



AUTHOR MEETS CRITIC

Proops’s ‘Nugget of Gold’ in Kant’s Dialectic

Desmond P. Hogan

Princeton University, Princeton, NJ, USA

Email: deshogan@princeton.edu

Abstract

The Fiery Test of Critique describes Kant’s indirect proof of idealism from the Antinomy of Pure Reason as the ‘nugget of gold’ in the *Critique of Pure Reason’s* Transcendental Dialectic. Here, I offer critical reflections on Proops’s reading of Kant’s indirect proof.

Keywords: Kant; idealist arguments; transcendental idealism; antinomy of pure reason; freedom

Ian Proops has written a hugely instructive and enormously enjoyable book on Kant’s Dialectic. In *The Fiery Test’s* closing reflection, he summarises the book’s main conclusion: ‘Kant regarded rational metaphysics as yielding two sorts of valuable residue when subjected to the fiery test of critique’ (p. 453).¹ My focus here is on what Proops calls the Dialectic’s ‘nugget of gold’: the indirect proof of idealism Kant extracts unexpectedly from the dialectical inferences of rational cosmology. ‘In recommending Transcendental Idealism as the uniquely correct way of resolving the antinomies’, the book concludes, Kant has, ‘or so he believes – provided the surest possible validation of his own Copernican revolution’ (p. 462).

1. Proof structure of indirect argument

Kant’s indirect argument for idealism proceeds from a view that, on the assumption of realism, inconsistent cosmological theses and antitheses ‘can be established through equally evident, clear, and incontestable proofs’ (*Prolegomena*, 4: 340). That *each* antinomy involves a contradiction entailed by realism, supposedly resolved by transcendental idealism, is itself strong evidence that the indirect proof is intended to advance from any one of Kant’s four (or five, depending how one counts) cosmological antinomies. It may therefore surprise that a majority position among leading contemporary scholars interprets Kant as holding that only his first and second antinomies indirectly prove idealism. Much is at stake here: if this majority view is correct, Kant’s third antinomy and resolution, with its famous discussion of human freedom, plays no role in his indirect argument. *The Fiery Test* reads Kant as holding that ‘*each* individual antinomy-resolution establishes Transcendental Idealism with apodictic certainty’ (p. 23). I concur, though it seems to me Proops’s argument would have benefited from a critical diagnosis of the majority view. I see two major sources

of error, one of which remains in place in Proops's own exposition. Both relate to Kant's official explanation of the indirect proof in a crucial passage I will label *Proof Structure*:

The [indirect] proof [of idealism] would consist in this dilemma. If the world is a whole existing in itself, then it is either finite or infinite. Now the first as well as the second alternative is false (according to the proof offered above for the antithesis on the one side and the thesis on the other). Thus it is also false that the world (the sum total of all appearances) is a whole existing in itself. From which it follows that appearances are nothing outside our representations, which is just what we mean by their transcendental ideality. (A506–7/B534–5)

One source of confusion arises from the *Critique of Pure Reason's* doctrine that the theses and antitheses of first and second antinomies 'are both false' (as mere contraries in the actual or idealist context), while thesis and antithesis of third and fourth antinomies 'may both be true' (A532/B560).² *Proof Structure's* reference to falsehood of a finitistic first and infinitistic second alternative may appear to echo this asymmetry, suggesting that only the first two antinomies underpin the indirect proof. Proops avoids this mistake and makes the confusion less likely. For he carefully distinguishes Kant's claim that dual falsity of finitistic and infinitistic positions follows *from transcendental realism* – that is, 'according to antinomy proofs' (p. 272), and his claim that both opposed assertions of the first and second antinomies are *in fact* false, that is, given transcendental idealism. This careful distinction between the realist locus of contradiction and the idealist locus of resolution is one of the many virtues of *The Fiery Test*.

If each antinomy gives rise to an indirect proof, why does *Proof Structure* refer only to the mathematical antinomies? The correct answer, I take it, is that it does not. Recognition of this would confirm Proops's general conclusion; however, he remains guilty of the usual mistake in claiming that *Proof Structure* 'is suited only to the mathematical and not to the dynamical antinomies' (p. 273). To see the error, consider again Kant's reference above to a provable falsity of the finitistic and infinitistic alternatives, 'according to the proof offered [respectively] for the antithesis on the one side and the thesis on the other'. It is natural to assume this refers only to the mathematical antinomies. Recall however that antithesis proofs in all four antinomies establish falsity of a finitistic view (equivalently, assuming realism, truth of an infinitistic view), while thesis proofs establish falsity of an infinitistic view (equivalently, assuming realism, truth of a finitistic view). To confirm that *Proof Structure* intends all four antinomies, compare Kant's warm-up discussion a page earlier:

The very same [resolution] holds of the series of causes ordered one above another [third antinomy], or of conditioned existence up to necessary existence [fourth antinomy], *which can never be regarded in themselves as either finite or infinite in their totality [i.e. as they are according to the third and fourth thesis/antithesis proofs]*, because, as series of subordinated representations, they exist only in the dynamical regress . . . (A505–6/B533–4)³

2. Role of thesis/antithesis arguments in Kant's indirect proof

If realism does not entail inconsistent cosmological theses and antitheses, there is no antinomy in Kant's sense of a necessary conflict, and his indirect argument fails. He sometimes states his idealist resolutions in a way that may suggest a short argument for idealism, bypassing the work of showing that the conflicts in question are unavoidable for the realist. Kant writes, 'the concept of a sensible world existing for itself is self-contradictory'; or again, 'The reason why the first two antinomies are both false is that I had to ground them both on a *contradictory concept*, namely, that of a whole in space and time that is also supposed to be an absolute whole, consequently a thing in itself' (*Prolegomena*, 4: 342; Refl 5962, 18: 401). It would however beg the question to offer either claim as itself justifying idealism. Kant's resolutions of the antinomies are indeed crucial to his indirect argument, insofar as he must show that alleged contradictions in realism do not also beset his idealist alternative. The actual 'unavoidability' of cosmological contradictions for the realist, however, underpins Kant's idealist resolution on which the empirical world 'is not given, and cannot be given, prior to the regression in it'. To Proops's credit, he resists the temptation to present Kant's resolutions as his idealist arguments. Only in his third antinomy discussion do I find some conflation of contexts of antinomy and resolution. It arises there from Proops's effort to find a place for freedom in Kant's indirect argument, an issue to which I will return.

3. Supreme principle of pure reason

If contradictions are not entailed by transcendental realism, Kant's indirect argument fails. Proops's verdict is the usual one that Kant indeed fails to show that realism entails contradictions. One especially mysterious source of this failure arises from the role in Kant's thesis arguments of a Supreme Principle of Pure Reason – here labelled *D* for 'descriptive'. Among its formulations is the following: 'If the conditioned is given, then the whole sum of conditions, and hence the absolutely unconditioned, is also given' (A409/B436). This Supreme Principle, which Kant claims is 'obviously synthetic', plays a central role not only in the Antinomies but also in the dialectical inferences of rational psychology and theology. For Kant portrays 'the immediate illusion lurking behind much of dogmatic speculative metaphysics as *the illusion that D is true*' (p. 48). In the cosmological context, the 'unconditioned' referenced in *D* is presented as representable in two inconsistent ways corresponding to antinomy theses and antitheses: as first member of a series of cosmological conditions, or as the series taken as a whole (A417/B445). In the first way, we arrive at the ideas of a world beginning and limit in space (first antinomy), simple constituents of the world (second antinomy), freedom in the sense of a first cosmological cause (third antinomy), and a necessary world cause (fourth antinomy).

Proops offers helpful though incomplete reconstruction of Kant's diagnosis of the source of the illusion that *D* is true (p. 46n16, p. 56). As he reads it, a prescriptive principle, inherent in reason, demands that we seek complete explanation. This in turn 'inevitably generates' the illusion that *D* is true. We slide illegitimately from reason's need for full explanation – and the fact that, 'if the conditioned is given, through it a regress in the series of all conditions is given to us as a *problem*' (A498/B526) – to the dogmatic *D*: 'Necessarily, for the object of any cognition,

the series constituting its unconditioned (real) condition *exists*' (pp. 47-9). Note that Kant himself does not deny the *existence* of all relevant unconditioned conditions. The illusion here is that we are justified in inferring the existence of 'the whole sum of conditions, and hence the absolutely unconditioned' from a given 'conditioned'.

Kant's own diagnosis of *D* as illusory gives rise to perhaps the central puzzle for his indirect proof. How is his antinomy 'unavoidable' on the assumption of realism if the dialectical cosmological syllogism, 'on which rests the entire antinomy', makes indispensable use of the illusory *D*? That syllogism is presented as follows: 'If the conditioned is given, then the whole series of all conditions for it is also given; now *objects of the senses* are given as conditioned; consequently, etc.' (A497/B525). Kant notes an ambiguous middle term: 'the major premise of the cosmological syllogism takes the conditioned in the transcendental signification of a pure category, while the minor premise takes it in the empirical signification of a concept of the understanding applied to mere appearances; consequently, there is present in it that dialectical deception called a *sophisma figurae dictionis*' (A499/B527-8). The conditioned in the major premise concerns things in themselves *qua* intelligible objects; in the minor, it picks out entities in space and time. The antinomy arises for the realist, then, insofar as entities in space and time are treated as things in themselves. But again, why cannot the realist avoid the 'unavoidable' contradictions by registering *D*'s illusory character, so simply rejecting it for things in themselves?

Proops answers as follows:

[Kant] seems to think that we know the major premise to be true; for he is prepared to assert it in his own voice. He says, for example, 'If the conditioned as well as its condition *are things in themselves*, then when the first is given not only is the regress to the second given as a problem, but the latter is thereby actually [*wirklich*] already given along with it.' (pp. 258-9)

Proops adds: 'It's an excellent question how Kant could be in a position to know that this claim is true, but there can be little doubt that he does take us to know it – a point that, in my view, deserves a higher profile in the Kant literature (p. 259n30).

This is not only an excellent but an essential question. If *D*, described as 'obviously synthetic', is knowable for real things in themselves, have we not contradicted the key doctrine that synthetic knowledge 'only arises' if understanding 'unites with' sensibility, thus is restricted to appearances (A51/B75)? As Proops himself notes, 'Kant is committed by his negative verdict on speculative dogmatic metaphysics to maintaining that the arguments for both the thesis and the antithesis *are not known to be sound*' (p. 289). Kant's reader faces a dilemma: if we do know *D* for things in themselves, why can we not hope to rehabilitate rational psychology and theology on its basis? If, on the other hand, Kant does not mean that *D* is knowable for things in themselves, what does he mean, and how are his antinomies unavoidable?

In agreeing (on p. 48) with Pedrag Cicovacki that *D* is best understood as one formulation of the Principle of Sufficient Reason, Proops implicitly rejects deflationary proposals taking it to be a version of the principle of determinacy. Such proposals are motivated in part by the desire to explain how Kant could claim to know *D*'s truth for things in themselves. Such deflations must however explain why *D*

is presented as ‘obviously synthetic’ and also how it can do the substantive work it apparently does in Kant’s arguments. Given a conditioned member of a cosmological series in a domain in which determinacy holds, we may certainly infer that the series either has or lacks a first member. In contrast to this disjunction, *D*’s inference to the unconditioned is presented as carrying explanatory force. Kant also often suggests that his explanatory unconditioned must take the form of an unconditioned first of the series (e.g. *CPrR*, 5: 48). Indeed *D* appears to do correspondingly heavy lifting in Kant’s thesis proofs. On the one hand, Kant’s antitheses proofs are also officially governed by *D*; a close reading however suggests they rest on imputed direct a priori knowledge of spatiotemporal order. (I agree that Kant’s third antithesis, for example, treats the causal principle, in Proops’s terms (p. 294), ‘as a basic principle known a priori’, rather than inferring it from PSR.) Several thesis arguments, by contrast, rest on minimally altered versions of *D*. Kant’s third antinomy thesis proof, for example, rests on a supposed ‘*Law of Nature*’ that without an unconditioned first cause of the world series, no event anywhere in the series would be ‘sufficiently determined a priori’ (A446/B474). Again, a determined cosmological series has or lacks a first member. The Principle of Determinacy does not permit us to pluck one disjunct for the sake of our thesis proofs, another for antitheses.

Rejecting such deflationary readings, Proops proposes instead that *D* may be ‘regarded as one formulation of the Principle of Sufficient Reason’. The PSR, he notes, does ‘have a claim to be a substantive principle which generates many of the claims of dogmatic speculative metaphysics’ (p. 48). He adds a caveat: ‘*D* is plausibly equivalent to the PSR only on the Kantian assumption . . . that nothing can be a condition of itself. If something could be its own condition, then the demands of the PSR might be satisfied while *D* was false’ (p. 48n20).

I would state this relationship slightly differently. Note first that ‘unconditioned first’ of a series is ambiguous, in a sense that Kant exploits, between ‘unconditioned by another series member’ and ‘unconditioned by another member *or by itself*’. Until Kant’s uncompromising rejection of self-conditioning becomes clear, Leibnizian readers would certainly be inclined to read *D*’s demand for an unconditioned first series member as a ‘formulation of PSR’. For Leibniz’s version of PSR holds that an infinite chain of cosmological conditions cannot constitute a complete sufficient reason of a subordinate series member. If one adds, however, that Kant’s unconditioned-by-another also cannot be its own condition, Leibniz, Wolff, and Baumgarten will immediately object that PSR *contradicts* Kant’s *D*. That is how Kant himself presents things: he objects that the PSR of Leibnizian predecessors excludes existence of an unconditioned. What Proops calls an ‘assumption’ that nothing can be its own condition is for Kant no mere assumption. He regards the incoherence of self-grounding as a priori and certain and as entailing the ‘obvious falsehood’ of unrestricted PSR.⁴

I suggest we must go beyond Proops’s view of Kant’s employment of *D* in his antinomies. Proops plausibly interprets the third antinomy thesis proponent as motivated by a demand for ‘comprehensibility’:

A cause’s being ‘comprehensible’ [*Begreiflich*] consists in its having a causality – that is, an operation or exercise of its power – that is not conditional on anything else. This feature is aptly described using the epistemic term

‘comprehensible’ because when a cause is comprehensible one can in principle understand how the effect comes about – how, in other words, it is possible. (p. 296)

‘Understanding how the effect comes about’ in this sense thus demands the kind of terminating series the thesis presents as necessary for an effect to be ‘sufficiently determined *a priori*’. Any cause, first or subordinate, explains its effect in a less demanding sense. What is missing from Proops’s story is Kant’s further claim that such comprehensibility is a speculative chimera, a theoretical demand that *cannot* be fulfilled. Appeal to an unconditioned first ground, he argues, inevitably fails to furnish the promised comprehensibility: ‘If the intelligibility [*Begreiflichkeit*] of a thing is to be complete, we need a first ground; however, we cannot posit a first ground through reason, and it follows there is no completely intelligible absolute positing for human reason’ (Refl 3976, 17: 372; cf. A613/B641). This complaint is closely associated with Kant’s own turn from the speculative to the moral ideas.

When Kant describes *D* as holding for things in themselves, I claim he does not mean the theoretical inference from conditioned to unconditioned is justified for real entities, but only for things in themselves conceived as ‘objects of pure understanding’ in a certain natural sense. The evidence follows his puzzling claim that when conditioned and condition are things in themselves, ‘then when the first is given not only is the regress to the second given as a problem, but the latter is thereby really already given along with it’. Kant adds the qualification: ‘... or rather, is presupposed ... Here the synthesis of the conditioned with its conditions is a synthesis of the *mere* understanding, which represents things as they are without paying attention to whether and how we might achieve acquaintance [*Kenntnis*] with them’ (A498–9/B526–7). This language echoes the Amphiboly chapter’s discussion of a natural but misleading *representation* of things in themselves through ‘mere understanding’ as subject to unrestricted conceptual intelligibility constraints Kant believes cannot govern reality.⁵ Such a reading is also supported by Kant’s claim that ‘Spinozism is all that remains’ if Transcendental Realism is endorsed (*CPrR*, 5: 101–2). As Proops rightly notes, here Kant ‘can hardly see Spinoza’s position as one that is balanced by equally powerful countervailing arguments from a Transcendentally Realist point of view’ (p. 219). The same conclusion applies to the Transcendental Aesthetic’s assertion that reality of spatiotemporal form entails its application to all existents, even God (B71–2). When Kant’s dogmatist thesis proponent argues with the help of *D*, we should conclude, he argues from a natural but misleading *representation* of things in themselves through ‘mere understanding’. Neither Kant nor the realist knows *D* to be true of *actual* ultimate existents.

4. Freedom and the indirect argument

If the realist does not know *D* to be true of existing things in themselves, we readily see why Kant should present the antithesis conclusions as true consequences of realism, in a manner wholly ‘unbalanced’ by conflicting theses. We are also relieved of the burden of explaining how the realist could know *D* to be true of actual existents. We are left however with another apparently intractable problem: saving the

supposed ‘unavoidability’ of Kant’s antinomies assuming realism. I hold that Kant’s answer here has two parts. First, he is claiming the antinomies are unavoidable on the assumption of realism for dogmatic proponents of *D*. The deeper aim here is however to *undermine D* and related speculative principles as supposed sources of insight into real things. Second, an antinomy is also in fact unavoidable *tout court*. This conclusion should be congenial to Proops, since it turns on *human freedom’s* role in Kant’s indirect argument, a role Proops seeks to uphold in the third antinomy context.

I believe Proops’s own account of this role is not quite right. The central problem is that the Dialectic’s third antinomy thesis officially establishes only ‘cosmological freedom’ or an unconditioned cause of the world. Kant does add that the thesis argument ‘removes an obstacle’ to the ‘possibility’ of human freedom in the world (A448–50/B476–8). His antithesis argues that cosmological freedom contradicts the requirements of time order. The problem is that a *necessary* conflict in realism is said to ground Kant’s indirect argument, and the Dialectic’s printed text locates this necessary conflict between spatiotemporal order and cosmological freedom, not human freedom. Any effort to make human freedom necessary to his indirect argument faces the issue that it is not essential to the Dialectic’s official third antinomy thesis, while Kant’s resolution argues only that spatiotemporal order, because merely ideal, does not exclude unconditioned causation. Now as Proops correctly registers, Kant does claim *elsewhere* that his ‘unavoidable’ antinomial conflict involves the freedom of humans.⁶ He does not regard the relevant claims as problematic, as he should, but apparently views them as restatements of the Dialectic’s official proofs and resolution.⁷

Let me conclude by sketching a different approach, summarising what I argue elsewhere. It sets out from Kant’s growing interest in the late 1760s in a *moral* deduction of various ‘unconditioneds’ mirroring his growing scepticism regarding speculative deductions of the sort exhibited in *D*. It then notes Kant’s identification of a moral analogue of *D* based on his claim that practical reason presupposes ‘its own unconditioned causation (in regard to nature), i.e. freedom, because it is aware of its moral command’ (*CPJ* §76, 5: 403). Practical reason also presupposes an unconditioned *end* – Kant’s ‘highest good’ as ‘unconditioned totality of the object of pure practical reason’ (*CPrR*, 5: 108). Texts from 1769 onwards develop moral deductions of three metaphysical ‘unconditioneds’: human freedom, God’s existence, and immortality:

Three intellectual (intelligible) [things] (noumenon) contain the unconditioned, and one can have knowledge of freedom and its laws and thereby *prove the objective reality of humanity as noumenon* in the midst of its mechanism as phenomenon. God as unconditioned necessary substance. Freedom as unconditioned causation, and immortality as personality . . . The categories applied to the intelligible can indeed ground practical-dogmatic knowledge, namely when they are directed towards freedom and determine the subject only in relation to this. (*Refl* 5552, 18: 221; cf. *CPrR*, 5: 132)

A crucial step registers Kant’s overlooked connection of these three unconditioneds of practical reason with the *theses* of his final three antinomies:

The three tasks of metaphysics: God, freedom and immortality fit on to [*passen auf*] the last three antinomies in reverse order, where simplicity, absolute causality, necessity can all be applied to the intelligible. (Refl 6212, 18: 497; cf. Refl 4757, 17: 703)

Notwithstanding Kant's rejection of *D* for real things in themselves, in short, realism still faces *unavoidable* antinomies within his own epistemology. For the unconditioned of practical reason are inconsistent with the antitheses of the last three antinomies. Human freedom is Kant's central case here, without which there is no morality, thus no deduction of his other moral unconditioned. We can now explain apparently anomalous summaries of Kant's antinomies in the *Prolegomena* and correspondence, where his 'unavoidable' conflict is said to involve the freedom of *humans*. Finally, we can explain Kant's crucial hint in introducing his theoretical ideas of reason as concepts of the unconditioned, namely that these ideas 'perhaps make possible a transition from concepts of nature to the practical and generate support for the moral ideas' (A329/B385). What does he mean? An infinite chain of conditions, Kant agrees with Leibniz, cannot satisfy reason's demand for full explanation. Reason is consequently led to posit 'first beginnings' as unconditioned explainers invoking principles such as *D*. Kant is clearly taken by the parallel with practical reason: in both cases we are led to posit a 'faculty of beginning a state spontaneously' (B561). Speculative first beginnings, however, cannot fulfil their appointed task of securing what Proops calls 'comprehensibility' of effects. We are faced instead with an antinomy of explanation: comprehensibility demands a first of the series, yet this first is itself *unbegreiflich*. This result is however in turn 'fortunate for the practical vocation of humanity' (A464/B492). Attention to such purely *speculative* conflicts, expounded in Kant's antinomies, 'strikes down the impertinent curiosity and presumptuousness of those who so far mistake the true vocation of reason that they make most of insight and knowledge just where insight and knowledge really cease, trying to pass off as further speculative interests what one should base on practical interests' (A470/B498).

In many though not all details I see this story not as controverting but developing Proops's interpretation in the *Fiery Test*. He has written a terrific book, for which we can all be thankful.

Notes

1 Ian Proops, *The Fiery Test of Critique. A Reading of Kant's Dialectic* (Oxford University Press 2021); citations throughout simply by page number. Citations from Kant's works, apart from the *Critique of Pure Reason*, cite the volume and page number of the Academy edition. Quotations from the *Critique of Pure Reason* are cited according to the standard A and B pagination for the first and second editions, respectively. Translations are my own.

2 Cf. 23: 41 (from *marginalia* to the A edition). When Kant says that the thesis and antithesis of the third and fourth antinomies 'may both be true', as Proops notes (p. 281) he is not claiming to have justified on theoretical grounds the logical or the real possibility of transcendental freedom (3rd antinomy) or God's existence (4th). He is rather simply arguing that spatiotemporal order, on his idealist proposal, does not exclude such realities.

3 This point has also been spotted recently by Alexander Buchinski.

4 'Taken in the unlimited universality in which it there stands, [the PSR] is obviously false if applied to entities; for it says there could be absolutely nothing unconditioned. To seek to avoid this embarrassing

consequence by saying of a supreme being that he does have a reason for his existence, but *it lies within himself*, leads to a contradiction' ('On a Discovery', 8: 198).

5 See also Kant's reference to the 'wholly natural illusion of common reason' arising from the mere 'logical demand . . . to assume complete premises for a given conclusion' (A500/B528; cf. A416-7/B444).

6 'Thesis: There exists in the world [*sic*] causes [*sic*] through freedom. Antithesis: There is no freedom but everything is nature' (*Prolegomena*, 4: 339). '[I]nvestigation of the existence of God, of immortality, etc, wasn't the point from which I set out, but rather the antinomy of pure reason: "The world has a beginning; It has no beginning etc. up the fourth [*sic*]: 'there is freedom in the human [*sic*], vs there is no freedom but everything in him is natural necessity'" (*Correspondence*, 12: 257).

7 See example on p. 303 (cf. p. 457): 'Kant treats the thesis position of the third antinomy as asserting specifically that 'the human will is free' (A475/B503). But again, the printed third antinomy thesis asserts no such thing, and it would be a remarkable error on Kant's part to claim it does.