendental subject. Denying the transcendental subject seems quite detrimental since we cannot derive the CI without assuming its existence.

My reading of the *Groundwork* did not point me to a distinctively Kantian objection against this argument. However, a Kantian might still be worried that Nietzsche's denial

of a transcendental subject is counterintuitive: when engaging in moral deliberation, most of us reason *as if* there is a subject with free will. If there is no doer "behind" a deed, it is unclear how we can talk about actions (Pippin, 2006, pp.132-133). For example, to explain why a person wears a mask in the tube during a pandemic, we need to introduce a subject who, for example, intends to protect herself and other people. If Nietzsche denies the existence of a doer behind the deed, there is no natural way to explain why this person wears a mask. On the other hand, Nietzsche often talks about different types of people who are interacting with each other and of their deeds rather than subject-less events, e.g., the strife between priestly and warrior nobles in Ancient Rome (GM I,13). So, the question is if Nietzsche can talk about deeds without postulating a transcendental subject behind that deed.

Pippin (2006) provides one answer to this quibble. He interprets Nietzsche's position as suggesting that one's self is only revealed in actions and cannot be pinned down outside of deeds. Often, we cannot distinguish between our intentions and our actions because our reasons for a deed may evolve over time (Pippin, 2006, pp.139-140): at first a person might be wearing a mask on the tube to protect herself but later she might be wearing it to protect vulnerable relatives. In that sense, it is not really a fixed intention that makes her wear a mask. While she is performing the deed, her intentions are being expressed in the deed and she becomes aware of what is truly motivating her and so of her *I*. To give a full account of her decision to wear a mask, we also need to appeal to psycho-physical facts such as convention that influence her behaviour (Pippin, 2006, p.140). Thus, while performing the action of wearing a mask every day, a person becomes aware of the different factors and intentions that are driving her behaviour. In that process she becomes aware of possessing an *I*. On this view, deeds are expressions of her *I* (Pippin, 2006, p.144) rather than caused by a transcendental subject outside of these deeds.

V Conclusion

Given the aforementioned, I have argued that Kant cannot defend what Nietzsche calls slave morality in GM I. Kant's ethics is a prime example of a slave morality because it prescribes a universal code of conduct and largely aims at negating rather than creating values.

There are two main objections that Nietzsche raises against Kant. Firstly, Nietzsche can attack the CI for being insufficient on its own to derive all principles in Kant's moral system. For CI to work, Kant needs to assume the validity of certain principles of slave morality such as intrinsic human dignity and universalizability without justifying them *a priori*. While a Kantian may object that Nietzsche's reason for rejecting these principles is subject to the Genetic Fallacy, the Nietzschean critique still stands because the best explanation for Kant's beliefs in FH does not imply their truthfulness.[7]

Secondly, Nietzsche also rejects the idea of a transcendental subject. Without a transcendental subject many of Kant's *a priori* arguments cannot work. Crucially, his deduction of the CI fails. One can argue that Nietzsche's denial of a transcendental subject might have some counterintuitive implications for describing actions. However, I have shown that Nietzsche can still talk about agents performing deeds since deeds are an expression of one's *I*.

Thus, Kant does not succeed in establishing a moral system of universally binding duties *a priori*. To achieve this, Kantians need to provide a stronger response to Nietzsche's rejection of a transcendental subject or clear Kant's systems of unjustifiable assumptions.

[7]While I have not discussed the origin of Kant's belief in FUL, a similar case can probably be made for its origin in Kant's Pietist upbringing.

Primary Sources

Kant, I (1997), *Groundwork of the Metaphysics of Morals*, trans. M. Gregor and J.

Timmermann, Cambridge University Press.

Kant, I (2017), *The Metaphysics of Morals*, trans. M. Gregor, Cambridge University

Press.

Nietzsche, F (1954), *The Antichrist*, *in* W. Kaufmann, ed. & trans., Portable Niet-

zsche, Viking.

Nietzsche, F (1966), *Beyond Good and Evil*, trans. W. Kaufmann, Vintage.

Nietzsche, F (1967), *On The Genealogy Of Morals*, *in* W. Kaufmann, ed. & trans.,

On The Genealogy Of Morals and Ecce Homo, Vintage.

References

Callanan, J. (2013), *Kant's Groundwork for the Metaphysics of Morals*,

Edinburgh University Press.

Crosby, D. A. (1994), 'Kant's ideas about ultimate reality and meaning in relation to his moral theory: Critique of an enlightenment ideal', *Ultimate reality and meaning* **17**(2), 117–136.

Johnson, R. and Cureton, A. (2021), Kant's Moral Philosophy, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy'.
URL: *https://plato.stanford.edu/archives/spr2021/entries/kant-moral*

Leiter, B. (2021), Nietzsche's Moral and Political Philosophy, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy'.
URL: *https://plato.stanford.edu/archives/sum2021/entries/nietzsche-moral -political/*

O'Neill, O. (2015), Kant: rationality as practical reason, *in* 'Constructing Authorities: Reason, Politics and Interpretation in Kant's Philosophy', Cambridge University Press, p. 38–55.

Pippin, R. B. (2006), Lightning and flash, agent and deed, *in* C. D. Acampora, ed., 'Nietzsche's On the Genealogy of Morals: Critical Essays', Rowman & Littlefield Publishers, pp. 131–144.

Reginster, B. (1997), 'Nietzsche on ressentiment and valuation', *Philosophy and Phenomenological Research* **57**(2), 281–305.

Sinhababu, N. (2007), Vengeful thinking and moral epistemology, *in* B. Leiter and N. Sinhababu, eds, 'Nietzsche and Morality', Oxford University Press, London, pp. 262–280.

Wood, A. (2007), The Moral Law, *in* 'Kantian Ethics', Cambridge University Press, Cambridge, pp. 66–84.