Is there a transcendental deduction on Kant’s *Groundwork* III?

The purpose of the third section of the *Groundwork* is to justify the categorical imperative, that is, to prove its possibility. In the end of the second section, Kant establishes that a proof like this is necessary to show that morality is “something” and “not a chimerical idea without any truth” or a “phantom” [G4: 445]. Since the categorical imperative was established as a synthetic *a priori* practical proposition, to prove its possibility it is necessary “to go beyond cognition of objects to a critique of the subject, that is, of pure practical reason” [G4:440]. To this kind of proof Kant gives the name *deduction*.

In the present paper I shall (1) show the argument whose purpose is to prove the possibility of the categorical imperative; (2) show that the argument is a transcendental deduction; (3) present the argument as it is reconstructed by Allison², and (4) show that, although it seems compelling, the position of the commentator is wrong.

I

The notion of a deduction plays a central role in Kant’s critical project. Nowadays, it points to a meaning quite familiar to us: it refers to the logical procedure by means of which a conclusion is established through the relationship between some premises. Kant was familiar with this logical usage of the notion ‘deduction’, but as Henrich³ remarks, it wasn’t the only nor the most common usage in the academic language in the 18th century.

During the 18th century, ‘deduction’ was a notion used by jurists to refer to the written claims exposed to the Court in legal proceedings. This meaning was established since the 14th century. According to the rules of the Holy Roman Empire, some controversial legal claims should be submitted to the Imperial Court and both parties should justify their claims in a written way. Those written documents, whose aim was to justify the controversial legal claims, were called ‘deduction writings’.

Now, considering the argumentative structure of a deduction, one of its peculiar characteristic is that it must refer to an origin. Since the aim of a juridical deduction was to justify the legitimacy of a possession or a usage, that is, the legitimacy concerning an acquired right, it was necessary to explain how this possession or usage came into being. With this, it could be possible to decide who between both parties in the controversial juridical claim was right. The origin of an acquired right should be find in a fact, which must exist before the right in question came into being. That’s why the argumentation

---

¹ The first version of this paper I wrote with a scholarship from CAPES/Brazil during a research period in New York under supervision of Béatrice Longuenesse, to whom I’m thankful for her extremely helpful criticisms. The final version I wrote with a scholarship from CNPQ/Brazil.

² The paper we will analyze is entitled ‘The deduction in *Groundwork III*’. It is the chapter 12 of Allison’s *Kant’s theory of freedom*. Cambridge, Cambridge University Press, 1995.

³ Henrich, D. *Kant’s notion of deduction and the methodological background of the first Critique*, p.31.
presented in a deduction should relate the origin to fundamental facts that constitute it. Kant used the term *deduction* having in mind the deduction writings and not the logical procedure\(^4\).

Kant called metaphysical deduction the task of referring to a non empirical origin, or to identify the origin before justifying the legitimacy of a possession or a usage. Once the origin is identified it is possible to go on to the task of justification, which Kant called transcendental deduction. What he was trying to justify by a deduction was the possibility of synthetic *a priori* judgments. And this is the main ground for calling his project the critical philosophy. According to him, when we come to *a priori* judgments we have to consider their origin in the nature of reason itself and so to justify them and explain their possibility. These tasks of justification and explanation belong to a critique of reason by itself.

In relation to the theoretic use of reason, his concern was to prove how some concepts can refer *a priori* to objects, and what was deducted in the first *Critique* were the categories. According to Kant, the philosopher should distinguish from one another the *a priori* and empirical elements of human cognition, and then consider what is the justification for accepting the *a priori* elements.

In relation to the practical use of reason, he is trying to prove that we, human beings, can act and, consequently, morally judge our actions. His starting point is the common rational cognition. Kant agrees with everyman that moral judgments are true or false, i.e., that we do can say that to explode a bomb in a kindergarten just because we don’t like kids is wrong! On the other hand, he agrees with positivists that moral judgments cannot be verified or justified by appealing to experience. He internalizes the origin of the moral law and says that human beings *ought* to act morally well because the moral law is a self-imposed one. So, he identifies the origin of the moral law in reason, and this means that the moral law is an *a priori* one. But since from the mere analysis of the concept of a human being does not follow the concept of acting morally well, or obeying the moral law, the moral law (for human beings) is said to be a synthetic *a priori* principle. And, as we saw above, it has to be justified by a transcendental deduction.

\[]\(^4\) It is worth noting that beyond the fact that Kant was familiar with the practice of deduction writing, as Henrich pointed out, Kant himself recognizes that he had in mind the works of jurists as we can see in the first sentences of the §13 of the Transcendental Deduction in the *Critique of Pure Reason*: “Jurists, when they speak of entitlements and claims, distinguish in a legal matter between the questions about what is lawful (*quid juris*) and that which concerns the fact (*quid facti*), and since they demand proof of both, they call the first, that which is to establish the entitlement or the legal claim, the *deduction*” [*CPR A84/B116*]. Moreover, he characterizes his critical project as “the *court of justice* [emphasis mine], by which reason may secure its rightful claims while dismissing all its groundless pretensions (…)” [*CPR A XII*] or as “the true *court of justice* [emphasis mine] for all controversies of pure reason” [*CPR A751/B779*] and “the universal human reason itself” as the only “*judge* [emphasis mine]” [*CPR A752/B780*]. That is, Kant himself uses legal metaphors to describe his project and this is something that can also be used to confirm that he had in mind the legal procedure when talking about deduction as a kind of proof.
Before we move on and analyze the deduction itself, it is important to have in mind that the content and the origin of the moral law is something Kant has expounded in the first two sections of the *Groundwork*. In the first section, from an analysis of common rational cognition, he arrives at the condition for a moral action, which is the obedience to a law. In the second section, he presents the content of the moral law by means of some formulations of it and identifies its origin in reason. For our present purpose the formula of autonomy is the most important. It states that the rational agent has to “act only so that the will could regard itself as at the same time giving universal law through its maxim” [G4: 434].

Autonomy is the principle behind moral judgments and intends to be the condition for moral action. It expresses the essence of moral law and it is the principle on which a rational agent would act if reason had full control over passion. That is, although Kant doesn’t make it clear, the principle of autonomy doesn’t need to take the form of a categorical imperative. Now, once we know which is the principle of morality, the next step is to ask about how it can be justified.

Kant first tries to justify it as a moral law and then as a categorical imperative. But since the autonomy expresses the essence of the moral law, which is the principle on which a rational agent would act if reason had full control over passion, and since it appears for a human being, who sometimes can act under influence of passion, as a categorical imperative, that is, as a principle on which she ought to act, the question about the justification of the principle of autonomy can be put as ‘how is a categorical imperative possible?’. And this is exactly the question Kant asks on the headline of subsection 4 of the third section of the *Groundwork*, and to which he will answer with a transcendent deduction.

III

In the third section of the *Groundwork*, Kant’s purpose is to justify how the moral law is possible to be a motive for human actions. Since the moral law under human conditions was proved to be a categorical imperative and hence a synthetic *a priori* proposition, and since, according to Kant, the proper way to prove a synthetic *a priori* proposition is by a transcendental deduction, we should look for this kind of argument in the third section.

Kant’s argument is given in subsection 4, under the title ‘How is a categorical imperative possible?’. The core of the argument, which is an answer to that question, is the following passage:

> But because the world of understanding contains the ground of the world of sense and so too of its laws, and therefore immediately lawgiving with respect to my will (which belongs wholly to the world of understanding) and must accordingly also be thought as such, it follows that I shall cognize myself as intelligence, though on the other side as a being belonging to the world of sense, as nevertheless subject to the law of the world of understanding, that is, of reason, which

---

5 Willing the good action is not necessarily contained in the volition of a human being, endowed with reason and sensibility, that is, it cannot be ‘analytically derived’ from the volition of such a being. In G4: 420n, Kant justifies that the categorical imperative is an *a priori* synthetic practical proposition, because it “does not derive the volition of an action analytically from another volition already presupposed (for we have no such perfect will), but it connects it immediately with the concept of the will of a rational being as something that is not contained in it”.

3
contains in the idea of freedom the law of the world of understanding, and thus cognize myself as subject to the autonomy of the will; consequently the laws of the world of understanding must be regarded as imperatives for me, and actions in conformity with these as duties. [G453-4]6

To analyze the argument, we can break it down in two main parts: the first one is stated by the element ‘because’ and provides a reason; the second is stated by the expression ‘it follows’ and provides a conclusion. The first part, or the premises of the argument, can be rewritten as follows:

P1 – Because the world of understanding contains the ground of the sensible world;

P2 – and so too of its laws [and because the world of understanding contains the ground of the laws of the sensible world];

P3 – and [because] therefore immediately lawgiving with respect to my will (which belongs wholly to the world of understanding) and must accordingly also be thought as such.

The second part, or its conclusion, that can also be broken down in two parts, as follows:

C – It follows that I shall cognize myself as intelligence, though on the other side as a being belonging to the world of sense, as nevertheless subject to the law of the world of understanding, that is, of reason, which contains in the Idea of freedom the law of the understanding world, and thus cognize myself as subject to the autonomy of the will;

C - consequently the laws of the world of understanding must be regarded as imperatives for me, and actions in conformity with these as duties.

The first premise is not difficult to understand. It states just that the intelligible world contains the ground of the sensible world. The meaning of the second premise is also compelling, and it actually follows from P1 as a corollary. The third premise, otherwise, is a little bit more problematic. One might think that because of the premises 1 and 2, that is, because of the relation between the intelligible world and the sensible world established in the first two premises, the intelligible world is also immediately legislative for my will, which belongs entirely to the intelligible world. But this cannot be true.

To better understand the third premise we could rewrite it in the follow way:

---

6 In German: “Weil aber die V estandeswelt den Grund der Sinnewelt, mithin auch der Gesetze derselben, enthält, also in Ansehung meines Willens (der ganz zur Verstandeswelt gehört) unmittelbar gesetzgebend ist, und also auch als solche gedacht werden muss, so werde ich mich als Intelligenz, obgleich [emphasis mine] andererseits wie ein zur Sinnenwelt gehöriges Wesen, dennoch dem Gesetze der ersteren, d.i. der Vernunft, die in der Idee der Freiheit das Gesetz derselben enthält, und also der Autonomie des Willens unterworfen erkennen, folglich die Gesetze der Verstandeswelt für mich als Imperative und die diesem Prinzip gemassen Handlungen als Pflichten ansehen müssen”. [G4: 453-4].
P3 – In relation to my will, which belongs entirely to the intelligible world, the intelligible world is directly legislative, and it must also be conceived as containing the ground of actions and laws of the sensible world.

The meaning of the premise itself is not problematic, what is problematic is the relation between it and the first two premises; better, the problem is relative to the element ‘because’ in the beginning of the first premise and the function of the expression ‘and therefore’, which introduces the third premise. That is, one might think that P3 is a conclusion that follows from premises 1 and 2. But it would be a complete misinterpretation of what Kant is arguing for. It would make no sense to support that because the intelligible world contains the ground of the sensible world and of its laws, it is also directly legislative for the will. It would make no sense because the will belongs entirely to the intelligible world; it doesn’t belong to the sensible world. So, there is no such a kind of relation between the premises 1 and 2 with P3. The simplest way to solve this misinterpretation is to put a ‘because’, or any other element that indicates a reason, in the beginning of P3. So, it would be properly read as a premise, additional to 1 and 2.

Moreover, it is important to remark, that the ‘it’, in the second part of P3, refers to the intelligible world. Hence, what P3 express is that (1) the intelligible world is directly legislative to the will, which is part of this world, and (2) in relation to my will, as part of the intelligible world, this world “must also be conceived as containing the ground of actions and laws of the sensible world”. Thus, from (1) and (2), it is possible to say that the will, as part of the intelligible world, contains the ground of the sensible world and its laws. And this is the gist of the deduction: that is, that the pure will as part of the intelligible world contains the moral law as a categorical imperative for this will as part of the sensible world.

The conclusion of the argument is a little bit easier to understand, although it is not completely evident. To analyze it, it is possible to break it down in two parts, which are separated by the occurrence of the element ‘consequently’.

[...] it follows that I shall cognize myself as intelligence, though on the other side as a being belonging to the world of sense, as nevertheless subject to the law of the world of understanding, that is, of reason, which contains in the idea of freedom the law of the world of understanding, and thus cognize myself as subject to the autonomy of the will;[...]

In this first part, Kant maintains that the human being (I), inasmuch as she considers herself as intelligence and at the same time as a being that belongs to the sensible world, is subject to the law of the intelligible world and to the autonomy of the will. It is important to have in mind the conjunction and at the same time since the beginning of the sentence, although it doesn’t appear in this passage. It is not the being only as intelligence that is subject to the law of the intelligible world and to the

---

7 It is better to substitute the expression ‘though’, which appears in the passage, by the expression ‘at the same time’ for a clearly understanding of this argument. In a lot of passages through the whole Groundwork Kant emphasizes the simultaneity of both perspectives (intelligible and sensible) in relation to human beings and the imperative character of the moral law. And this simultaneity plays a fundamental role here.
autonomy of the will; it is the being that is both: intelligence and sensible, that is subject to reason (which contains the law of the intelligible world) and to the autonomy of the will. It would be a miscomprehension to take Kant to be supporting that a rational being only as intelligence is subject to the law of the intelligible world and to the autonomy of the will. If it were this, it would be impossible to explain how a human being is subject to the moral law and can take it as a motive to her actions, that is, it would be impossible to prove morality under human conditions. Moreover, the element ‘nevertheless’ wouldn’t be necessary. For a being who is only intelligence it is not necessary to consider itself as nevertheless subject to the law of the intelligible world and to the autonomy of the will.

In relation to the second part of the conclusion, “consequently the laws of the world of understanding must be regarded as imperatives for me, and actions in conformity with these as duties”, it is possible to say that what is a law of the intelligible world must appear, or be considered by a human being (me), as an imperative to such a being. This because she is not only intelligence, she is also part of the sensible world, so her will can also be affected and hence motivated to act by a law of this last world. And since an imperative is just “the formula of a command of reason and is expressed by an ‘ought”8, the actions according to it are called duties.

Thus, almost without realizing, the reader is faced with a complete deduction9. This deduction is the answer Kant provides to the question ‘How a categorical imperative is possible?’. The gist of the deduction is that the pure will as part of the intelligible world contains the moral law as a categorical imperative for this will as part of the sensible world. Behind the deduction there is the idea that, in somehow, the intelligible world is superior to the sensible world, that is, that the rational perspective of a human being is superior to her sensible perspective and the pure practical will contains the supreme condition of the will affected by sensible desires.

In the paragraph that follows the one we analyzed Kant gives again an answer to question ‘how a categorical imperative is possible?’, and this we can took as a summary of the deduction just presented, one of the formal characteristics of a good deduction as Henrich10 remarks. And then, Kant says: “the practical use of common human reason confirms the correctness (Richtigkeit) of this deduction” [G4:454].

8 See G 4: 413.
9 According to Liddell, the paragraph we analyzed and took to be the whole deduction is just the second part of Kant’s deduction in the Groundwork. The author affirms that the deduction starts in the subsection ‘Freedom must be presupposed as a property of the will of all rational beings’, and its second part is presented in the subsection ‘How is a categorical imperative possible?’. (LIDDELL, B. Kant’s ‘deduction’ in the ‘Grundlegung’, p.401-2). We disagree with Liddell. For us, the argument presented in the subsection ‘Freedom must be presupposed as a property of the will of all rational beings’ is important for the deduction, but it is just a preparatory argument together with the subsection ‘Of the interest attaching to the ideas of morality’. This seems to be also Kant’s idea, since he affirms, at the end of the first subsection ‘The concept of freedom is the key to the explanation of the autonomy of the will’, that he cannot yet answer the question of how a categorical imperative is possible, because “some further preparation is required”. So subsections 2 and 3 are the preparatory argument to the answer that will be given in subsection 4, where Kant presents the deduction itself.

Now we can conclude that (1) if the purpose of a transcendental deduction, to Kant, is to justify the legitimacy of a possession or a usage of a synthetic \textit{a priori} judgment or principle; (2) if the autonomy, under human conditions, was proved to be a categorical imperative and hence a synthetic \textit{a priori} principle; (3) if the membership in the intelligible world is what justifies that the human being, finite rational being, is autonomous and hence can act morally well, and (4) if the paragraph we analyzed is the argument that proves that the human being belongs to/is member of the intelligible world, so this paragraph presents a complete transcendental deduction.

IV

It is important to remark that our interpretation is distinguished from the interpretations of great commentators. In the present text my purpose is to analyze Allison’s thought in relation to the deduction of the third section of the \textit{Groundwork} and show why we think he is wrong.

According to Allison the deduction, whose pivotal point is the move from possession to reason to membership in the intelligible world, can be presented in seven steps:

1. “Now I assert that every being who cannot act except under the idea of freedom is by this alone – from a practical point of view – really free”.
2. “And I maintain that to every rational being possessed of a will we must also lend the idea of freedom as the only one under which we can act”. (G4:448)
3. All laws “inseparably bound up with freedom” are valid for every being with reason and will.
4. But the Reciprocity Thesis establishes that the moral law is inseparably bound up with freedom”.
5. Therefore, the moral law is valid for every being with reason and will.
6. Since beings such as ourselves have reason and will, the moral law is valid for us.
7. Since we do not necessarily follow the dictates of the law (these dictates being “objectively necessary” but “subjectively contingent”), the law for us takes the form of a categorical imperative, that is, we are rationally constrained, although not causally necessitated, to obey it.

Steps 1 and 2 consist in a preparatory argument, as the commentator calls it. Step 7 consists in “a distinct deduction of the categorical imperative” (Allison. p.224). Steps 3 to 6 consist, therefore, in a deduction of the moral law. And here it is the first difference between ours and Allison’s analysis.

Allison notes that the third section of the \textit{Groundwork} is one of the most enigmatic of Kantian texts. Although it is clear that its main purpose is to justify the supreme principle of morality, articulated in the first two sections, and for that Kant appeals to a deduction; it is not clear whether the deduction is of the moral law, the categorical imperative, freedom, all three, or even whether it can be properly characterized as a deduction at all. Allison’s argument is for a deduction of the moral law, and his underlying presupposition is the reciprocity thesis.
As already indicated, our first disagreement with Allison is about what Kant is trying to justify by a deduction. We argue for a deduction of the categorical imperative, he argues for a deduction of the moral law.

In section two of the *Groundwork*, Kant claims that the “principle of autonomy is the sole principle of morals can well be shown by mere analysis of the concepts of morality” [G4:440]. However, that it is a categorical imperative “cannot be proved by mere analysis (...), because it is a synthetic proposition” [G4:440] and for such a proof “one would have to go beyond cognition of objects to a critique of the subject, that is, of pure practical reason” [G4:440], a business that “does not belong in the present section” [G4:440] says Kant. Moreover, in the end of the same section, he emphasizes that the proof that morality is not a chimera to human beings “follows if the categorical imperative (...) is true and absolutely necessary as an a priori principle” and this “requires a possible synthetic use of pure practical reason, which use, however, we cannot venture upon without prefacing it by a critique of this rational faculty itself, the main features of which we have to present, sufficiently for our purpose, in the last section” [G4:445].

In section three, right after the argument that justifies the possibility of the categorical imperative, Kant claims that “the practical use of common human reason confirms the correctness of this deduction” [G4:454].

Now, according to the passages we quoted it seems reasonable to regard Kant’s attempt to prove the possibility of the categorical imperative, and not of the moral law, by a deduction. Moreover, since a deduction is to prove the transcendental conditions of a possession or an usage of a synthetic a priori judgment or principle, and since the moral law follows analytically from the concepts of morality, it wouldn’t be necessary a deduction to justify it.

But, even if Allison’s suggestion that the deduction is of the moral law and not of the categorical imperative was right, he argues for a failure of the deduction due to a fatal ambiguity in two central notions. The first is in that of the intelligible world and the second in that of the will.

In relation to the ambiguity in the notion of an intelligible world, Allison supports that Kant refers to both a *Verstandeswelt* and an *intelligibelen Welt* and slides from the former to the second without sufficient justification.

The *Verstandeswelt* is to be understood negatively as encompassing whatever is nonsensible or “merely intelligible”, that is, whatever is thought to be exempt from the conditions of sensibility (the noumenon in the negative sense). The *intelligibele Welt*, on the other hand, is to be understood positively as referring to a supersensible realm governed by moral laws, a “kingdom of ends” or “the totality of rational beings as things in themselves” [G4:458] (the noumenon in the positive sense).

Allison notes that Kant’s goal is to show that human beings are members of an *intelligibelen Welt* because this would entail that they stand under the moral law. The problem is that the possession of reason only gets us to a *Verstandeswelt*, and since this world is an indeterminate concept, it cannot provide any conclusion about the nature of the rational being as a whole nor about her will.
The second difficulty, related to an ambiguity of the notion of the will, is a corollary of the former. The main point is that given the identification of will and practical reason the claim that rational beings possess a will can mean (1) merely that reason is practical or (2) that pure reason is practical. The former (practical freedom) is sufficient for us to affirm that we are genuine rational agents rather than automat; but it is the second (transcendental freedom) that is necessary to establish our autonomy.

The problem, again, is that the membership in the Verstandeswelt provides support just for practical freedom, but it is transcendental freedom that is necessary and sufficient to establish morality on the basis of a nonmoral premise about our rationality.

Thus, Allison concludes for a failure of the deduction, which, according to him, Kant himself may have recognized. Assuming that he claims that

we can see why he <Kant> would abandon the attempt to establish the practicability of pure reason on the basis of any nonmoral premise. Thus, instead of beginning with the concept of a rational agent and moving from this first to the presupposition of freedom and then, via the Reciprocity Thesis, to the moral law, Kant there <Critique of Practical Reason> moves directly from the consciousness of the moral law as the ‘fact of reason’ to the practicability of pure reason and the reality of transcendental freedom [p.228]. (Emphasis mine)

My critic to Allison is based on three points. First is that knowing the way in which Kant uses the concepts, it is more reasonable to take him using the concepts of a Verstandeswelt and of an intelligibelen Welt not in a univocal sense, but, sometimes, exchanging them. Second, that in the Groundwork, it is not Kant’s purpose to prove the reality (objective validity), but the possibility of the categorical imperative and for that the noumenon in the negative sense is necessary and sufficient.

The noumenon in the negative sense is a being of understanding “insofar as it is not an object or our sensible intuition” [CPR B307], but it can be under determination of space and time, that is, it is the object that can appear (as a phaenomenon). Admittedly, is the noumenon in the negative sense that Kant needs to justify the possibility of the categorical imperative because the categorical imperative is the way the moral law, a law of a being of understanding, appears to a being that is not only a being of understanding, but it is also a being of sense, a being that is under determination of space and time. To prove the possibility of the categorical imperative is to prove that the sensible affected will can give meaning, through its actions, which appear in space and time, to the rules of the pure rational will, and for that, the noumenon in the negative sense is necessary. If we have just the noumenon in the positive sense, that is, “the object of a non-sensible intuition (...), namely intellectual intuition” [CPR B307], we would have two different worlds and no connection between them. Hence, it would be impossible to justify how a being that is also part of the sensible world could be motivated to act by a law of the intelligible world. It is the noumenon in the negative sense that allows us to understand the intelligible world and the sensible world as a double standpoint of the same world, a double standpoint that the human being considers herself and allows to understand why she has to act morally well.

Finally, our last critic to Allison, is that what gets us to the intelligible world it is not just the possession of reason but the consciousness of the spontaneity of reason. And this seems to give a
positive content to our thought of ourselves as members of the intelligible world and, hence, a positive content to the concept of an intelligible world itself. That is, it seems reasonable to support that the consciousness of the spontaneity of reason presupposes a law different from that of nature, and this allows a positive characterization of the intelligible world and, therefore, a characterization of the human being as a noumenon in the positive sense.

Notwithstanding, Almeida\(^\text{11}\) notes that the characterization of the human being as a noumenon in the positive sense, allowed by the consciousness of the spontaneity of reason, can take us to another problem. The problem is that although the rational being can ‘transport’ herself to the intelligible world by the consciousness of the spontaneity of theoretic reason this is not sufficient to ascribe the same spontaneity to practical reason. And this, according to him, is the unsolvable problem that made Kant abandon his attempt to prove the supreme principle of morality by a deduction and appeal to a \textit{fact of reason} in the second \textit{Critique}. But to face this problem is something we will do in another work.

In the beginning of this paper we stated as our objectives to show the argument whose purpose is to prove the possibility of the categorical imperative; to show that the argument is a transcendental deduction; to present the argument as it is reconstructed by Allison, and, finally, to show that, although it seems compelling, the position of the commentator is wrong. Now we can say that our objectives were achieved.

BIBLIOGRAPHY


PROOPS, I. *Kant’s legal metaphor and the nature of a deduction*. In Journal of the History of Philosophy; Apr. 2003; 41, 2, Platinum Periodicals.