Can Civil Disobedience be Justified through a Kantian Theory?

Sosveen, Aurora

Awarding institution:
King's College London

Can Civil Disobedience be Justified through a Kantian Theory?

Aurora Johanne Holland Jevanord Søsveen

Thesis submitted for the consideration of obtaining the degree

Master of Philosophical Studies in Philosophy

King's College London

University of London

2023

Abstract

This thesis examines the possibility of whether the concept of civil disobedience can be justified through a Kantian theory. This represents a tension in Kantian political philosophy that has hitherto been largely overlooked by commentators. Civil disobedience raises the question of which set of laws we should obey in cases where the civil law demands something contrary to the moral law. Such a situation could potentially result in a conflict of duties, which for Kant is impossible. Hence, the concept of civil disobedience poses a challenge for Kant's philosophy as the concept could potentially render Kant's basic claim of obeying the law incomprehensible if it recommends obeying two conflicting laws at once. Civil disobedience provides an interesting alternative to the concept of revolution, as it does not aim to overthrow the constitution, rather, it fully accepts the authority of the state, yet it still deliberately breaks laws to achieve the goal of changing the legal framework. The background for addressing this question is the fact that existing literature mostly addresses Kant's arguments and attitudes concerning revolution. Consequently, there is a gap in the literature concerning Kant's political philosophy. This thesis aims to fill this gap by assessing different forms of potential justifications for Kantian civil disobedience, and argues that Kant's political philosophy, when properly understood, already encapsulates the essential concept of civil disobedience as part of Kant's existing theory.

Table of contents

List of abbreviations

The abbreviations used for Kant's works in this thesis is as follows and will follow standard guidance for pagination when cited (e.g., "MM 6:302"). When referred to in the text, the abbreviation will be as stated in the brackets. For full citation, please see bibliography.

E – An Answer to the Question: What is Enlightenment? (*What is Enlightenment?)*

GM – Groundwork of the Metaphysics of Morals (*Groundwork*)

Idea - Idea for a Universal History with a Cosmopolitan Intent (*Universal History*)

MM – The Metaphysics of Morals

TP – On the Proverb: That May be True in Practice, but is of No Practical Use (*Theory and Practice*)

TPP – To Perpetual Peace: A Philosophical Sketch (*Perpetual Peace*)

Introduction

The aim of this thesis is to address the question of whether there could be a Kantian theory justifying civil disobedience. While there is extensive literature on the topic of Kant and revolution, the topic of civil disobedience has so far been overlooked. Kant did not address the question himself, and so this leaves a textual gap that a Kantian theory should seek to fill. The question is thus twofold: one concerns whether there is scope for a Kantian justification of civil disobedience, the second question concerns what a Kantian theory of civil disobedience could look like. My emphasis will be the first part of the question, in which I shall argue that there is not scope for a Kantian justification of civil disobedience. In answer to the second part of the question, I propose my alternative conceptualisation which I argue achieves the same goals as civil disobedience but remains firmly within the Kantian framework.

The argument will proceed as follows. First, I give a definition of civil disobedience directly drawn from Rawls. This definition will remain unchallenged throughout the thesis as it is the starting point for our enquiry, and the aim of the thesis is not to assess the adequacy of the Rawlsian definition. Second, I present what I consider intuitive cases of civil disobedience. These are both historical and contemporary cases in which the agents broke the law, yet our intuitive judgement is to side with them. These cases represent the unclear grounds we could

have for believing what they did was right when it was contrary to law, and function as test cases against the proposed solutions. The next chapter introduces aspects of Kant's political philosophy. The aim in this chapter is purely exegetical, and covers only what is most relevant from Kant's political philosophy in relation to the question on civil disobedience. Following this, I begin examining Kant's opposition to revolution by examining his argument. I then argue that the concept and discussion of revolution shares enough common ground with civil disobedience to form the basis on analogical inference; if we can find a Kantian justification for revolution, we can find a Kantian justification for civil disobedience.

I spend the majority of the thesis assessing the possibility of a moral justification for revolution, and subsequently civil disobedience. This is because I find this approach most intuitive, as we would typically agree that there was some moral justification in the proposed test cases for civil disobedience. In this chapter, I first examine Korsgaard's approach, before considering two different approaches presented by Hill. I reject both of these on the grounds that they neglect crucial aspects of Kant's theory. In a brief interlude, I write on the potential of some of these aspects, but conclude that this alone is not enough to address the question of neither revolution nor civil disobedience. In response, I discuss what I consider the most viable approach to an essentially moral justification for revolution, which is the Enthusiasm model. With this as basis, we seem to have a convincing account for a moral justification for civil disobedience. In response, I raise a series of objections related to its sufficiency.

The final part of the thesis examines potential political justifications for revolution and civil disobedience. Here, I introduce additional aspects of Kant's political philosophy which shed additional light on the problem at hand. Following this, I conclude that there cannot be a Kantian theory of civil disobedience because there cannot be a concept of civil disobedience under a Kantian framework. The final section considers how we might conceptualise the qualities found in civil disobedience under the Kantian framework.

What is civil disobedience?

Definition

For the purposes of this thesis, I will rely on a definition of civil disobedience as set out by Rawls, which reads as follows: civil disobedience is "a public, nonviolent, conscientious yet political act contrary to law usually done with the aim of bringing about a change in the law or policies of the government" (Rawls 1971, p. 329).

The civilly disobedient act need not breach the same law it protests, only some law. For instance, if I oppose what I believe to be a discriminatory law by breaking into public office, there is an asymmetric relation between the law I break and the law I protest. But equally, the two can coincide, such as Rosa Parks' protest of discriminatory laws executed by breaching one of those laws. Further, whilst typically engaged in after legislation has passed, one can mark one's resistance to a statute which is not yet in place, even if it should be constitutionally upheld later. However, protests before legislation has passed would usually fall underneath legal means of protest or discussion (provided no laws are broken) and thus not take the form of civil disobedience. I believe protests prior to passing legislation which breaches laws would be considered an illegal protest rather than civil disobedience, and hence a distinct category. The point of civil disobedience is, after all, to express one's dissent to an existing law or policy; it is not simply disobedience. Rawls includes a qualification that legal means should have been tried prior to civil disobedience, however, they need not have been exhausted as one risks an eternal cycle of rejection and appeal (Rawls 1971). The civil nature also commits one to nonviolence, and hence civil disobedience can often take on a symbolic form as opposed to merely forming a mob.

Further, Rawls argues that civil disobedience does not appeal to morality or religious acts; it is necessarily a public act (Ibid). In short, this allows us to set aside specific quarrels and opinions between private persons. For instance, the law may protect the right to religion and allow people to practice their faith freely, however, one specific religion could not rightfully, nor by civil disobedience, argue for their specific faith to be dominant by law. The wider point is that one may be dissatisfied with certain laws on an individual level, but that alone is not enough to warrant civil disobedience; the law is not there to suit my every desire, it is there to protect and serve the public. However, if the state put a blanket ban on any religious freedom, that would

be grounds for civil disobedience as it violates what Rawls would identify as a fundamental right through his veil of ignorance.

Finally, Rawls argues that acts of civil disobedience express a willingness to accept the legal consequences that may follow. That is, one engages in civil disobedience fully aware that one can be legally punished for this action, it is not an act from ignorance. Thus, civil disobedience does not aim at undermining the juridical system or the state as a whole, only to express dissent from the omnilateral will which supposedly gave rise to a given law or policy. This structure will be discussed in the next section.

Rawls also distinguishes between civil disobedience and conscientious refusal. While he recognises that there is some overlap, and that one need not preclude the other, conscientious refusal is "noncompliance with a more or less direct legal injunction or administrative order" (Ibid, p. 323). This, unlike civil disobedience, can be grounded in religious or ethical beliefs, as it does not necessarily invoke a sense of public justice. Evasion also falls under this category, which can demonstrate more of a personal refusal rather than a bid to change the law. To take a relatively innocent example, suppose I am riding my bicycle in a quiet area with no cars or pedestrians nearby. I am aware that I should still follow traffic rules, but I cannot be bothered to wait for the signal to turn green, and so I carry on pedalling on red. This could be considered an act of conscientious refusal or evasion as I let my pragmatic desires trump the traffic rules due to there not being anyone to stop for. However, in doing so, I do not intend to signal that the law should be changed and that we should disband all traffic rules. Rather, it is a minor offence, done for a personal and not public reason, that crucially does not negatively affect others. In sum, civil disobedience expresses a desire or need to change the law it protests and it does so on public grounds.

Potential cases of civil disobedience

In this section I present cases that we would intuitively categorise as cases of justified civil disobedience as a means to conceptualise just what civil disobedience is. These examples will be referred to throughout the thesis as test cases.

The suffragettes

The suffragettes' fight for women's suffrage took many forms ranging from fully legal discourse to violent acts. Consequently, many of their actions would fall outside the scope of the definition of civil disobedience. One event, however, that should intuitively qualify is that of the Grille Incident in which members of the Women's Freedom League smuggled a banner demanding rights for women into the Ladies' Galley in the House of Commons and hung it from the grille covering the galley windows. Two women proceeded to chain themselves to the grille, which subsequently had to be removed to get the women unchained. This protest was non-violent, but contrary to law and of a deeply symbolic value due to its demanding women's suffrage in the very Houses of Parliament that they were not allowed to participate in, only watch.

Hiding Anne Frank

Anne Frank and her family have become representatives of the many Jews that were hidden during World War II. This case, however, captures the rightfulness of those who hid the Franks and other Jewish families. The people offering their assistance were not inherently at risk themselves, but still chose to jeopardise their security by helping Jews. In doing so, they were breaking the laws imposed by the Nazis, yet were simultaneously doing something we today must unanimously consider as right and moral.

Rosa Parks and civil rights

Rosa Parks was one of several protestors who deliberately refused to give up her seat at an Alabama bus for a white person. Although this was later found to be in conflict with the US constitution, the local Alabama law decreed such discrimination. Hence, Parks' action was contrary to law, non-violent, and had the aim of changing the law that allowed racial discrimination and intuitively stands as a clear example of civil disobedience.

Gandhi's actions for equality

Mahatma Gandhi spent the majority of his life actioning for equality within India and independence from British Rule. His means consisted of boycotts and non-cooperation that were contrary to British law but always non-violent in nature. Although more violent events

later unfolded in Gandhi's name, Gandhi himself only every advocated for non-violent resistance and thus serves as another intuitive example of civil disobedience.

The vigil for Sarah Everard

A more recent example would be the vigil for Sarah Everard after her brutal murder by a serving police officer. The nature of the crime became a symbol for the many charges of discrimination and inadequacy of the Metropolitan Police, and the participants argued that the streets of London were not safe for women. While protests and large gatherings are normally lawful when permission is obtained, this vigil – a mass gathering outside on Clapham Common – was during a government lockdown which temporarily restricted this right. As a result, many of the participants were found to be breaking the law and were arrested by the very same police force they were expressing their discontent with in what many considered a disproportionate response to a non-violent event. Hence, this would be an example of civil disobedience, as they were breaking the lockdown rules which were legally enforceable by the police, yet the aim for the vigil was to highlight the violence against women in particular and the inadequate response from the Metropolitan Police.

Climate justice campaigns[1]

Other recent examples would be campaigns and actions from various climate justice groups such as Extinction Rebellion, Just Stop Oil, Insulate Britain and so forth. The exact means vary depending on the specific campaign and group, but their goal is always related to demanding more action to combat climate change, and they use disruptive means that are disruptive and contrary to law. However, their means are significantly more aggressive than the other examples mentioned, and one might therefore argue that they do not comply with the criterion of non-violence. In response, the campaigners might argue that the urgency and severity of climate change is significant enough to warrant such means. Whilst I think this example truly highlights the tension between what is morally right and what is demanded by law, the scope of climate justice and climate ethics extends beyond the aim of this thesis. In short, there is simply too many factors and considerations from the field of environmental ethics that extends beyond the scope of this thesis to fairly consider this example. Hence, climate justice campaigns will not

[1] This section addresses the intuitive appeal of climate justice campaigns, before bracketing them off for the purpose of this thesis. Hence, this example will not be used as a test case.

be considered as part of the set of intuitive cases of civil disobedience in this thesis, as the topic warrants a much more thorough exploration than what is directly related to the question of Kant and civil disobedience. For an introduction to environmental ethics, please see Lee (2022).

Kant's political philosophy

In order to address the question of the possibility of a Kantian approach to civil disobedience, we must first establish some basic tenets of Kant's political philosophy. This section will only establish what is necessary to understand and answer the question, and is thus limited in both scope and aim. It should not be considered a broad and critical assessment of Kant's political philosophy as a whole, only exegetical.

The innate right to freedom and the state of nature

Kant positions himself in the tradition of Locke, Hobbes, Rousseau by using the idea of the state of nature as a term and basis for what human society looks like, without the features which makes it a civil condition. The state of nature is thus not an empirical claim about human history, but a tool in conceptualising how we achieve the civil condition. Kant begins his argument for this from the assumption that there is only one innate right, and this is the right to freedom for all humans:

> "Freedom (independence from being constrained by another's choice), insofar as it can coexist with the freedom of every other in accordance with a universal law, is the only original right belonging to every human being by virtue of his humanity" (MM 6:238).

This independence of being constrained by another's choice is subsequently defined as including the innate equality of humans allowing them to be their own master, and being beyond reproach "since before he performs any act affecting rights he has done no wrong to anyone" (Ibid). Since there is only one innate right, every other right that may follow is an acquired right, such as rights to property. Kant thus differentiates himself from natural rights theorists by reducing the number of innate (or natural) rights to only freedom. From this basis, any acquired

right must be consistent with the innate right of freedom, which consequently provides some restrictions on the *form* that acquired rights may take.

In the state of nature, the right to freedom exists, however, it cannot be enforced. This is because the state of nature lacks the coercive authority necessary to enforce laws. There may be private relations and agreements that allows some distribution between land and resources, consequently we can find society in the state of nature, but in the event of conflict there is no public arbitration available to settle disputes in accordance with the law (see e.g., footnote in MM 6: 246). This means that any private agreement can only be upheld as long as the other party is willing to do so, effectively making each person dependent on the private will of another. Should the other party not fulfil their obligations, there are no legal sanctions available in the state of nature. Kant argues that we can have divisions of what is "mine or yours" in the state of nature, but that such distinctions can only be provisional. That is, one can only have provisionally rightful possession of something external as one's own, any conclusive possession can only be realised in the civil condition (MM 6: 256-7). Thus, the provisional rightful possession of this can only express a unilateral, or private, will.

Theorists sometimes use innate freedom to argue that the civil condition or society is essentially an instrument for coordinating private wills, that is, a system of private freedom. Such views often look at the civil condition as having certain benefits that we consequently trade our freedom for. Kant, however, makes an argument strongly in favour of the necessity of the public as such. Ripstein illustrates this by reconsidering roads from a Kantian perspective to address a structural problem with a system of equal private freedom. In this thought experiment, Ripstein asks us to assume a system in which all land is privately held, each plot belonging to a different individual, divided equally in both size and quality (Ripstein 2009). Following the postulate of private right, each individual is entitled to do as they see fit with their land (Ibid). Under such a system, we quickly encounter a problem with the distribution of land, namely, the fact that land is immovable, and so, "to get from one location to another, you need to traverse all of the locations in between" (Ibid, p. 245). This means that you would need the landowner's permission to cross each patch of land on your way and back, and each person is entitled to exclude you from passing through their land. As a result, you risk being landlocked in your own land, or even being unable to get back to your land should one neighbour refuse to let you cross their land on your way back.

This highlights a structural problem with a system of private freedom: your freedom of movement is dependent and restricted by the goodwill of others. You might have a very kind

and sociable neighbourhood that generally allows the use of their land, but there is no rightful guarantee for them to uphold this permission. Effectively, you are no longer independent from the choice of another and so you are no longer able to exercise your innate right to freedom. One might object by claiming that if one neighbour refuses to let you pass, you can simply ask another. But the point remains: there too, you are dependent on someone else's goodwill to let you pass. Moreover, this second points highlights the structural problem of having no public spaces: if every piece of land is privately owned, the owner is fully entitled to block you from it at their leisure. The only way to guarantee free passage, and by extension free association between individuals, is through public roads: "With a system of public roads in place, everyone can enter into voluntary transactions with whomever they wish, without being subject to the choice of any other person" (Ripstein 2009, pp. 248-249). Rather than seeing this as a compromise on private freedom, Ripstein argues that public roads actually gives you, along with everyone else, "an entitlement to get from place to place," a right you could not have in a system of private freedom (Ibid, p. 250).

Crucially, this example only showcases the structural problem with a system of private freedom, and thus only gives us the *form* of how decisions of the public must be made. That is, Kant does not aim to instruct us on how roads should be build, by whom, and with what material; it merely follows from the innate right to freedom that an entitlement to get from place to place can only be conclusively secured through public roads. Hence, Kant does by no means aim to prohibit private land, nor would he prohibit all private roads; conceivably one could have a private road leading to one's driveway that would indeed only be used with the owner's permission. However, as discussed, a system based *only* on private roads would reveal a significant structural fault. The solution to such systems for Kant is to enter the civil condition where private rights may be conclusive and public right may exist.

Entering the civil condition

A system of private freedom as sketched above, would generate what Kant calls the postulate of public right. This is "when you cannot avoid living side by side with all others, you ought to leave the state of nature and proceed with them into a rightful condition, that is, a condition of distributive justice" (MM 6:307). By a rightful condition, Kant means the civil condition, which is subject to distributive justice (MM 6:306). On this view, an "action is *right* if it can coexist with everyone's freedom in accordance with a universal law" (MM 6:230). A rightful condition

is a system of reciprocal limits on freedom, in which one can only exercise one's freedom to the extent that it does not constrains that of another, thus connecting "reciprocal coercion with the freedom of everyone" (MM 6:232).

Our incentive for entering this condition is what Ripstein refers to as three defects that we find in the state of nature. The first is the problem of unilateral choice, the second the problem of enforcement, and the third the problem of indeterminacy. The first becomes clear as soon as we acquire some property. Your acquisition of some property is a unilateral act that generates your provisional rightful possession. Following that, you are entitled to exclude others from accessing or using your property, however, "the mere fact that you act unilaterally raises the question of how that action can bind *me*" (Ripstein 2009, pp. 150-151). That is, a unilateral will cannot rightfully bind others and thus any act on my part that might be compatible with your unilateral will is voluntary and may be withdrawn at my leisure. This means that although a system of private freedom mentioned above - a result of such unilateral acts - might function practically, but there is no guarantee or stability for this function to be upheld, as there is no obligation for other parties to be bound by your unilateral act. Kant's solution to this is to introduce an omnilateral will, which allows for a united legislative authority which can rightfully pass laws in the name of everyone.[2] Only an omnilateral will can authorise and make unilateral acts conclusive. A public act must be omnilateral as it is not for some individual's end but instead "exercised on behalf of all citizens considered as a collective body" (Ripstein 2009, p. 196). A public act thus aims to provide the form for what the rights in a rightful condition are, which must be decided by the omnilateral will.

The problem of enforcement, simply put, is that unilateral acts provides to assurance, that is, "nobody can rightfully be compelled to serve the purposes of another unilaterally" (Ripstein 2009, p. 164). As with the problem above, there is a lack of obligation to respect any claims made in relation to acquired rights such as property. However, it is not merely an issue of binding others based on your unilateral acquisition, it is also a problem of being able to enforce the acquisitional right you assume you have. In the state of nature, any right or acquisition is merely provisional and so "your entitlement to use force is limited to the case in which interfering with your possession thereby interferes with your person" (Ibid, p. 165). This clause remains independently of a rightful condition as it is part of one's innate right to freedom from being constrained by the choice of another. We might accept a norm not to interfere with the

[2] Kant also uses the term *"omnes et singuli"* to highlight that by "everyone", he means everyone both as individuals and as a united entity. See MM 6:315.

possessions of another, however, a norm does not guarantee enforcement in space and time. Consequently, if we wish to enforce external rights such as property right, we need assurance based on an omnilateral will that secures the division between what is mine and what is yours: "you have what is yours, because if another wrongs you, you will be able to get it back" (Ibid, p. 167). With a system of private freedom there can be no such assurance, because there is no public system of enforcement of rights.

Finally, the problem indeterminacy recognises that rights are subject to dispute. Consequently, we need some form of arbitration that, just like our rights, is omnilateral and universal, that is, it applies to everyone and it is the same for everyone. The indeterminacy problem lies in the inconsistency of attempting to resolve such disputes without an omnilateral will, as individuals or groups of individuals will necessarily disagree on the interpretations of the terms of their agreements. The solution is having a judiciary, "a body that has omnilateral authorisation to apply the law to particular cases" (Ripstein 2009, p. 172). That is, a public institution which is rightfully authorised to decide on behalf of everyone and may thus settle disputes both rightfully and consistently.

Following these three defects, we not only have three incentives to enter a rightful condition, but also one which entails a republican institutional system of three branches comprising the sovereign: the legislative, the executive, and the judiciary (MM 6:316). This distinction is made on non-instrumental grounds, the justification of which is well formulated by Ripstein and thus worth quoting at length:

> "Anything that the state does has to be properly authorised by law: the making of law, the taking up of means to give effect to the law, and the passive classification of particulars. Failure to separate the legislative from the executive function turns into a form of despotism through which some rule over others. The failure to separate the judiciary from the executive and legislative branches creates another version of the same problem: a dispute can only be resolved consistent with the right of the parties if its particulars are brought under a general rule; if the rules can be changed in response to a particular case there is only force, not law" (Ripstein 2009, p. 175).

Hence, all three branches of the sovereign are dependent on one another and thus a civil union must have all three to be considered a rightful condition, and consequently the state is necessary for any rightful interaction between individuals.

Sovereignty and the structure of the state

We have so far explored the necessity of entering a rightful condition and the need for a sovereign state whose powers must be separated into the legislative, executive, and judiciary branches. This separation is necessary to provide the formal conditions for creating a civil union which can enforce "the conditions under which the choice of one can be united with the freedom of the other in accordance with a universal law" (MM 6:230). With the formal conditions for the existence of a rightful condition in place, we can now examine the possible constrains on the conduct of the sovereign.

Kant argues that the exit from the state of nature and formation into a state comprises "the original contract". While this is only an idea, not an empirical claim, this is what allows us to determine the legitimacy of a state:

> "In accordance with the original contract, everyone (*omnes et singuli*) within a people gives up his external freedom in order to take it up again immediately as a member of the commonwealth, that is, of the people considered as a state (*universi*). And one cannot say: the human being in a state has sacrificed a part of his innate outer freedom for the sake of an end, but rather, he has relinquished entirely his wild, lawless freedom in order to find his freedom as such undiminished, in a dependence upon laws, that is, in a rightful condition, since this dependence arises from his own lawgiving will." (MM 6:315)

The last sentence highlights a distinction between Kant's approach and the concept of a general will inherited from Rousseau. For Rousseau, the civil condition is an instrumental solution to ensure that all humans are able to exercise at least *some* of their freedom, which requires that we give up our initial freedom of self-governance. For Kant, however, the civil condition cannot be seen as a sacrifice but is instead something transformative which allows human beings to eventually exercise their full rational capacity in accordance with the freedom of everyone. In this passage, Kant does not claim that every legitimate state must be a perfect realisation of this idea, but rather that the original contract serves as a guiding principle for how a legitimate state should conduct itself.

Weinrib argues that the idea of the original contract provides a principle in which the public authority "must be arranged and exercised in a manner that adheres to its own justificatory logic" (Weinrib 2018, p. 34). That is, the state must be committed to improving the existing legal order "into the deepest possible conformity its own internal moral standard" (Ibid). As a

result, a legitimate state cannot undermine the innate right to freedom without simultaneously undermining its own authority. Thus, the original contract can show how it is not merely a case of the citizens being subject to rule, but that the sovereign has certain obligations to perform: "Specifically, it obligates every legislator to formulate his laws in such a way that they could have sprung from the unified will of the entire people and to regard every subject, insofar as he desires to be a citizen, as if he had joined in voting for such a will" (TP 297). This gives the sovereign a clear criterion for how laws must be formulated in order to be rightful, and consequently if a people could not possibly agree to a law, it cannot be considered rightful.

Further, the structure of the state and constitution must be republican (TPP 350). This, Kant argues, is because republicanism is the only system which conforms to the innate right to freedom alongside reciprocal coercion, and the only form which ensures that the law is authorised by a common united will. The latter is justified through the idea of the original contract and the need for an omnilateral will with a division of state power as discussed in the previous section. However, perhaps surprisingly, Kant does not argue for democracy in any direct sense. This is because he considers it a form of despotism because it sets up an executive power which is also the legislator which creates internal inconsistency when there is dissent or disagreement. The only way a people could possibly be bound by a universal law would be if they represent the legislative but not the executive function of the state. Consequently, "[e]very form of government that is not representative is properly speaking without form, because one and the same person can no more be at one and the same time the legislator and the executor of his will" (TPP 352).

Finally, Kant gives a practical justification for his republicanism:

> "Since unanimity cannot be expected of an entire people, the only attainable outcome that one can foresee is to obtain a majority of the vote, and (in a large population) not even a majority of the direct vote, but only a majority of those delegated as representatives of the people. Thus, we will assume that the principle of allowing this majority to suffice [for legislation] has received universal agreement, that it is a matter of contract, and that is must be the ultimate basis for establishing a civil condition" (TP 296).

Within the realm of practical reason, it is unreasonable to assume that each individual citizen can reach an agreement of what should constitute law, as this was an issue we found in the state of nature. Hence, in order to ensure the practical capacity of enforcing and executing laws, a

representative system of majority votes is not only practical but justified on Kant's institutional basis.

Kant and revolution

Revolution and civil disobedience – a possible analogy

With the preliminaries of Kant's theory of the state in place, we can begin examining sources of tension. This section sets out the problems related to Kants extreme stand on revolution and the potential inconsistencies we may find with his view.

In a rightful condition that ensures that the form of the state's decisions is correct, Kant argues that the citizens have a duty to obey the coercive laws. More controversially, "all resistance to the supreme legislative power, all incitement of subjects actively to express discontent, all revolt that breaks forth into rebellion, is the highest and most punishable crime in a commonwealth, for it destroys its foundation" (TP 299). Kant continues by claiming that this prohibition is absolute (Ibid). Following this, citizens have no right to resist unjust laws or even express discontent, and thus there is absolutely no room for justifying active resistance. Kant's treatment of the subject and most commentary considers only the case of revolution, and not civil disobedience. Whilst the aim of a revolution is different, it still shares some common ground with civil disobedience that could provide insight into how we might justify civil disobedience.

Based on the quotation above, I take revolution to be active, often violent, resistance against the sovereign, with the aim of overthrowing the existing constitution. Civil disobedience, on the other hand, is "a public, nonviolent, conscientious yet political act contrary to law usually done with the aim of bringing about a change in the law or policies of the government" (Rawls 1971, p. 320). The difference between the two concepts is thus found both in the end and the means used to achieve that end; civil disobedience is necessarily non-violent and only aims at challenging a law within the existing constitution, whereas revolution often uses violent means to overthrow the constitution. However, they differ mainly in degree and not kind, as both revolution and civil disobedience use extra-legal means to achieve the end of challenging the state's authority either in part or completely. Hence, if we can find adequate grounds for justifying revolution, this should consequently also apply to civil disobedience.

Kant's opposition to revolution and its legal grounds

Kant's absolute prohibition is particularly puzzling when one considers the fact that Kant himself was an avid supporter of the French revolution, often being referred to as "the Old Jacobin" (Korsgaard 2009, p. 236). While we tend to accept some inconsistencies between a historical philosopher's personal beliefs and their philosophical theses, the fact that Kant in theory was such a strong opponent to revolution, yet in practice was very supportive of the French revolution does raise the question of whether there could be some room for justifying revolution after all. The question is thus twofold; there is first a question of consistency in how Kant could hold these beliefs simultaneously, and second, a question of how, if at all, revolution can be justified.

Kant's immediate justification in *Theory and Practice* is the structure of the scope of power for the sovereign or the head of state. The people of a commonwealth *cannot* have to coercive rights over the head of state, as in order to be subject to those laws someone must have the authority to give laws. As we saw in the state of nature, a private agreement between two equals holds no assurance as this is something only a universal authority can give. In order for such an authority to be rightful, it must go through the institutions comprising the sovereign, and through the idea of the original contract the people can envision themselves as being both authors of and subjects to the law. However, despite having gone through the process of a universal law being made rightful, the enforcement and judiciary of the law cannot rest with the people themselves, as this too results in a form of despotism. In order for the public to be subject to coercive laws, there must be an authority that can formally decree them. Hence, the head of state secures the rightful institutions by taking on the role as the symbolic author of the coercive law and is not bound by coercive laws himself although he has an obligation to act in accordance with the original contract.

Wit (1999) argues that there is some textual confusion surrounding the head of state and whether they are to be considered a physical or moral person. If we consider the head of state a physical person, then there are no good reasons for why they should not be subject to coercive laws as everyone else. If, however, we consider

> "the sovereign in the *moral* sense of the term, that is, the sovereign *as giving the law to herself* (and to those who are subordinate to her, the people), or, the sovereign *in the function of* a sovereign, regardless of who she is, she cannot by any *external* force be coerced by the law because she gave it to *herself*" (Wit 1999, p. 291).

Following this interpretation, the head of state is merely a role or function that a physical human being can carry, and note a modal property of personhood bestowed on e.g., an arbitrary lineage of presumed royal blood. Because of the symbolic nature of the head of state, I argue, contrary to Wit, that this allows Kant to consider a rebellion against the head of state the same as a rebellion against the constitution. The head of state is not physically the constitution, that would be absurd, but they are the symbolic representative of the constitution, and so a challenge to the authority of either is a challenge to both. Therefore, I take revolution to be a rebellion against the sovereign with the aim of overthrowing the constitution.

Although Kant's personal view is puzzling, his argument for the logical necessity of prohibiting a right to revolution is relatively simple and can be found in *Theory and Practice*. The constitution, either in its pure or representative form, is the basis of the civil condition. A right to revolution, defined as aiming to overthrow the constitution, would contradict the very idea of a coercive authority, and thus cannot be permitted. Kant's argument consequently takes a legal form, essentially asserting that a law permitting self-destruction is inconsistent with the concept of a state authority which cannot possibly endorse such an end. As a result, any sound understanding or justification of revolution cannot be made on a legal ground.

The French revolution

Given Kant's personal conviction, some treatment of the French revolution is warranted. The French Revolution was not a singular event as one can look at it as two instances of revolution: one in 1789 and one in 1792. This division is not uncommon, but Maliks (2022) argues that the difference is a crucial way to understand Kant's position. The 1789 revolution was a peaceful, procedural transfer of power, in which the King abdicated his executive function. The 1792 revolution was a violent rebellion with the people taking the law into their own hands, culminating with the execution of the former monarch. Maliks argues that Kant endorsed the first revolution, but rejected the second.

A question which immediately arises from this approach is whether the 1789 revolution can in fact be considered a revolution due to its procedural character which intuitively is closer to reform. One way to account for this is through Kant's distinction between different forms of transformative political action. Here, Kant uses *metamorphosis* as the term for reform through legislation, allowing the state to continually exist under change, which Maliks argues is applicable to the 1789 events (Maliks 2022). This stands in contrast to *palingenesis*, which is a

new beginning after a mob overthrows the existing powers, in which society is returned to the state of nature. *Palingenesis* would hence apply the 1792 revolution.

Maliks argues that the reason why Kant prefers *metamorphosis* is due to the supremacy of law. That is, the legal system has the final say as the authorised judiciary of the omnilateral will. If this legal authority can be overruled, there is no conclusive way to establish law. Thus, if one wants to change the law, this must be done through established procedures, otherwise the result will be unlawful and unjust. *Metamorphosis*, then, allows for reform and progression within the established framework, which on Maliks' interpretation includes the effective abdication of the King in 1789, without resulting to abolishing the state and returning to the state of nature.

However, the unlawfulness of *palingenesis* only applies to the existing or overthrown system. And so, if a *palingenesis* comes about, it will be unlawful in the existing sense, however, that does not mean that the new state cannot create a new legal system. If it does, then that will take on the role as having the judiciary authority. This reflects what Kant holds about the formation of states, in that this has rarely, if ever, happened in a procedural and lawful way, but the historical origin does not release us from the obligation to create a true republic afterwards (MM). Consequently, as we always exist within a legal structure of society when we are in a rightful condition, it will always be unlawful to overrule the legal authority and incite a revolution or *palingenesis*. Hence, Kant denies a right to revolution.

Maliks then discusses how this frameworks applies to the French revolution.[3] While *metamorphosis* is preferable to *paligenesis*, it is dubious that the 1789 event did not involve the dissolution of the state. This is because the very idea of temporary sovereignty is a contradiction. Maliks demonstrates that the French King's believed he only granted executive power to the assembly temporarily, and could thus reassume his power once their task was finished. But this does not work logically, as one cannot both be sovereign and not sovereign at the same time. Thus, as soon as sovereignty is alienated, the king is now the *former* sovereign, and he no longer has any claim to reclaim it. The people he designated as sovereign in his place are now the actual sovereign, and so only they can choose to abdicate or bestow sovereignty upon him again. Moreover, sovereignty cannot be divided and distributed at will while maintaining sovereignty. Kant likens this to treating it as a property right, which can be alienated as it bears no personal connection to the agent. But sovereignty is a personal right, meaning that it is inherently tied up with your person. This is most obvious in the form of a despotic monarch,

[3] For a detailed discussion of the events of the French revolution, see Maliks 2022.

in which his bidding is in fact law; he *is* the state. Thus, "treating sovereignty, or the state, like property implies setting a private will above the general will" (Maliks 2022, p. 34). This means that the will of the sovereign as an individual trumps that of which he represents and embodies through the omnilateral will through his abdication as sovereign.

Finally, Maliks examines the unlawfulness of the French King's execution in 1792. Kant considered regicide the most serious crime related to revolution due to the monarch's identification with the constitution (see e.g. TP 300). Some might argue in response, that he was not on trial as king, but as Louis Capet, an ordinary citizen. However, according to Maliks, Kant argues that he was tried for actions that occurred when he was sovereign king, not for any actions or alleged crimes happening after his rule (Maliks 2022). Thus, he was essentially on trial for his perceived wrongdoings as king. However, the sovereign can do no wrong. This is meant in a strictly legal sense; a sovereign can undoubtedly make poor decisions, but that does not make them illegal, because such decision-making and legality is part of his sovereignty (TP 299-300). And so, claiming the King committed an illegal action which were in fact within the scope of his power becomes a logical fallacy. This is probably also why Kant saw the trial as a mere pretext for justifying the execution, which was in fact driven by fear of a counter-revolution. This attempt of disguising murder as justice was thus particularly to Kant's distaste, as it was not only immoral, but also a perversion of the very principle that drives justice. For this reason, Kant considered regicide as diametrically opposed to justice, despite his sympathies for the initial spirit of the French Revolution.

The moral justification for revolution

The previous chapter considered Kant's legal basis for opposing revolution, in addition to a brief interpretation of Kant's understanding of the French revolution. From this, it became clear that there cannot be a legal right to revolution, as this is contradictory to the very concept of having enforceable rights. Hence, any justification for revolution must be extra-legal. This chapter examines the potential for a moral justification for revolution. The chapter is based on the assumption that potential solutions for Kant's opposition to revolution will also apply to civil disobedience.

Korsgaard and Flikschuh: The supra-moral revolutionary

Korsgaard argues that we must accept that legitimate governments are imperfect without that endangering their legitimacy: actual governments are "imperfect participants, in the Platonic sense, in the form of justice, a form that is given by the ideal of the republic" (Korsgaard 2008, p. 246). Korsgaard consequently accepts Kant's idea of a no right to revolution, but argues that there are still cases where a good person might have to take the moral law into their own hands. Korsgaard's interpretation of the duties of justice, duties of virtue, and revolution, are all compelling, which makes her conclusion all the more sudden and confusing. She writes that the "imperfections of the actual state of affairs are no excuse for revolution – if they were, revolution would always be in order" (Korsgaard 2008, p. 258). However, in the following paragraph, Korsgaard outlines an analogy that appears to be integral to her argument, namely, that there are cases where we might rightfully infringe one someone's autonomy to save them from themselves, in cases such as addiction. Because there is no clear *a priori* requirement we can use to decide when such intervention is necessary, it becomes an act of paternalism (if directed at a person) or a revolutionary (if directed at a government), and we find ourselves in a territory in which neither the civil nor moral law can offer us the substantial guidance we need, leaving the agent entirely alone in their decision-making (Ibid., p. 259).

It is unclear why we should accept that saving someone from themselves could be analogous to saving a state from bad governance. Cases where we might legally infringe on someone's autonomy are extreme, and usually only a result of the person posing a significant risk of harm to themselves or others. What is implicit is that the person in question is no longer at their full rational capacity, which can invoke Korsgaard's use of paternalism. The argument is then that a government that similarly falls below its rational capacity may also require intervention. But this raises the question of who am I as an individual to interfere with this? Even if we do recognise that the government is in a sense "sick", what could possibly give me as an individual the right for this kind of paternalistic or revolutionary interference? In the case of a person, the interference is legitimised on account of a doctor or someone of a full rational capacity, perhaps even with the ability to enforce the civil law, it is no longer a relation between equals but one of rightful dependence. However, by assuming the same position as an individual towards the government, the revolutionary suggests that they have access to some superior rational capacity that allows them to surpass the general will.

This thought may be precisely what Korsgaard is arguing for, as her afterword invokes a Nietzschean will to power. However, it remains unclear why we should accept such a sentiment

within a Kantian framework. One consideration is the megalomaniac thought of knowing better than the general will *and* the moral law, but Korsgaard also argues that this intervention is rooted in a *feeling*, the same feeling which allows us to root for the hero that casts his principles aside and picks up a gun. Abandoning any rational principle and act merely based on a feeling or inclination of what is right completely abandons Kant's need for procedural justice, leaving us with no justification (moral, or otherwise) for such an act. In fact, this subversion of justice becomes akin to regicide which Kant opposes due to its extreme violation of both a person and a principle. Of course, Korsgaard is right when pointing out the injustices of the world that have not or did not become rectified through procedural justice in accordance with the moral law, but such examples are not enough to abandon the principles that make up Kantian legal and moral theory. Such an approach might be sufficient within a Nietzschean framework, but this kind of individualism based of a feeling of will to power is decidedly not Kantian and consequently unconvincing.

Flikschuh discusses similar issues with Korsgaard's interpretation and offers a potential way of solving the issues without falling back onto Nietzsche. Her critique encompasses my discussion above, but also questions the contractual interpretation Korsgaard employs:

> "The implication of Korsgaard's contractualist position then is that my claim to the object is valid insofar as others agree to recognise it as such, not, as Kant himself claims, that others are obliged to acknowledge the object as mine merely because I claim it as such on the grounds of my capacity for external freedom" (Flikschuh 2008, p. 131).

That is, any claim to an object is dependent on others being willing to acknowledge my claim. In relation to property, this problem is solved by creating a rightful condition. However, a similar story about the origin of nations is unconvincing. If a regime can only be recognised as legitimate if recognised as such by all individuals, establishing a legitimate regime becomes near impossible. From a moral point of view, such a thought is noble, after all a regime would probably be better if it did have such a recognition on an individual level. But this goes against Kant's own position on two accounts. First, the fact that Kant fully recognises that most regimes have been established in an unjust manner, but that this is no hindrance to their legitimacy (MM 6: 340). Second, the dependence on the approval by individual or private wills undermines the concept of the omnilateral, which is not merely an aggregate of the individual wills, but also its unification. This would further deprive any elected representatives of their power to act on behalf of the omnilateral will, as any such action would be dependent on the approval of all private wills. This might be possible if we envision a direct democracy, but this is far from

Kant's republican vision. Further, as Flikschuch notes, the collection of private wills does not amount to a people. A people can only exist in relation to a state, i.e., in the civil condition. Without the legitimacy of the state, a people is just a collection of private wills under a system of private freedom. This means they can be highly organised, but organisation alone is not sufficient for a legitimate civil condition and state.

Consequently, Flikschuh argues that Korsgaard's interpretation blurs the distinction between "legitimate but unjust regimes and those which are both legitimate and just" (Flikschuh 2008, p. 131). And without this distinction, Korsgaard misses out on the possibility for reform. If we begin from the assumption that a legitimate but unjust regime may exist, we can start to reform the regime towards a more just one without calling its legitimacy into question (Ibid, p. 139). The revolutionary, on the other hand, must always challenge the legitimacy of the regime and plunge us back into the state of nature before reestablishing a legitimate state. Constantly challenging the legitimacy of a regime is not a very effective way of change, especially not from a Kantian perspective as it undermines both our juridical and ethical duties.

Of course, it is important to note that Korsgaard does not argue for an ethical duty to revolt, rather, she argues that there might be cases where there are exemptions to the ethical duty to not revolt (Flikschuch 2008, p. 134). In other words, Korsgaard is by no means defending a positive right to revolt on ethical ground. However, she is defending the possibility of exemption from the ethical duty which results in her Nietzschean spin. The very idea that one sole revolutionary can achieve a supra-moral standpoint where he does not only make a decision on behalf of everyone else, but claims that his private will should be the will of everyone, is absurd within a Kantian framework.

Further, there seems to be an implicit assumption that the revolutionary will be good, despite using these means. But this hinges on the assumption that everyone agrees that the current regime is unjust, an approach that fails logically. If everyone agreed that the current regime is unjust, there is plenty of scope for reform under a republican constitution, hence there is no need to revolt. If, however, only the revolutionary believes the regime to be unjust, there is no basis on which he can legitimately claim this injustice; it will necessarily be his unilateral will.

It is easy to root for a hero that we believe is doing right, even when breaking the law, but "the fact that not everyone favours the same revolutionary is significant: revolutionaries make private judgements about what is just for everyone else" (Flikschuch 2008, p. 142). Flikschuch cites Khmer Rouge as a less favourable example of revolutionaries, and it demonstrates the fact

that we cannot make concessions to the revolutionaries we root for (whether fictional or real) without being forced to make similar concessions to those we do not root for. So either we let exceptions be part of our theory, regardless of which side they are on, or we do not allow for exceptions. The only thing we can know *a priori* about a given revolutionary is that they revolt against the state, which under a Kantian framework remains impermissible both from a juridical and ethical standpoint.

Hill: Approaching hard cases from a Kantian perspective

Another possible approach to a moral justification for revolution can be found by Hill, who sketches out possible circumstances where we might make exceptions without abandoning the principle. He uses the case of terrorists immorally threatening the lives of innocent hostages and examines what might be permissible responses from a Kantian point of view (Hill 1991, pp. 196-7). While this is not a case of neither revolution nor civil disobedience, it is a case where two equally important maxims - saving the lives of innocents and not harming another person - are being pushed, which might help us navigate similar difficult territory within a Kantian framework.

Hill begins from the starting point that the state has coercive power which includes enforcing violence when necessary, i.e., a monopoly on violence. From this, Hill notes that Kant was not opposed to the death penalty for the most serious crimes, due to the punishment being proportional to the crime (Hill 1991, p. 210). Hence, juridically speaking, the state could have a legitimate right to execute, and may consequently have a right to shoot terrorists keeping innocent hostages. The difference would be that in our case, the terrorists would suffer capital punishment without a trial, which could damage its legitimacy. Hill further highlights the state's coercive power as a necessary "hindrance to the hindrance of freedom". That is, "the coercive power of the state must provide incentives so that even without conscience everyone will have a clear and sufficient reason not to violate the liberty of others (as defined by just laws)" (Ibid, p. 209). This is consequently a way of externalising Kant's innate right to freedom. Because human beings do not always act in accordance with reason, they may consequently infringe on someone's autonomy and freedom. This is of course immoral, but cast under pure moral terms there is no external punishment which can be applied in response apart from the person feeling guilty. Hence, if we want to ensure that certain infringements on another's freedom is

punishable, we must cast it under legal terms as well, allowing the coercive powers to be enacted in a way that respects the autonomy of both the perpetrator and the victim.

Following this, Hill argues that there may be occasions where "hindering hindrances to freedom" will legitimately involve termination with extreme prejudice, despite its violation of the maxim to not kill human beings. Hill considers a variety of possibilities within his case, and writes that "the dignity principle implies a reluctance to kill or even risk killing human beings, and it implies that we should make every effort not to be forced into situations where these actions might be required" (Hill 1991, pp. 213-14). But assuming that every effort to preserve lives has been made, Hill argues that, following Kant's jus talionis, the dignity principle "does not absolutely prohibit killing human beings" (Ibid).

Setting our judgement on the rightness of capital punishment aside, Hill argues that Kant's rationale for this is in fact rooted in the dignity principle. That is, a legal system that "expresses its respect for the incomparable worth of each rational agent by seeking to secure for each, in advance of particular contingencies, a full fair opportunity to live as a rational agent" (Hill Jr., p. 211). Hence, violating this incomparable worth, or dignity, or another rational agent is so severe that the only proportionate legal response is capital punishment. This may at first seem like a paradoxical position, claiming that the dignity principle can be violated on grounds of the dignity principle. But this is why we must be clear about the assumptions in Hill's argument, namely, that this rests on the assumption that the state has done everything it should to prevent such a crime from happening, the law is available to the public and made through the general will, the law officials are not corrupt, and so the agent in question has been given a "full fair opportunity to live as a rational agent" (Ibid). Under such circumstances, if the agent still commits the crime of murder, then a formal execution could be justified, and, assuming the agent remains rational, this line of reasoning would be understood and accepted by the condemned. Such a punishment would consequently not be accepted in a corrupt and Kafkaesque legal system. Consequently, assuming these conditions for Hill's terrorist case, we can envision the possibility of this principle being extended so that in rare and extreme cases law enforcement can legitimately shoot a terrorist.

While this approach is interesting, it seems to be almost a mirror image of our civil disobedience. Rather than asking what is permissible for state officials to do in response to illegal and immoral actions, we are trying to find if there are any permissible grounds for committing such an action in the first place. Although my focus is on less extreme cases of civil disobedience, crucially cases of non-violent yet illegal acts, it is not implausible to envision a

case where civil disobedience is characterised as terrorism by the powers that be. And, as discussed earlier, we cannot use the ideological beliefs of a given revolutionary or dissident as grounds for our argument. So, let us assume that Hill's terrorist does not intend to kill their hostages, but only use them as leverage to get their case across, and that their case is one of noble intent to change society for the better. Is there any way we can justify this? The burden of proof certainly falls on the terrorist in proving that their unilateral will is in fact the omnilateral will, and that they are merely trying to ensure that this omnilateral will can be enforced. Such an argument would be implausible and at odds with the procedural character of the omnilateral will.

Hill is aware of this problem, and later begins sketching out potential ways of thinking about political violence "against one's government when it is perceived as being grossly unjust" (Hill 1997, p. 106). Hill uses a broad definition of political violence in which the aim is to achieve some political end and that the "primary aim is not merely profit, revenge, personal grudge, and the like, but at least in part to gain or retain control of legal and political institutions, to express an ideology, to gain or assert a perceived right, etc." (Ibid, p. 108). While Hill's focus is on more extreme cases when violence is used to achieve these aims, it is not unthinkable that civil disobedience could fall under some of these categories or be part of it as it would exemplify a less extreme case of a similar sentiment. Just like we have been looking at revolution, using more extreme cases can help us think about less extreme ones. Hill works on the assumption that all "nonviolent avenues of reform would have been exhausted, all reasonable appeals and compromises rejected" and that a "more just, less oppressive regime would very likely result from the change" (Hill 1997, p. 109). In other words, these acts fulfil Rawls's criterion of first attempting legal reform, however, he takes a different route when exploring potential solutions for such a case.

First, Hill argues that, at the very least, Kant should have left whether it is ever morally right to rebel an open question (Hill Jr. 1997, p. 115). While this is in part based on Hill's familiar position that Kant neglects empirical considerations and therefore suggests a more applied variant of Kantian ethics, this is still a relatively minimal claim that one could be inclined to support as Kant could in no way have foreseen some of the horrors of the 20th Century. Assuming that the question is more open than Kant would have it, Hill looks at whether the Categorical Imperative have the resources to answer the question, and how, if at all, it can offer us guidance in hard cases such as this. After some discussion, he mostly sets aside the Formula of Universal Law (FUL) of Formula of the Kingdom of Ends (FKE). On this basis, the argument

is that FKE shifts the focus from maxims and their consistency to a broader consideration of humanity as a whole (Ibid, p. 121). The Kingdom of Ends is "Kant's most explicit use of a political model for deliberation about moral principles" as it sets out our moral framework in which all members are both "authors of the law and subject to the laws" (Ibid, p. 128). This, of course, mirrors Kant's structure of the rightful condition and the omnilateral will. With the FKE as basis, any policy must honour the basic value of human dignity, which does not have a price, as an end in itself.

Hill then argues that "as a direct guide to particular action, the formula has 'gaps'; in dilemma-like cases, it pulls us very strongly towards opposite courses of action, without (by itself) giving advice on which course to take" (Hill Jr. 1997, p. 132). This effectively captures an intuitive problem that many have with Kantian ethics in that it is too abstract and offers little solution in dilemma-like cases. Assuming, for instance, in trolley problems, the Kantian directive is clear in that we should not sacrifice any human being, but offers no guidance on how to reason when not doing so is simply not an option. It is consequently within this space that Hill attempts to find Kantian ways of approaching such cases. What he finds, however, is only, as Hill himself admits, speculative (Ibid, p. 137). One possible outcome is simply to say that such matters are too complex for basic Kantian principles to grapple with. This is shown clearly in Hill's discussion of maxims, in which we risk having to go back and forth between maxims until we achieve our desired maxim. The problem with this is that in doing so, we take more and more empirical evidence into consideration, making the maxim less universal in the process. This, then, undermines the pure rational principle that is meant to guide FUL.

Similarly, we might choose to approach the matter by giving ourselves maxims with inbuilt exceptions. However, if our maxim can only be used if we make sure to build in a multitudes of exceptions, it is no longer a maxim we can follow as a principle. While there may be cases where other considerations matter, that does not mean we can inherently consider them as moral under certain circumstances. For example, lying will always be immoral on a Kantian view, even if it is done in circumstances that we would find excusable, such as saving a friend. Naturally, the latter is a good thing, but that still does not change the immorality of the lying in question, rather, you have now done one thing that is morally bad (lying) and one thing that is morally good (saving a friend). Similarly, from a legal perspective, stealing remains a crime regardless of the circumstances; we cannot have a law forbidding stealing except under circumstances a, b, c. There might be understandable reasons for why someone did steal, but that does not make it moral, something that Hill does recognise (Hill 1991). Morality with such

a broad scope for exceptions is not morality on a Kantian view, although I am sympathetic to Hill's view of the ever increasing complexities of the moral dilemmas we face.

Following this, it seems that Hill is unsuccessful in deriving clear moral guidance from Kant in the case he outlines. Consequently, he suggests a route in which we are more lenient with the Kantian moral principles in the face of empirical reality as with "too much optimistic faith in inevitable progress and too little tolerance for exceptions in principles, Kant himself went too far in his absolutist stance against political violence" (Hill 1997, p. 140). It is important to note that Hill does by no means advocate for a positive attitude towards political violence, but rather an understanding that despite the tragic human cost, such cases happen and we ought to know how to best respond to them; which for Hill points to some lenience in principle. While such lenience offers a sort of solution to cases of political violence, the solution is one that involves a compromise on both Kantian morality and his general and, admittedly, absolutist framework. Hence, there are two ways of evaluating Hill's proposal. One is to accept that Kantian morality has its limits, and in order to achieve some response to hard cases, we should be willing to accept a level of lenience in our guiding principles; they are, after all, guiding principles, not absolute laws; and the moral law cannot be enforced. Another would be to reject the compromise on this very basis: Kantian morality is indeed all about absolute laws, and it is following them autonomously that makes us free.

The potential of *Universal History*

While Hill's approach is substantial and promising in many ways, I do not think we should accept his solution of compromise, as it goes against the most fundamental aspect of Kantian morality and thus risks undermining the very principles it seeks to uphold. Further, Hill repeatedly emphasises the lack of empirical uptake of these principles, for instance when writing that "though it might be useful to imagine what principles would be reasonable if everyone followed them, we have to modify these in the light of the fact that we know that not everyone will" (Hill 1997, p. 136). This clearly demonstrates Hill's key position that lack of empirical uptake can justify compromising on our principles. This is a pragmatic approach which in some cases would be laudable. However, it is not clear why someone as well-read and familiar with Kant as Hill neglects the potential found in an *Idea for a Universal History,* especially considering that Hill mentions this work in his footnote on page 125. Following this, I think the issue is not so much about our stringent principles but rather our human failings, as

the crooked piece of wood that we are. However, that is not to say we cannot make progress. On the contrary, that is exactly what Kant argues for in *Universal History*, as we move closer and closer to a fully just and moral civilisation.

A crucial claim in *Universal History* is that mankind's capacity and use of reason can only be developed on a species level, not an individual one (*Idea*, 19). This means that there will not be one individual of supreme reason (or a supra-moral capacity), but rather that we all as individual contribute to this aggregate of developing reason. Moreover, this development does not happen within a generation, but requires time as "reason itself does not operate on instinct, but requires trial, practice, and instruction in order to *gradually progress* from one stage of insight to another" (Ibid, my emphasis). What drives this development is not some Hegelian *Geist*, but rather man's own *unsocial sociability*, which is our "tendency to enter into society, combined, however, with a throughgoing resistance that constantly threatens to sunder this society" (*Idea*, 20-21). This can be demonstrated with our often conflicting desires of wanting to live together in a peaceful society while at the same time also wanting everything to be organised according to our own desires, making a "civil society administered in accord with the right" both mankind's greatest achievement and challenge (Ibid).

Universal History argues that the unsociable nature that is disciplined results in developing arts and culture as a way of civilly expressing those desires. However, one could argue that a similar expression (albeit less restrained) can be found in acts of civil disobedience or revolution. That is, even though we desire a rightful civil order, every so often someone will force their private will upon the public by means of a revolution, thus creating discord. By pairing this with the gradual species development of reason, Kant could argue that a revolution can signify progress from a historical perspective, without having to commend it as a singular event. Such an approach could allow us to say that changes in agriculture, cartography, revolutions, discovering oil, developing capitalism, etc., all are part of this gradual development without saying that one event was particularly immoral.

This development would result in humanity coming together as a "moral whole" (*Idea*, 21). Kant does not elaborate much on what this entails, but it is reasonable to assume that this would in fact be a Kingdom of Ends in which we are all members of a moral "Kingdom" where we do not treat other human beings as a mere means to an end (GM). From this, it seems that Kant's ultimate hope for history is that humanity will develop their morality so that this internal legislation of morality can also be achieved externally. This development can be seen as occurring in multiple stages where first the individual must emerge from "his self-imposed

immaturity" (E, 35). Following this, a people can ensure that their society reflects what is morally right so that it is elevated from the individual to a people. Once we have a state of morally good people, we can begin aiming for a "federation of peoples", i.e., a world that is governed and lived in accordance with both the moral law and the postulate of public right, resulting in the Kingdom of Ends.

However, Kant recognises the practical difficulty in achieving this as the "perfect solution is impossible; from such warped wood as man is made, nothing straight can be fashioned" (*Idea*, 23). This can be interpreted as pessimism suggesting that there is no point trying to achieve this goal. However, I would argue it is more accurate to interpret this as a normative claim arguing that our current state is far from imperfect, and perhaps we will never achieve it, but that does not mean we should not try. But verifying these claims would require a deep empirical investigation into human psychology and the future of the world as a whole, both of which are impossible, hence both of the claims are speculative. This leaves us with an ambiguity that can be interpreted pessimistically, but also with an optimism to improve. And this is where I think Hill's second argument fails to some extent. Hill seems to argue that we must adjust our normative goals because our empirical reality falls short, when in fact this should inspire us to develop our own rationality and morality to get closer to that normative goal. Further, by distinguishing between what is our current empirical reality and what is our normative goal, we are not forced to compromise on the principles to the extent that Hill proposes. This also reflects how some laws can be rightful but morally inadequate; we have to start from somewhere, and a starting point for reform is to acknowledge the basic legitimacy of the state and the civil law.

This, however, raises the question of how we secure such a development. An intuitive response to the above would be that this normative goal and gradual development encourages nothing but passivity as humanity progresses regardless of what we do. A second objection could be that trusting the process is too slow, we want justice now and we have the means to achieve it, and one such means is civil disobedience or revolution (in this very general discussion I treat them as the same as they both involve breaking the law to achieve some political end). The latter would demonstrate the need for action and the fact that it is humanity itself, not *Geist*, that ensures this progress.

However, I believe both of these objections can be answered by looking at *An Answer to the Question: What is Enlightenment?,* which emphasises the individual's duty to develop their reason and the public's freedom of the pen. Kant distinguishes between one's public character and private character, insisting that the "public use of one's reason must always be free, and it

alone can bring about enlightenment among mankind" (E, 37). Although somewhat counterintuitively phrased, Kant then argues that this public use reflects one's scholarly character whereas a private use of reason will often be reflective of one's work and therefore somewhat restricted. For instance, a civil servant should not criticise or prevent executing orders while at his post as this would hinder achieving the public ends of the general will. However, as a scholar (i.e., outside of work) he is free, perhaps even obliged, to offer relevant criticism. With this distinction, Kant allows for our unsocial sociability to express itself in a regulated manner.

Further, protecting this right to public discourse is essential to improving the society and facilitating the aforementioned development of humanity as a whole. Given Kant's insistence on the role of reason in morality, all peoples must be granted the opportunity to develop their rationality both as individuals and as a people. Using an argument that should be familiar by now, Kant argues that imposing a guardianship on a people that restricts their rational development is impossible, and is from a contractual perspective "absolutely null and void, even if it should be ratified by the supreme power, by parliaments, and by the most solemn peace treaties" (E, 39). Doing so would prevent future generations to rid themselves of errors and "generally to increase [their] enlightenment," something that would be a "crime against human nature" (Ibid). Thus, no people could reasonably want restriction on their public freedom, and imposing such a restriction would be an illegitimate (null and void) move.

Kant anticipates criticisms from existing monarchs and argues that he will soon realise that there is "no danger to his legislation in allowing his subjects to use reason publicly and to set before the world their thoughts concerning better formulations of his laws, even if this involves frank criticism of legislation currently in effect" (E, 41). This point is twofold. On the one hand, it seeks to convince a monarch that would be afraid of losing his power as a result of public freedom that he too would benefit from this. On the other hand, Kant implicitly puts forward his point about the legitimacy of a given regime and its laws; that is, one can (and should) criticise the laws of a regime but that does not automatically translate into revolt. Following this, one could argue that Kant makes a strong case in favour reform, as emphasised by Flikschuch (2008), and that any rightful government must make allowance for even frank criticism of its existing laws. Additionally, one could argue that it is one's duty as a scholar (i.e., one's public identity) to provide such criticism for the betterment of society, thus making the overarching development of society an active undertaking.

Golob: The revolution as a sign of moral progress

While the interpretation in the previous section has some potential in how we understand Kantian progress, it does not directly answer the question of how revolution can be justified on a Kantian basis. I have rejected the approaches to this question from both Korsgaard and Hill; Korsgaard on the basis that a single person cannot possibly have access to the supra-moral standpoint required for her argument, and Hill on the basis of the neglected potential of *Universal History* which follows from his arguments. This section examines Golob's Enthusiasm model as a potential justification-

Golob argues that the most common interpretation of Kant's prohibition of revolution combined with his personal enthusiasm for the French revolution is the "Cunning of Reason" model (COR). This claims that "acts which are morally problematic will nevertheless contribute to long-term moral progress" (Golob 2021, p. 979). Whilst this approach has both merits and textual support, it can lead us to either Hegelian or consequentialist inferences which could go against the basic Kantian framework. What is crucial about Golob's insight, however, is that the COR model does not adequately demonstrate what is distinctive about revolution in particular, as opposed to e.g., war or other forms of discord. Moreover, Kant's position on war varies enough between texts to call into question its consistency. And, finally, it is not clear how the COR model found in e.g., *Perpetual Peace* explains the *moral* development he argues for in *Conflicts of the Faculties*, as opposed to a merely legal one (Golob 2021, p. 981). Hence, it is possible that Kant has a consistent argument with the COR when it comes to explaining legal and political development, however, it is then unclear how we can adequately explain the following passage: "It is simply the spectator's mind-set which … (due to its universality) demonstrates a character of the human race as a whole and also (due to its unselfishness) a *moral character* of the latter" (*Conflict of Faculties* 7:87, as found in Golob 2021, p. 978). This suggests that not only is there something special about revolution as an event following discordance, revolution is also unique in that it can bring about moral progress based on the spectator's reaction, neither of which is captured by the COR model alone.

In response, Golob proposes what he calls the "Enthusiasm model" which can supplement COR (Golob 2021). Following definitions from the *Third Critique* and *Conflict*, enthusiasm for Kant is the feeling of affect paired with the idea or sympathy of the good (Ibid, p. 982). Kant considers affect a particularly strong emotion which makes the mind "unable to engage in free deliberation about principles with the aim of determining itself according to them" (*Critique of Judgement* 5:272, from Golob 2021, p. 982). An example that springs to mind is someone who

is blinded by rage and thus simply acts based on this feeling without thinking. Such affect paired with the idea of the good results in an intense commitment to this idea that can be observed in the actions of revolutionaries (Golob 2021). Further, this Enthusiasm is able to move the spectators of these revolutionaries to such an extent that they, too, experience a feeling that "borders on enthusiasm" (Golob 2021, p. 983). This reaction is both universal and unselfish something Golob explains clearly:

> "It is unselfish because the spectators are not themselves suffering the injustices of the *ancien régime*: they are safely housed in Germany or England or elsewhere. It is universal not because everyone exhibits it, something Kant knew was not the case, but because its presence points to a common moral capacity: the fact that those with nothing to gain and a great deal to lose show such intense fellow-feeling is a 'sign' of their 'innate' commitment to the ideal of the good" (Ibid).

Hence, the fact that people other than the French, the spectators, were so moved by the revolution suggests that there is not only a capacity of moral feeling, but one that allows us to set aside our personal risks and differences to side with the idea of the good. Golob then argues how this approach differs from COR.

First, he argues, this account is not neutral on the motives of both the revolutionaries and the spectators, as it requires them to be in a state of enthusiasm, thus separating Golob's account from the more passive COR. Second, the spectators are not in an epistemically superior position or have special insight into history, "[a]ll that is required is their spontaneous emotional reaction" (Golob 2021, p. 983). Third, this approach accounts for the supposed moral development, not merely a legal development. Fourth, it is not the consequences of the revolution as such that are significant, but rather "its motives and the spectators' *contemporary* reactions to them" (Ibid, my emphasis). Consequently, it is not because we today consider some benefits of the French revolution that we may speak of progress, but because of the enthusiasm found in the contemporary agents, thus evading consequentialist justifications. Fifth, there is a tripartite structure required to ensure that a revolution is, in fact, a sign of moral progress: "There are the actors, spectators, and finally the philosopher observing the reaction of those spectators: it is that reaction which is the 'sign'" (Ibid, p. 984). This final evaluation by the philosopher ensures that Kant makes a theoretical claim about progress: "it is effectively an inference to the best explanation, by the philosopher, seeking to explain the behaviour of the spectators" (Ibid). Golob argues that enthusiasm essentially provides a loophole to our judgement of the revolution: "the revolutionaries are exempt from moral condemnation insofar

as they operate under the influence of an affect and one deeply linked to morality. And Kant's epistemic framing of the issue provides another. His position is not that the uprising itself was acceptable; instead, the target of his approval is the spectators and the reason for his approval is what they teach him about human psychology." (Golob 2021, p. 984).

The concern here, which Golob subsequently addresses, is that this loophole is a very powerful exemption to grant even to spectators of the Revolution, yet we currently have no resources to determine whether the revolutionaries truly acted from the motives warranting such a reaction or whether they were false prophets, so to speak. A possible solution is to look at Kant's subsequent turn to pedagogy in *Conflict*, which upon first reading strikes one as unrelated and a different topic altogether. However, Golob argues that the shift to the schoolroom demonstrates the pedagogical role of enthusiasm (Golob 2021). The details of this is not relevant for the purposes of this thesis, but the key claims in relation to revolution are the epistemic position generated by the revolutionaries, the spectators, and the philosopher's analysis of this interaction, particularly of those spectators that did not have any personal gain from supporting the revolution. With this interaction and tripartite structure, we gain a lasting epistemic insight which constitutes a sign of moral progress, because our capacity for morality is revealed in the world.

Golob's aim is limited to providing an additional account of the role of revolution both epistemically and pedagogically, and so ends with further observations. Kant's aim is to provide a pedagogical argument which rests on the effect of the revolt on the viewer and the theorist. Due to the intense emotional affect of the revolutionaries themselves, they cannot be morally commended, but also not condemned since they are not acting on their rational capacity. Golob argues that this provides a loophole in relation to revolution, however, he also writes that such a position also highlights the potential limits of Kant's transcendental idealism and its epistemic restrictions, such as our inability to be sure that there has ever been an act of a truly good will.

Similarly, even with the tripartite structure Golob offers, we still do not know for certain that the revolutionaries themselves were moved by the idea of the good. Given their exemption from moral judgement, one could argue that this is not necessary. However, as Golob notes, this also reveals Nietzschean critiques about human motives, which suggest that any noble or good motive is merely human impulses in disguise. This becomes particularly worrying given the scale of revolution, if we cannot determine with certainty the motivations of one, how could we possibly determine with certainty the motivations of so many?

Despite this, Golob's interpretation is convincing as he improves upon several historical problems and coherently justifies the existence and use of the moral loophole for the revolutionaries. This should, in other words, be a particularly convincing account when it comes to civil disobedience. If there is a moral loopholes for the revolutionaries, surely there must be one for civil disobedient acts which only seeks to break laws in order to improve the existing civil order, particularly given that civil disobedience is by definition non-violent. If we apply this interpretation to civil disobedience, we can argue that the agents of civil disobedience are in a state of enthusiasm, motivated by the idea of the good, which is why they end up breaking laws as they get carried away by their affect. Because of this, they can be rightfully punished by the civil law for breaking those laws (hence, there is no legal right to civil disobedience), but they gain an exemption, morally speaking, for doing so. In other words, we have a clear moral justification for civil disobedience.

However, there are still possible objections to this approach. First, as mentioned above, we do not know for certain that revolutionaries truly act from the idea of the good and we thus risk endorsing false prophets. This is significant as not all revolutions have been commended by history as "not everyone favours the same revolutionary" (Flikschuh 2008, p. 142). In response, one could argue that these events were not truly revolutions for that very reason, but assuming the reactions were the same, we quickly end up with a Gettier case of what makes a true revolutionary and a true display of enthusiasm. For while e.g., the storming of the US Congress was widely condemned, they were supported by people in the US who did not themselves participate. Further, there is no criterion in the Enthusiasm model that necessitates every existing human being to experience the enthusiasm of a spectator in order for it to qualify as such; it is merely the universal capacity that the spectators supposedly unveil. One could make a similar argument for the French revolution; the enthusiasm of the spectators was revealed in Europe, which was enough to make it 'true', regardless of whether that same enthusiasm (or any reaction, for that matter) was found in e.g., Asia. This issue is exacerbated by the subsequent moral exemption, as we cannot pass any moral judgement on those revolutionaries because they acted in a state of enthusiasm, moved by the idea of the good. Without a clear way of distinguishing true revolutionaries from those that merely appear true, we risk falling into this vicious circle.

Second, we must return to Maliks' interpretation of the French revolution. He distinguishes between the 1789 and the 1792 revolution, in which the former was a procedural transfer of power and the latter a violent rebellion. Maliks argues that Kant only endorsed the first

revolution of 1789, not the second (Maliks 2022). If we accept this distinction, it becomes even less clear which revolution would qualify as the sign of moral progress. Perhaps, then, the enthusiasm we saw was in fact directed at the first, peaceful, revolution rather than the violent one. However, if the first revolution was in fact a legal reform, which did not destroy the civil condition but rather changed its governance, it is difficult to see how it could be conceived as a revolution and not merely reform. Consequently, it must have been the second revolution of 1792 that was the "true" revolution which gave rise to the enthusiasm and supposed moral progress.

However, one could conceivably argue that the first revolution was still a substantial enough change to qualify as a revolution, given that Kant's definition relies on the use of rebellion instead. If this is the case, it seems plausible that the enthusiasm should have been directed at the 1789 revolution which brought forth significant legal reform which would more appropriately represent the people of France. But the enthusiasm And so, it is possible that the enthusiasm Kant speaks of was in fact misdirected to the 1792 revolution. Thus, it seems that we could give an argument for both sides individually. Given that the two events are interlinked, this is not enough to call into question the enthusiasm model as a whole, however, it further support my argument for the difficulty of discerning which revolutions are "true" and which merely appear that way.

Finally, while the solution of a moral exemption is presented coherently, it does not give us the tools we need to answer the question of whether a revolution is ever right. It improves upon Korsgaard's initial position by not claiming that the agents have some supra-moral insight they act on, but rather that they cannot be judged by their actions due to their affect. However, an exemption is merely suspending judgement due to lack of knowledge. Perhaps the state of enthusiasm by the revolutionaries was forged, even if the spectators' reaction was real; if so, how can the philosopher tell what is true if one or more of the tripartite structure could be false? While falling back on scepticism provides a theoretical loophole, it remains unsatisfactory from a moral point of view as we consequently do not have the resources to morally condemn those who either knowingly or not are not "true" revolutionaries. In other words, we must either give up on a firm epistemic ground for making such judgements, or give up some of our ability to condemn wrongful actions on a grand scale. I believe the gravity of a revolution should make us want to err on the side of caution by not giving up our ability to morally condemn wrongful acts, even if they should qualify for an exemption based on enthusiastic affect.

Civil disobedience is a much less significant act compared to revolution, and so it could potentially escape this critique. Firstly, the intentions of the civilly disobedient is generally to improve the existing state by changing some laws they believe are fundamentally unjust. In order to see that change, they believe they are justified in breaking some laws to make way for better ones as long as this is done through non-violent means. Prima facie, it might seem more likely that the civilly disobedient are truly motivated by the idea of the good but that they get carried away by their enthusiasm which is why they end up breaking some laws for their cause.

To elaborate, we can try describing the case of Rosa Parks under this model. As part of the civil rights movement, Parks was motivated by an idea of the good, namely, equal rights. Given the unequal rights, Parks believed that any legal action to change this would not be enough, as the rights and opinions of Black Americans were already dismissed by the law. Hence, it seemed that legal routes would be insufficient. Due to her conviction, Parks then get somewhat carried away and ultimately breaks the law, resulting in civil disobedience. The enthusiasm of the spectators are shown through both the immediate and subsequent support for the cause. First, the organised boycott of the bus company that enforced the apartheid law, drawing support from the local community and more Black Americans; and second, the change in the local law permitting such discrimination. One could argue that the philosopher's observation could be akin to the international recognition that discrimination on the basis of ethnicity is both morally wrong and unlawful. The reason why I would not describe the latter as the spectators' reaction is due to the fact that racial discrimination remained prevalent across the world afterwards, most notably, perhaps, in South Africa.

However, the immediate reaction of support, not just from the people suffering from the discrimination, but also from those who did not (as illustrated by e.g., Harper Lee), could be a case of enthusiasm revealing a universal moral capacity for recognising the equality of all human beings. The fact that white people in the South were willing to support the cause, often with the risk of losing their jobs, financial insecurity, and social alienation, shows that Parks' civilly disobedient act was able to generate the necessary enthusiasm from spectators who were not directly affected by the issues. That is not to say that the risks of white people supporting the civil rights movement were in any way comparable to the suffering of Black Americans, but rather that their support was risky enough to warrant my claim that this was a genuine case of enthusiasm from the spectators. This immediate reaction thus signified an epistemic shift, which we benefit from today by the largely universal recognition that racial discrimination is fundamentally wrong, thus constituting a sign of moral progress. The fact that this view has

stood the test of time also shows that the enthusiasm expressed was not merely a case of "*Schwärmerei*", which Golob describes as "an over-wrought fanaticism" (Golob 2021, p. 985). Hence, the case of Rosa Parks seems to be fully justified as an act of civil disobedience following the Enthusiasm model.

As it stands, we now have a convincing case for a moral justification for civil disobedience. The qualms that we had with revolution, namely its violent and rebellious nature, are avoided in civil disobedience by definition; it is non-violent and does not aim at destroying the state, only improving it, a process which sometimes requires breaking laws. We also have a convincing case with regards to human rational development over history, in which humans as a species gradually develop over time by collectively developing our rational capacity. Additionally, there is less risk of false prophets with civil disobedience, as the legal order and state remains, and so if the act was not one of true civil disobedience, it would merely be a criminal act. The true civilly disobedient act then showcases this epistemic and moral development of our rational capacity which takes a significant step forward in response to the act. Of course, the civilly disobedient act would remain unlawful, however, it would have a strong moral justification which avoids the pitfalls of other accounts I have discussed.

Is this approach sufficient when it comes to civil disobedience?

Due to the development of human rational faculties, it is possible that we do not fully know what is right. This is evident in human history, with e.g., slavery being seen as acceptable from ancient times, but now we know that it is morally abhorrent. The moral law has not changed regarding this, enslaving human beings has always been a clear violation of their autonomy and freedom, but the civil law has taken a lot of time to catch up with this, due to our lack of rational understanding. Hence, in order to make sure that the civil law can in fact catch up with, or reveal the full extent of, the moral law civil disobedience may be a necessary act to ensure this rational progress is enforced by the civil law.

Let us suppose, for instance, that Rosa Parks and the members of what would become the civil rights movement were individuals whose rational faculties were already developed beyond the species average. They had consequently understood that such laws could not be permitted by the moral law and thus sought to change the civil law. Yet due to their standing in society, such legal attempts did not give rise to the necessary change. Hence, they resorted to civil disobedience to ensure that their epistemic insight would be made clear and enforceable by the

civil law. Under this interpretation, the goal was not to spread this moral knowledge, but to ensure the legal rights for everyone so that moral knowledge could come from it afterwards, making the goal political.? In short, moral knowledge does not ensure that it is reflected in the civil law. Kant makes it clear that this should be the goal in works such *as Universal History, Theory and Practice*, and *To Perpetual Peace*, but it is not a *necessary* criteria for a state to be legitimate; it merely reflects the sovereign's obligation to aim for conformity with the original contract. With continuous political development, human beings as a species will in all likelihood continue to progress towards the moral ideal, but it is only the civil law that can ensure such insight is enforced, and so civil disobedience must be directed at political change.

The moral ideal is our goal, however, getting there required political action with political justifications. Consequently, it is more suitable to look for a justification of civil disobedience within the political and legal sphere as opposed to the moral one. The moral approach shows us that acts of civil disobedience (and revolution) can be part of human progress towards a moral end, but it does not show us what is significant about a single act of civil disobedience in that time. One could argue that a legitimate republican state will inevitably develop in this direction without the interference of civil disobedience, however, this would undoubtedly take a long time, even assuming the affected people would get heard through e.g. legal protests, open questions etc. Hence, civil disobedience is a political strategy and device aiming to bring forth change at a much more rapid pace than any belief in nature's providence can give. Civil disobedience is thus very much an act within the political and legal sphere, and would thus be better justified by an argument within that same sphere. Further, we as human beings do not know that a morally perfect world will ever exist, nor if we will live to see it; but what we do know is that our present political circumstances are imperfect and should be changed for the better, which is why some choose civil disobedience as a strategy to further improve the potential uptake of their epistemic insight.

The political justification for revolution

I have now attempted to firmly establish civil disobedience as a political device within the political sphere, thus warranting a potential justification from the same sphere. Now onto what that might look like. In short, by following this approach, I do not think we have room for a Kantian story of justified civil disobedience, a counterintuitive and somewhat radical claim. I will explain how this does *not* result in a view endorsing oppression and discrimination contrary to Kantian moral law and human decency by appealing to the distinction between passive and active citizens and the concept of barbarism.

Kant's idea of citizenship and the duties of the state: Active and passive citizens

Active citizens are those that contribute to making the law while also enjoying its benefits, whereas a passive citizen enjoys the protection of the law has no right nor obligation to contribute to making it. A contemporary parallel could be that of those who have "Settled" status in the UK, which grants them certain right but not a right to e.g. vote, and citizens that enjoy such a right. While this distinction has sometimes been dismissed by scholars due to Kant's immediate disenfranchisement of women, Weinrib (2008) presents a clear starting point for a re-evaluation of this aspect, which is gaining more traction amongst Kantian scholars. The recurring argument here is that Kant's core argument does not *a priori* discriminate against women, but that this is rather an addendum stemming from the views of Kant's time. It is of course possible that Kant did in fact hold this view, however, that does not mean that this view necessarily follows from his arguments. The goal of this section is to see whether our examples of civil disobedience from the first chapter would be that of active or passive citizens. If they are active citizens, one could argue that their obligation to comply with the civil law is greater than that of passive citizens, because they are part of the omnilateral will which authorises the civil law. If, on the other hand, the civilly disobedient are passive citizens, it could be the case that the civilly disobedient act is merely an expression of their desire to become active citizens, and that they are not bound by the civil law to the same extent as their active counterparts thus making the passive citizen's approach justified.

Both *Theory and Practice* and the *Doctrine of Right* discuss citizenship. Weinrib argues that it is the discussion in the former that often proclaims these discriminatory conceptions, and that the latter is what best represents Kant's mature view. Based on a variety of criticisms of Kant's

account, Weinrib sets out two intuitive criteria for active citizenship. The conventional criterion that "full citizenship must not be withheld or awarded on the basis of economic status", and the natural criterion that "full citizenship must be attainable for all persons and cannot be denied on contingent factors such as gender" (Weinrib 2008, p. 4). Weinrib's interpretation thus aims to show that Kant does indeed fulfil both criteria.

If we follow the "Doctrine of Right" in *The Metaphysics of Morals*, Kant clearly demonstrates that economic status is not the sole criterion of citizenship. For instance, Kant argues for government support of the poor through taxation (MM 6:326). This clearly goes back to his fundamental principle of freedom, as the poor are left to the individual will of others since they have no means of their own to be independent. Hence, they are in a perpetual state of dependency, which the state ought to correct. When such a policy is enforced, the poor are independent, and gain a civil personality which qualify them as citizens (Weinrib 2008). Consequently, "the poor have the status of active citizens for the same reason as civil servants: all independent persons are dependent on the general will of the state. Dependency on the general will is consistent with contributing to the general will, which requires that one is not dependent on the choice of a private will" (Weinrib 2008, p. 13). Likewise, a rich person is not necessarily a citizen, as they could be economically dependent on another individual. The empirical fact that this was often the case for women seems to be a more logical reason for their disenfranchisement as opposed to some natural defect in their faculties. Kant's juridical conception is minimal, and does not exclude women, hence we can also fulfil the stipulated natural criterion.

If we accept this, then Kant's conception of active citizens does not automatically disenfranchise women and the poor. However, this is not enough to get away from the empirical fact that women in particular were passive citizens in Kant's time, which is far from ideal. It is important to note that some instances of dependence are rightful, such as a child's dependence on their parents or guardian. In those cases, the parent has a duty to act in the child's best interests and develop them so that they can become active citizens as adults. Thus, there seems to be a minimum criterion of adulthood, and perhaps also of some mental capacity, to be considered an active citizen. This may seem at odds with the stipulated natural criterion, however, it is worth noting that "full citizenship must be *attainable* for all persons" does not entail that all persons must be in possession of it at all times. That a child has the capacity to attain full citizenship, but does not currently qualify is not inconsistent with the natural criterion. Some might object to this kind of reasoning on general grounds, which is a different discussion,

but this is obviously a very Kantian way of reasoning, as it parallels Kant's arguments for e.g., who counts as a moral agent. Regardless, common sense dictates that there is no need for children to have a right to vote (i.e. be active citizens) but that they should be raised with the expectation of becoming active citizens and be educated accordingly.

How can a passive citizen become an active citizen? Weinrib argues for a, perhaps somewhat controversial, Kantian solution. Rather than merely granting a right to vote to passive citizen, Weinrib argues that the Kantian solution must be to fix the underlying issue of dependence.[4] That is, a symbolic right to vote means little if one is still dependent on the goodwill of another individual. Instead, we must aim to create a state of universal independence through institutions rather than merely allowing the dependent to vote: "The state must establish the conditions in which passive citizens can shed their dependency and become active citizens" (Weinrib 2008, p. 21). When remaining passive citizens, they are essentially forced to live under a rightful condition they cannot participate in, and this intermediary state (not the state of nature nor part of the omnilateral will) is something that ought to be corrected. This would bring us closer to the ideal that Kant puts forward: rather than merely satisfying the postulate of public right, the state has an obligation to approximate the idea of the original contract.

Do passive citizens have any rights to resistance? That is, if the active citizens do not aim to lift the passive citizens up to their status, do the passive citizens have any means of retaliation? It seems that while the passive citizen is forced to live under a rightful condition they cannot participate in, they are subject to coercion and punishment like any other citizen, but without the right to be part of the lawmaking body. Barring the exceptions, this generates unrightful dependence for the passive citizens. Clearly, there should be legal and institutional means to correct this, but it is not obvious what happens if this is not the case. I believe this results in one of two possibilities. The first is that the state ensures that all passive citizens have the opportunity to become active citizens, but still allows the existence of passive citizens. This allows those who are visiting, working there for a limited period etc., to still be bound by the coercive rights of the state without the need to consider them part of the omnilateral will. If, however, they want to be part of the legislative authority, they must be given the opportunity to do so.

Such a distinction would remain rightful as the state would still be under an obligation to correct unrightful dependence, but allows for widening the scope of rightful dependence to

[4] For a detailed discussion on rightful and unrightful dependence, see Moran 2021.

accommodate those who do not wish to enjoy a full set of rights in what is not their national state, but agree to still be bound by this state's laws in exchange for a limited set of rights to e.g., sell their goods, effectively amounting to a contractual obligation between the private citizen and the state. As long as they remain passive, I do not see why these citizens should have any justification for civil disobedience: passive citizens are not part of the omnilateral will and consequently can have no say in whether it is right or wrong. Consequently, they are not part of the civil union and thus cannot be civilly disobedient. An act of presumed civil disobedience based on a moral justification could be construed by the state as a declaration of war as it essentially takes the form of an outsider invading the legal space of a sovereign state. Hence, as long as the opportunity to become an active citizen is there, this conception of passive citizenship does not violate the postulate of public right and would not have a justification for civil disobedience.

The second possibility is that the passive citizens are in fact unrightfully disenfranchised and do not have the opportunity to become active citizens. If this is the case, the state would violate the postulate of public right and could therefore not be considered a legitimate state. It is important to establish why a group is disenfranchised and whether this infringes on their freedom as a human being. After all, we do not want to say that a state that disenfranchises children is not legitimate. If, then, the state actively keep a group of people as passive citizens when they should in fact be active (or at least have the option of becoming an active citizen), we are no longer in a rightful condition but rather one of barbarism. Such a condition might not only affect passive citizens but could also violate the rights of the active citizens. The next section is devoted to discussing the concept of barbarism using the distinction between active and passive citizens as basis.

The condition known as barbarism

The concept of barbarism can be found in Kant's *Anthropology*. It is the condition of force without freedom and without law, as any semblance of law violates the postulate of public right (recall definition): "As a defective form of a state of nature, a condition of barbarism can have no united will. All force is merely unilateral; when force is organized in a condition of barbarism, it take the form of rule by prerogative, even if members of the more powerful group have elaborate procedures for decision-making" (Ripstein 2009, p. 341). This means that all exercises of power are necessarily for private purposes as there is no public end; the fact that it

is organized does not amount to it being public: a "highly organized mob is still a mob" (Ibid, p. 342). When looking at this with regards to revolution, there is no argument for preserving the general will and the rightful condition, as this does not exist. Hence, a revolutionary act would not nullify the rightful condition and bring us back into the state of nature, as we already exist in a defective state of nature under barbarism. The people of power in a state of barbarism cannot claim to represent the people and thus act unilaterally with private agreements. When the state cannot claim to be acting from a united will, it thus violates the criterion for what makes a rightful condition and cannot be a legitimate state.

Ripstein uses this framework to argue that Nazi Germany, which is often the prime counterexample of Kant's claim to the rightful authority of the state, was not in fact a rightful condition but rather a state of barbarism. This provides a crucial step in a potential resolution of the problem of civil disobedience. If we were to assume that Nazi Germany was a legitimate state, Kant's position would be that the laws they decreed must be resisted. This would lead to the abhorrent position in which we had a duty to comply with the laws that so obviously and intuitively contradicted any concept of human value. If this was indeed a counterexample, Kant's position would be indefensible. However, if we accept Ripstein's argument that this system violated not only the original contract but also the postulate of public right. A barbaric state is hence not a legitimate state as it does not provide the institutions for a rightful condition. This allows us to distinguish between states that are legitimate but imperfect and those who claim legitimacy on a unilateral, and therefore false, ground.

Barbarism could on my interpretation exist both if the rights of active citizens are violated or if the rights of passive citizens are violated. For the latter case, it would be a state of barbarism if the opportunity of becoming active citizens are precluded. The former would be a bit more complex. If we accept Weinrib's argument that it does not follow from Kant's claim that women should be disenfranchised, then women are rightfully entitled to active citizenship which includes the right to vote. A state that violates this thus conflicts with right because the state could not exercise an omnilateral will with the exclusion of a certain group of people which by right should be entitled to vote. Similarly, all cases of civil disobedience that we covered in the first chapter would fall under this interpretation where a group of people are denied the ability to exercise their innate right to freedom. Intuitively, this should give us a political justification for civil disobedience in which those who are unrightfully disenfranchised must resort to extra-legal means to put forward their claims.

However, a more sound approach would be to say that a state of barbarism cannot accommodate any concepts related to a rightful condition because its very existence is dependent on the violation of right. That is, conceptually, the state delegitimises itself when it violates the innate right to freedom by disenfranchising those who should be active citizens. Thus, a state of barbarism is not a legitimate state. Both revolution and civil disobedience requires a civil or rightful condition to be conceptually sound and barbarism is a violation of this criterion.

Civil disobedience and barbarism

As argued in the previous section, the states which violate the postulate of public right which could justify civil disobedience cannot be considered legitimate states. As a result, the Kantian framework does not have room for a concept of civil disobedience. However, there is still a strong case for actions that *look* like civil disobedience. My argument is that our examples of civil disobedience merely appeared like cases of civil disobedience as they were enacted in a state of barbarism. This raises the question of how we conceptually makes sense of the actions that look like civil disobedience under the premise that there is no room for civil disobedience in the Kantian framework.

One such approach could be to argue for a provisional civil disobedience. This would be similar to provisional acquired rights in the state of nature, the difference being that a provisional right to civil disobedience cannot be made conclusive. The function of such a proviso would be to give conditions under which a state appears to be legitimate but is not due to its violation of the postulate of public right. A general formulation which would encapsulate our examples could be that "any state which violates the innate right to freedom for human beings is a condition of barbarism and must be brought into a rightful condition". The criterion is not arbitrary as it is the fundamental aspect of Kant's political philosophy, and if the reciprocal coercion of the civil condition is not authorised through the correct means, a state cannot be legitimate because it cannot claim to represent the people. Further, this evaluation can be combined with a call to action, that is, the duty to create a rightful condition by uniting with others. However, as Ripstein remarks, in a state of barbarism there may not be a people as such to unite with due to the excessive use of force and lack of freedom. In response, he argues that "[h]uman beings in such a situation are not required to engage in futile self-sacrifice to create a rightful condition, but they are permitted to use force to create one" (Ripstein 2009, p. 343). That is, the barbaric

attempt to keep people in a state of barbarism can be resisted with right because such a condition violates the innate right to freedom.

One might argue that a provisional civil disobedience would indirectly comply with the moral law, and thus have a moral justification as the violations under barbarism would violate the concept of human dignity. I do not aim to preclude this option, however, in response I would reiterate the fact that the goal of the nature of civil disobedience remains political, whether that is explained by a desire to change existing laws on the assumption that the state is legitimate, or whether it is explained by the need to create a rightful condition. Moreover, in order to approximate the original contract and achieve a truly moral constitution, we must necessarily begin from the existence of a legitimate constitution; a Kingdom of Ends cannot exist in the state of nature and is thus dependent on a rightful condition. In other words, we need the coercive law if we are to have any hope that a moral law can be put into practice.

This view is also compatible with Kant's idea of providence in *Universal History*, in which nature seems to have some secret plan to gradually ensure that humans as a species will progress to their full rational and moral capacities. As Wit argues, such a claim cannot be empirical, neither about the potential moral progress throughout history nor about the future prospects of continuing such development (Wit 1999). Instead, he argues, this idea represents a moral concept of progress in which we have a duty to believe (Ibid). This belief is put into practice through the duty of publicity: "The public use of one's reason must always be free, and it alone can bring about enlightenment among mankind" (*What is Enlightenment?* §37). By "public use" Kant refers to our rational capacity as individuals, encouraging us not only to think for ourselves but to actively voice our critiques so that the world may be improved. This allows Kant to avoid critiques of quietism in addition to distinguishing himself from Hegelian conceptions of history: "we *cannot* leave progress to nature because, first, there is no such concept of empirical progress in Kant's philosophy and, secondly, we can only hope for natural progress *if* public opinions do get expressed" (Wit 1999, p. 304). Thus, any hope for a just and moral society requires hope that it can be realised paired with the knowledge that such an option can only possibly be realised if we use our rational agency to express ideas and critiques. This approach allows the "sole palladium of the people's rights" to not be restrictive, but in fact generative of public right. Consequently, cases in which the state does not violate right so as to plunge itself into barbarism, would be covered by the freedom of the pen, as the public would not only be permitted to voice their critiques, but it would be a civic duty for them to do so. Hence, the civil aspect of civil disobedience would be covered by the freedom of the pen, and

the disobedient aspect by the postulate of public right demanding the creation of a rightful condition.

Conclusion

The aim of this thesis has been to answer the question of whether there could be a Kantian theory of civil disobedience. After introducing preliminary definitions and exegetical interpretation of Kant's political philosophy, I began examining the potential for a moral justification for civil disobedience. I considered three different approaches that fell under this scope. I rejected the first attempt due to its reliance on a supra-moral standpoint which would supposedly warrant imposing a unilateral will on a people, arguing that this approach fails as such a perspective is not available under a Kantian framework. I then considered Hill's more practical, yet distinctly moral approach of allowing some lenience in principle. This I rejected on the basis of drawing conclusions about the validity of a principle based on people's inability to follow it; the fact that a stringent moral principle is difficult to follow does not invalidate it nor makes it impossible to comply with. Finally, I presented the Enthusiasm model, which *prima facie* provided a convincing account in favour such a justification. However, upon further inspection, I presented a set of objections, and argued that being exempt from moral judgement is not the same as giving a positive moral justification. While endorsing such an exemption is plausible from a Kantian view, I argued that a political approach would be better suited to answer the question directly. My justification for this was the fact that civil disobedience is inherently a political action with a political aim, and should consequently have a political solution. Additionally, the potential for realising a moral society rests on the assumption of a legitimate civil condition; the Kingdom of Ends cannot be realised in the state of nature and thus our priority must be to enter a rightful condition before aiming to approximate the original contract.

I supported this argument by drawing on Kant's distinction between active and passive citizens, and the condition of barbarism. Here, I argued that the conditions under which we intuitively would consider civil disobedience justifiable are in fact conditions of barbarism that violates the postulate of public right. Following this, I demonstrated how barbarism generates a duty to enter a rightful condition which is conceptually akin to civil disobedience. However, since these actions do not occur in the civil or rightful condition, we can no longer speak of it as being

neither civil (as it does not conform to any idea of civil society) nor disobedient (barbarism is a defective form of the state of nature in which there is no authority one could disobey). Instead, the actions which we might call civil disobedience under different framework as simply actions that contribute to creating a rightful condition under a Kantian framework. Finally, I outlined the potential for freedom of the pen, arguing that the citizens in a rightful condition have a duty to actively critique and improve their state so to ensure at least the possibility of progress in history, thus accounting for the reformist aspects found in our intuitive conception of civil disobedience as well. Following this, I concluded that there cannot be a Kantian theory of civil disobedience as such, because the essential nature of the concept can be adequately explained with existing Kantian theory without giving rise to inconsistencies.

Bibliography

Flikschuh, Katrin "Sidestepping Morality: Korsgaard on Kant's no-right to Revolution" *Annual Review of Law and Ethics*, Vol. 16, 2008, pp. 127-145.

Golob, Sacha "Kant on Revolution as a Sign of Moral Progress" *Journal of Philosophy of Education,* Vol. 55 No. 6, 2021, pp. 977-989.

Hill, Thomas E. "A Kantian Perspective on Political Violence." *The Journal of Ethics,* Vol. 1, No. 2, 1997, pp. 105–40. *JSTOR*, http://www.jstor.org/stable/25115541 Accessed 7 Apr. 2023.

Hill, Thomas E. "Making Exceptions Without Abandoning the Principle: or How a Kantian Might Think About Terrorism" in Frey, R. G., and Christopher W. Morris. *Violence, Terrorism, and Justice*, Cambridge University Press EBooks, Aug. 1991. Cambridge University Press, pp. 196-229, https://doi.org/10.1017/cbo9780511625039.002.

Kant, Immanuel *Groundwork of the Metaphysics of Morals*, Mary Gregor and Jens Timmermann (Eds.), Cambridge University Press 2012, Reprinted 2018.

Kant, Immanuel *Perpetual Peace and Other Essays,* Trans. Ted Humphrey, Hackett Publishing 1983.

Kant, Immanuel *The Metaphysics of Morals,* Ed. Lara Denis, Trans. Mary Gregor, Cambridge University Press 2017.

Korsgaard, Christine M. "Taking the Law into Our Own Hands: Kant on the Right to Revolution" *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*, Oxford Scholarship Online January 2009, pp. 234-261.

Lee, Wendy Lynne *This is Environmental Ethics: An Introduction (This is Philosophy)*, Wiley-Blackwell 2022.

Maliks, Reidar *Kant and the French Revolution (Elements in the Philosophy of Immanuel Kant)*, Cambridge University Press 2022.

Moran, Kate "Kant on Traveling Blacksmiths and Passive Citizenship", *Kant-Studien,* Vol. 112, No. 1, De Gruyter 2021, pp. 105-126.

Rawls, John. *A Theory of Justice : Original Edition*, Harvard University Press, 1971. *ProQuest Ebook Central*, https://ebookcentral.proquest.com/lib/kcl/detail.action?docID=3300425.

Ripstein, Arthur. *Force and Freedom : Kant's Legal and Political Philosophy*, Harvard University Press, 2009. *ProQuest Ebook Central*, https://ebookcentral.proquest.com/lib/kcl/detail.action?docID=3300841.

Weinrib, Jacob "Kant on Citizenship and Universal Independence", *Australian Journal of Legal Philosophy,* Vol. 30, No. 1, 2008, pp. 1-25.

Weinrib, Jacob "Sovereignty as a Right and as a Duty: Kant's Theory of the State" *Sovereignty and the New Executive Authority*, Claire Finkelstein and Michael Skerker (Eds.), Oxford University Press 2018, pp. 21-46.

Wit, Ernst-Jan C.. "Kant and the Limits of Civil Obedience" , *Kant-Studien,* Vol. 90, No. 3, 1999, pp. 285-305. https://doi.org/10.1515/kant.1999.90.3.285