

Defending Kant's Antinomy of Practical Reason

Christopher Benzenberg
University of Cambridge

This is a penultimate draft. Please refer to the published version.

Abstract — This paper defends Kant's 'Antinomy of Practical Reason' (APR) against the objection that it fails to meet the standards of a strict Kantian antinomy. Critics claim that the APR cannot be a true antinomy because its thesis and antithesis do not contradict each other logically, unlike the 'Antinomies of Theoretical Reason' (ATRs). In response to this objection, it is argued that the theses and antitheses of the ATRs do not contradict each other logically either; they only contradict each other relatively. It is then shown that the thesis and antithesis of the APR, properly understood, also contradict each other relatively, thus satisfying the same standard as the ATRs. This paper offers broader insights into the structure of antinomical conflict within Kant's Critical philosophy.

1. Introduction

Kant's antinomies of theoretical reason (short: ATRs) are well-known. They constitute a cornerstone of the critical discussion in the first Critique. Less famous but equally interesting is Kant's "Antinomy of Practical Reason" (KpV, AA 5: 113; short: APR),¹ which is placed at the heart of the second Critique's Dialectic. While there is some controversy as to what the thesis and antithesis of the APR ought to be,² the most natural reading can roughly be summarised as follows:

*The Natural Reading of the APR*³

- (thesis) Happiness is the cause of virtue.
(antithesis) Virtue is the cause of happiness.

The APR, thus understood, was criticised by several interpreters. Beck argues that the APR — a "poor thing" (1960, 246), as he describes it — is not an "antinomy in any strict sense" because its thesis and antithesis "are not contradictories" (1960, 247). The same argument is advanced by Albrecht (1978, 184) and Milz (2002, 15–16), authors of the only two monographs ever written on the APR. In honour of the authors' initials, and because of its punch, I call this the BAM-Argument:

*The BAM-Argument*⁴

- (P1) If X is an antinomy in the strict Kantian sense, then the thesis and antithesis of X contradict each other.
(P2) The thesis and antithesis of the APR do not contradict each other.
(C) *Thus*, the APR is not an antinomy in the strict Kantian sense.

¹ Translations are my own.

² Milz provides a synopsis of the different readings (2002, 12–98).

³ See KpV, AA 5: 113. The Natural Reading is assumed by Beck (1960, 246), Albrecht (1978, 95), Wike (1984, 425), Faizzada (2017, 280–1), and Gardner (2017, 40).

⁴ For a similar argument, see Wike (1982, 126–129), (1984, 426), Watkins (2010, 154n.), Himmelmann (2015, 126), and Marwede (2018, 70n.).

The BAM-Argument seems plausible. The standard for an antinomy in the strict Kantian sense is set by the ATRs, which give rise to proper contradictions. The First Antinomy — to give just one example — *affirms* and *denies* that the world has a beginning in time and space (KrV, A 426–427/B 454–455). By contrast, the thesis and antithesis of the APR posit two different directions of an assumed causal relation between virtue and happiness. This doesn't appear to yield a contradiction similar to the ones expressed by the theses and antitheses of the ATRs.

In this paper, I defend the APR against the BAM-Argument. After distinguishing between logical and relative contradictions (§ 2), I argue that the theses and antitheses of the ATRs contradict each other relatively and not logically (§ 3). I then propose that the thesis and antithesis of the APR also contradict each other relatively and not logically (§ 4). Finally, I suggest that an alternative reading of the APR fails to address the BAM-Argument (§ 5).

2. Logical and Relative Contradictions

Kant explicitly states that the thesis and antithesis are “two propositions, which contradict each other” (Prol, AA 4: 341).⁵ But what does it take for two propositions to contradict each other? One popular way to explicate the sense in which two propositions are contradictory is to say that they are members of an exclusive disjunction: proposition p and q contradict each other iff either p or q , or formally “ $p \vee q$ ”. As Wike puts it, two contradictory propositions “*exclude* one another, and together *exhaust* the field [...] in question” (1984: 428; my emphasis; see also A 74/B 99; Log, AA 9: 107).

Yet this general definition of *contradictions as disjunctions* does not provide the conceptual precision necessary to analyse the Kantian

⁵ See also KrV, A viii, B 8, A 405/B 433, A 407/B 433, A 421/B 449; Prol, AA 4: 338, 343. The contradictory relation was reasserted by Beck (1960, 247), Strawson (1966, 184), Albrecht (1978, 184), Wike (1984, 426), Malzkorn (1999, 88–94), Förster (2002, 174), Milz (2002, 133), Allison (2004, 358, 364, 390), and Watkins (2010, 146).

antinomies. We must further distinguish between, what I call, *logical* and *relative* contradictions. Let's start with logical contradictions. The two claims "*S* is *P*" and "*S* is not *P*" contradict each other logically because they contradict each other in virtue of their logical form. Put differently, the disjunction "*S* is *P* $\dot{\vee}$ *S* is not *P*" is logically true. Since logical truths can be deduced from the empty set, I define:

Logical Contradictions

(LC) Two propositions *p* and *q* contradict each other logically iff $\emptyset \vdash "p \dot{\vee} q"$.

But there is another way in which two propositions can contradict each other. We might want to say that, given certain assumptions, it has to be the case that either *p* or *q*. For instance, given the assumption that *S* is alive, it must be true that either *S* is healthy or *S* is ill. But of course, "*S* is healthy $\dot{\vee}$ *S* is ill" can't be deduced from the empty set, for *S* might very well be dead. I propose that this second sense of contradiction is best captured in the following definition:

Relative Contradictions

(RC) Two propositions *p* and *q* contradict each other relative to a set of propositions Σ iff $\Sigma \vdash "p \dot{\vee} q"$.

Two notes: First, all logical contradictions contradict each other relative to any set Σ . Everything that is deducible from the empty set can also be deduced from any other set Σ . To demonstrate that two propositions contradict each other *merely* relatively, one must therefore show that they aren't logical contradictories. Second, for any two propositions *p* and *q*, there is always a set Σ such that *p* and *q* contradict each other relative to Σ — just take the set $\{"p \dot{\vee} q\}$. This is no cause for concern, however. Most of these sets are irrelevant because they contain absurd propositions.

The distinction between logical and relative contradictions can be used to address the BAM-Argument. Wike, who introduces a similar distinction (1984, 426), claims that, while only ATRs constitute a

logical contradiction, the APR also constitutes a contradiction, albeit a relative one. Wike then argues that both types of contradictions, logical and relative, are instances of contradictions in the general disjunctive sense. That's why, according to Wike, the APR qualifies as an antinomy in the strict Kantian sense (1984, 431).

I agree with Milz, who criticises Wike's strategy as extending "the concept of an antinomy [...] to the borders of a sheer equivocal meaning" (2002, 80). Surely there is some broad sense of contradiction exhibited by the ATRs, which also applies to the thesis and antithesis of the APR. But this way of refuting the BAM-Argument is trivial and uninteresting. To properly dismiss the BAM-Argument, one must either show that both the ATRs and the APR constitute logical contradictions, or that both constitute relative contradictions. That's the challenge.

3. The Antinomies of Theoretical Reason

Like Wike, many interpreters — especially the BAM-authors — argue that Kant has logical contradictions in mind when he ascribes a contradictory relation to the ATRs. Milz, for example, claims that the thesis and antithesis of the ATRs are "opposed in the manner of a logical contradiction; [...] following the form: $a \vee \sim a$ " (2002, 132–133).⁶ This requirement for logical contradictions explains why Beck, Albrecht and Milz think that the APR fails to meet the high standard of an antinomy in the strict Kantian sense. But what if the ATRs didn't exhibit logical contradictions after all?

In this section, I argue that the theses and antitheses of the ATRs contradict each other relatively and not logically. Following RC and LC, I first show that the disjunction between the theses and antitheses of the ATRs can be deduced from assumptions in the first *Critique's* Dialectic. I then argue that the disjunction cannot be deduced from the empty set. Rather than focusing on each of the ATRs individually,

⁶ See also Beck (1960, 247–248) and Albrecht (1978, 184).

my argument draws on Kant's general discussion of the ATRs. Let's start with the *sophisma figurae dictiones*:

The entire antinomy of pure reason rests on the dialectical argument: If the conditioned is given, then the whole series of all its conditions [i.e. the totality of conditions; C.B.] is also given; now objects of the senses [i.e. appearances; C.B.] are given as conditioned; therefore, etc. (KrV, A 497/B 525)

To be valid, the argument must be augmented in two ways. First, Kant states that the totality of conditions is only given if the conditioned is a thing in itself: "[I]f the conditioned [...] is a thing in itself, [...] then the complete series of conditions is [...] given at the same time" (KrV, A 498/B 527). Second, the argument is only valid if we add the essential claim of transcendental realism, which "represents [...] appearances [...] as things in themselves" (KrV, A 369).⁷ The *sophisma*-argument therefore rests on the following three assumptions:

- (1) If X is a thing in itself and if X is conditioned, then the totality of conditions is given.
- (2) If X is an appearance, then X is conditioned. (And there is at least one appearance A .)
- (3) X is an appearance iff X is a thing in itself.

At this point, one might object that a valid *sophisma figurae dictiones* is an oxymoron. Kant defines the *sophisma* as a fallacy "in which the middle term is used equivocally" (Log, AA 9: 135; see also Allison 2004, 361). Although the *sophisma* is invalid if we do not assume transcendental realism, the argument must become valid if transcendental realism, i.e. assumption (3), is added, "for the antinomies, unlike the paralogisms and the ideal, are not supposed to offer us a clear case of formally invalid argumentation" (Grier 2001, 172).

⁷ "Transcendental Realism – the view that appearances and things in themselves are identical" (Watkins 2010, 146). See also Malzkorn (1999, 101).

From {(1), (2), (3)}, we can deduce that the totality of conditions is given. Yet the totality of conditions, i.e. the unconditioned,⁸ is either *finite* or *infinite*: “[Either] every member without exception is conditioned, and only their whole is absolutely unconditioned, or else the absolutely unconditioned is only a part of the series [...] that itself stands under no other condition” (KrV, A 417/B 445). Let us therefore add one further assumption:

- (4) If the totality of conditions is given, then there is either a finite totality of conditions or there is an infinite totality of conditions, i.e. *either* the world is finite *or* the world is infinite.

From {(1), (2), (3), (4)}, we can deduce that *either* the world is finite *or* the world is infinite. Now, Kant distinguishes between four different conditioning relations R: space-time, parts, causality, and necessity (KrV, A 411–415/B 439–442; see also Willaschek 2018, 73–86). For each of these conditioning relations R, it is the case that *either* the world is R-finite (thesis) *or* the world is R-infinite (antithesis). It follows for all four ATRs: {(1), (2), (3), (4)} ⊢ “thesis ∨ antithesis”. The theses and antitheses of the ATRs therefore contradict each other relative to {(1), (2), (3), (4)}.

Two clarificatory notes: First, to show that the ATRs constitute relative contradictions, I don’t have to justify any of the four claims (1), (2), (3), or (4); I only have to show that the disjunction “thesis ∨ antithesis” logically follows from them. In fact, (3) is the essential claim of transcendental realism and thus false. Second, my reconstruction of Kant’s argument is only a rough outline. However, I hope to have shown that the contradictory relations between the theses and antitheses follow from substantive metaphysical assumptions (no matter the details).

I now turn to the second claim of this section: The theses and antitheses of the ATRs do not contradict each other logically. Let’s

⁸ See KrV, A 322/B 379, A 409/B 436, A 417/B 444–445. Willaschek suggests that, “[f]or all x, x is unconditioned [...] if x is the totality of [...] conditions” (2018, 96). For further discussion, see also Wike (1982, 48) and Watkins (2010, 149).

assume the opposite for *reductio*. In this case, the disjunction “thesis $\dot{\vee}$ antithesis” could have been deduced from the empty set, and thus stated without further ado. But this is not what happens. As we have seen, Kant provides a long and complex argument in support of the disjunction. If the ATRs were to constitute logical contradictions, Kant’s discussion of (1), (2), (3), and (4) would be superfluous, which is exegetically unattractive.

Moreover, it is perfectly possible that the world, understood as the totality of conditions, isn’t given at all. In this case, both the theses and the antitheses would simply be false. In fact, this is precisely how Kant resolves the ATRs. He suggests that the theses (“the world is finite”) and the antitheses (“the world is infinite”) “can both be false [...] if the world [...] wasn’t given” (KrV, A 504/B 532). That’s why Kant concludes that the ATRs don’t constitute a logical, or as he calls it, “proper contradiction” (KrV, A 503/B 531).

Against this, one might object that the theses and antitheses of the ATRs have to constitute logical contradictions, or else Kant’s indirect proof of transcendental idealism wouldn’t succeed. The indirect proof is a *reductio* argument which builds on the fact that transcendental realism, which gives rise to the ATRs, is an inconsistent theory (KrV, A 506–507/B 534–535). Yet a theory (i.e. a set of propositions Σ) is inconsistent only if it entails a logical contradiction. Put formally: Σ is inconsistent only if there is a $\Sigma \vdash \neg p$.

I agree that transcendental realism must entail a logical contradiction. But I disagree that this contradiction holds between the theses and antitheses of the ATRs. Above I have reconstructed the dialectical argument for the disjunction (a) “thesis $\dot{\vee}$ antithesis”. The apagogic proofs further show that the theses and the antitheses of the ATRs are false: (b) “ \neg thesis” and (c) “ \neg antithesis”. Transcendental realism is inconsistent because (in theory) it is committed to (a), (b), and (c);⁹ claims (a) and (c) entail (d) “thesis”, which *logically* contradicts (b) “ \neg thesis”. But note that the logical negation of the thesis is not the antithesis.

⁹ See KrV, A 506–507/B 534–535, as well as Strawson (1966, 110) and Malzkorn (1999, 116–118).

To summarise, I have argued that the ATRs, which provide the standard for an antinomy in the strict Kantian sense, only exhibit relative contradictions. The theses and antitheses contradict each other only relative to a series of metaphysical assumptions (in particular transcendental realism). Based on textual and philosophical considerations, I have then rejected the commonly held view that the ATRs give rise to logical contradictions of the form “ p ” and “ $\neg p$ ”. I have further suggested that my reading doesn’t jeopardise Kant’s indirect proof.

4. The Antinomy of Practical Reason

In this section, I argue that the thesis and antithesis of the APR also contradict each other relatively and not logically. Following RC, I first show that the disjunction between the APR’s thesis and antithesis can be deduced from assumptions in the second *Critique’s* Dialectic. There Kant claims that we have the duty to attain the highest good.¹⁰ As the inability to do our duties would invalidate the moral law, Kant argues: “If [...] the highest good is impossible [...] the moral law [...] must be false” (KpV, AA 5: 114); yet the moral law is valid (KpV, AA 5: 47):

- (5) If the moral law is valid, then the highest good is possible.
- (6) The moral law is valid.

From (5), (6) we can infer the possibility of the highest good. Drawing on this conclusion, Kant formulates a lengthy argument at the beginning of the second *Critique’s* antinomy-chapter.¹¹ An abbreviated version of that argument will suffice for the purposes of this paper:

¹⁰ See KpV, AA 5: 113, 119, 122, 129, 135. See also Willaschek (2016) and Benzenberg (2018).

¹¹ For a comprehensive discussion, see Milz (2002) and Watkins (2010).

In the highest good [...], virtue and happiness are thought as necessarily combined [...]. Now, this combination [...] must be thought [...] as the connection of cause and effect [...]. Thus, either the desire for happiness must be the motivating cause to maxims of virtue, or the maxim of virtue must be the efficient cause of happiness. (KpV, AA 5: 113)

In this argument, Kant introduces two further assumptions:

- (7) If the highest good is possible, then virtue and happiness are necessarily combined.
- (8) If virtue and happiness are necessarily combined, then either happiness is the cause of virtue or virtue is the cause of happiness.

Recall the Natural Reading of the APR, which I have introduced at the beginning of this paper. On this reading, the thesis and antithesis of the APR roughly state "Happiness is the cause of virtue" and "Virtue is the cause of happiness". Assumptions (5), (6), (7), and (8) therefore entail the disjunction between the APR's thesis and antithesis: $\{(5), (6), (7), (8)\} \vdash \text{"thesis} \dot{\vee} \text{antithesis"}$. This is to say that the APR's thesis and antithesis contradict each other relative to $\{(5), (6), (7), (8)\}$.

I agree with the BAM-authors that the thesis and antithesis of the APR don't contradict each other logically. For if they were logical contradictories, the disjunction "thesis $\dot{\vee}$ antithesis" would follow from the empty set. But then Kant's lengthy argument for the disjunction would be redundant, which it isn't. Moreover, the thesis and antithesis only posit two different directions of an assumed causal relation between virtue and happiness. However, it is perfectly possible that there simply is no such relation between virtue and happiness.¹²

¹² Milz points out that "one generally doesn't get contradictory sentences by inverting the causal direction" (2002, 132). Wike makes a similar point (1982, 16–17, 126–129).

5. Concluding Thoughts

In this paper, I have defended the APR against the BAM-Argument. After distinguishing between logical and relative contradictions (§ 2), I have argued that the thesis and antithesis of the ATRs, which set the standard for an antinomy in the strict Kantian sense, contradict each other relatively and not logically (§ 3). I have then suggested that the thesis and antithesis of the APR also contradict each other relatively and not logically (§ 4). That's why the BAM-Argument is not sound.

Against my general line of argument, one could object that Kant himself doesn't distinguish between logical and relative contradictions. While this might be true,¹³ I think that the distinction is justified as an interpretative tool that helps to analyse the Kantian antinomies. This being said, I also think that the burden of proof is now on those that wish to resurrect the BAM-Argument to provide a non-gerrymandered account of contradictions that applies to the ATRs but not the APR – the account of logical contradictions won't do the job.

At this point, it is also worth noting that some authors have tried to avoid the BAM-Argument by rejecting the Natural Reading of the APR, which is the reading I assumed in this paper. As I have noted in the introduction, there is some controversy as to what the thesis and antithesis of the APR ought to be. In contrast to the ATRs, Kant never explicitly labels two sentences as *the* thesis and *the* antithesis of the APR. While the Natural Reading enjoys the most textual support, many interpreters have rejected it on systematic grounds in favour of the following reading:

¹³ Though I conjecture that the distinction between logical and relative contradictions roughly maps onto what Kant calls "*dialectical*" and "*analytical opposition[s]*" (KrV, A 504/B 532).

*The Systematic Reading of the APR*¹⁴

(thesis) The highest good is possible.

(antithesis) The highest good is not possible.

The main motivation for the Systematic Reading is that it is thought to avoid the BAM-Argument. Milz, for example, suggests that it is only on the Systematic Reading that the thesis and antithesis of the APR exhibit an “explicit contradictory form” (2002, 174). Watkins argues along the same lines that the Systematic Reading “allows one to avoid the problem that [...] [the thesis and antithesis of the Natural Reading; C.B.], which one might naturally take to be the Thesis and Antithesis, do not in fact contradict each other” (2010, 154n.).

It should be clear from this paper’s argument that the motivation for the Systematic Reading is misguided. An antinomy in the strict Kantian sense (as defined by the ATRs) requires a relative rather than a logical contradiction. Yet on the Systematic Reading, the thesis and antithesis of the APR constitute a logical contradiction of the form “ p ” and “ $\neg p$ ”. This is too strong. The Systematic Reading should thus be rejected in favour of the Natural Reading, which is not only textually intuitive but also satisfies the criterion of exhibiting a *merely* relative contradiction.

In closing this paper, I want to emphasise that I have only shown that the APR survives the BAM-Argument. I have not shown that the APR *is* an antinomy in the strict Kantian sense. There might be further desiderata, which the APR fails to satisfy. In fact, there are two additional BAM-Arguments. The first suggests that the APR doesn’t exhibit the required proofs for its thesis and antithesis.¹⁵ The second criticises the APR based on its idiosyncratic resolution.¹⁶

¹⁴ See Beck (1960, 248), Förster (2002, 182), Milz (2002, 196), Watkins (2010, 154n.), Marwede (2018, 76), and Godess-Riccitelli (2019, 121n.).

¹⁵ See Beck (1960, 247), Albrecht (1978, 184), and Milz (2002, 130).

¹⁶ See Beck (1960, 247), Albrecht (1978, 119), and Milz (2002, 17).

While I think that these further objections can be addressed, I leave this task for another occasion.¹⁷

References

- Albrecht, Michael (1978): *Kants Antinomie der praktischen Vernunft*. Hildesheim: Georg Olms Verlag.
- Allison, Henry E. (2004): *Kant's Transcendental Idealism, An Interpretation and Defense*. New Haven: Yale University Press.
- Beck, Lewis W. (1960): *A Commentary on Kant's Critique of Practical Reason*. Chicago: The University of Chicago Press.
- Benzenberg, Christopher (2018): "Attaining the Highest Good". *GAP.10 Conference, Gesellschaft für Analytische Philosophie*, Cologne.
- Faizzada, Walid (2017): *Autonome Praxis und intelligible Welt, Die transzendental-praktische Freiheit in Kants Lehre vom höchsten Gut*. Leiden: Brill.
- Förster, Eckart (2002): "Die Dialektik der reinen praktischen Vernunft". In: Höffe, Otfried (Ed.): *Immanuel Kant, Kritik der praktischen Vernunft*. Berlin: De Gruyter, 173–186.
- Gardner, Sebastian (2017): "Kant's Practical Postulates and the Development of German Idealism". In: Gabriel, Markus et al. (Eds.), *German Idealism Today*. Berlin: De Gruyter, 27–72.
- Godess-Riccitelli, Moran (2019): "The Nature of Moral Faith: From Natural Beauty to Ethico-Theology in Kant's Third Critique". In: *Revista Portuguesa de Filosofia* 75, 117–144.
- Grier, Michelle (2001): *Kant's Doctrine of Transcendental Illusion*. Cambridge: Cambridge University Press.

¹⁷ I wish to thank Ralf Bader, Angela Breitenbach, Andrew Chignell, Pirachula Chulanon, Maya Krishnan, Bernhard Milz, Michael Oberst, Tobias Rosefeldt, and Myriam Stihl for their helpful comments on earlier drafts of this paper. I also received helpful feedback from the participants of the Berlin Kant Colloquium and the 13th International Kant Congress in Oslo.

- Himmelmann, Beatrix (2015): "Antinomie der praktischen Vernunft". In: Willaschek, Marcus et al. (Eds.), *Kant-Lexikon*. Berlin: De Gruyter, 125–127.
- Kant, Immanuel (1900–): *Kant's Gesammelte Schriften*, Königlich-Preußische Akademie der Wissenschaften (ed.), Berlin: de Gruyter.
- Kant, Immanuel (1998): *Kritik der reinen Vernunft*, Jens Timmermann (ed.), Hamburg: Felix Meiner Verlag.
- Malzkorn, Wolfgang (1999): *Kants Kosmologie-Kritik, Eine formale Analyse der Antinomienlehre*. Berlin: De Gruyter.
- Marwede, Florian (2018): *Das höchste Gut und Kants deontologische Ethik*. Berlin: De Gruyter.
- Milz, Bernhard (2002): *Der gesuchte Widerstreit, Die Antinomie in Kants Kritik der praktischen Vernunft*. Berlin: De Gruyter.
- Strawson, Peter F. (1966): *The Bounds of Sense, An Essay on Kant's Critique of Pure Reason*. London: Routledge.
- Watkins, Eric (2010): "The Antinomy of Practical Reason: Reason, the Unconditioned and the Highest Good". In: Reath, Andrews and Timmermann, Jens (Eds.): *Kant's Critique of Practical Reason, A Critical Guide*. Cambridge: Cambridge University Press, 145–167.
- Wike, Victoria S. (1982): *Kant's Antinomies of Reason, Their Origin and Their Resolution*. Washington: University Press of America.
- Wike, Victoria S. (1984): "Kant's Practical Antinomy". In: *Southern Journal of Philosophy* 22, 425–433.
- Willaschek, Marcus (2016): "Must We Believe in the Realizability of Our Ends? On a Premise of Kant's Argument for the Postulates of Pure Practical Reason". In: Höwing, Thomas (Ed.): *The Highest Good in Kant's Philosophy*. Berlin: De Gruyter, 223–244.
- Willaschek, Marcus (2018): *Kant on the Sources of Metaphysics, The Dialectic of Pure Reason*. Cambridge: Cambridge University Press.