Kant's Attempt to Establish a Principle of Causality
（カントによる因果律の確立の試み）

Christofer BULLSMITH

Abstract

This paper examines Kant's 1781 and 1787 attempts to establish a principle of causality. I offer an interpretation of Kant's position using a discussion of object-concepts and argue that this position avoids criticisms from Strawson and Friedman. While providing a valuable account of subjectivity and objectivity and placing useful limits on possible future theories of causality, Kant's attempts ultimately fail because he underestimates the flexibility of understanding.

Keywords

causality, objects, object-concepts, objectivity, determinism, Kant

In the B edition of the Critique of Pure Reason (1787), Kant purports to be seeking to establish the 'Principle of Succession in Time, in Accordance with the Law of Causality': 'All alterations take place in conformity with the law of the connection of cause and effect' (B232). In the A edition (1781) we have instead the 'Principle of Production': 'Everything that happens, that is, begins to be, presupposes something upon which it follows according to a rule' (A189). Since all happenings are for Kant merely alterations in substance (A187/B230-1), and the 'law of connection of cause and effect' is the Principle of Production, the formulations are equivalent (Allison 1983: 216). So Kant is arguing that any event has some cause or other. The question is then what kind of 'rule' Kant seeks to establish as holding between the cause and the event. Kant calls it 'necessary a priori', but this is presumably not logical necessity, and there is so far no clue as to whether it concerns event-types or only event-tokens, etc.

Kant's main argument proceeds from the fact that we take certain stretches of our experience to be objectively ordered — notably, when we watch a happening, such as a ship sailing down a river. Other stretches of our experience are merely subjectively ordered — notably, when we let our imagination play, or look back and forth at the coexisting (objectively simultaneous) parts of house (A 192-3/B237-8, A201/B246). There must be some way for us to distinguish between subjective and objective orderings within experience, despite the fact that all our representings are ordered in time (B234), and we cannot perceive time in itself as a metric over them (B233, A200/B245). Kant's answer is that we must think the objectivity into our experience, project a causal order onto the experienced world. For two states to be objectively ordered, 'the relation between the two states must be so thought that it is thereby determined as necessary which of them must be placed before, and which of them after, and that they cannot be placed in the reverse relation' (B234; see also A192/B237).

'Objective' is the modifier corresponding to 'object': for an ordering to be objective is just for it to appertain to successive states of an object. Since knowledge of things-in-themselves is closed to us, 'objects' are objects of experience — themselves appearances or representations (A190-1/B236). The problem is then to distinguish 'objects' from mere appearances — the general problem of which the task at hand is a special case.

> [A]ppearance, in contradistinction to the representations of apprehension, can be represented as an object distinct from them only if it stands under a rule which distinguishes it from every other apprehension and necessitates some one particular mode of connection of the manifold. The object is that in the appearance which contains the condition of this necessary rule of apprehension. (A191/B236; see also A197/B242-3.)

The object is 'that which prevents our cognitions from being haphazard or arbitrary', giving them 'that unity which constitutes the concept of an object' (A104; see also B137). So 'object' is elucidated in terms of the 'concept of an object', which is elucidated in terms of its function as a principle of unity under which an intuited manifold can be thought.

So, an object in this sense (this cup, say) is presumably a bundle of appearances united by the concept 'cup', by being related in a regular rule-governed way to one another and to my previous experience of cups (originally, appearances used in ostensively teaching me the concept). The set of associated rules will include (for example, and as for all spaced objects) rules relating the succession

of appearances of a cup with the cup's motion relative to my empirical self. Thus my freedom to move my head in relation to the cups on my table — or more specifically, to bring the bare appearances-at-a-time to vary over-time in ways which accord geometrically with their being spaced — is constitutive of my experience, and readies me to bundle these appearances as objects. This bundling brings the bare appearances-at-a-time into connection with other parts of my visual field (indeed, enables it to be differentiable into parts), my previous experience of cups (indeed, enables my experience to be of cups), and what would be forthcoming were I to change my perspective (enabling future experience of cups). So cups are objects in virtue of the concept 'cup' organising or systematizing my experience coherently (and consistently with other organising concepts).

For Kant, objectivity is therefore tied in with 'judgement' (see for example A68/B93, B141-2), implying that it goes beyond the impressions or feelings of individual believers, and stating what must be (or better, ought to be, since even all actual cognizers could err) accepted by all judging subjects. To a first approximation, to judge that an object is a cup is to (tacitly) claim that the concept 'cup' would play a role in organising or systematising the experience of anyone; judgements of experience connect to the necessary structure that attaches to the experience of any cognizer. In this way, to judge a succession to be 'objective' is to claim a universality to one's judgement. Though cognizers sometimes err, the ordering I impute is in this sense necessary.

This analysis of objectivity — its link to the universal or necessary — can be applied to causality. That is, its objectivity is elucidated in terms of the 'concept of the causal relation' (B234), which is elucidated in terms of its function as a principle of unity under which an intuited manifold can be thought. Whereas the empirical concept 'cup' has its ground in organising contingent experience, the pure concept of the understanding 'causality' (the concept of a relation between an event and some cause from which the event follows necessarily, according to a rule) is required for the possibility of objective experience (A194-5/B239-40, A201/B246-7). So given that we have objective experience, the causal relation holds a priori (A91-2/B123-4, A112-3) of all events (A195-6/B240-1). That is, '[a]ll alterations take place in conformity with the law of the connection of cause and effect' (B232) — which is what Kant was trying to establish.

It remains, however, to examine exactly what this law connects in a 'necessary relation'. Suppose the objective succession consists in the succession of state of affairs B upon state of affairs A. Strawson argues that Kant commits a 'non sequitur of numbing grossness' (1966:137) in moving from

conceptual necessity (that our perceptions be ordered 'of A', 'of B') to causal necessity (that the change from A to B is itself necessary, that 'an event of that type invariably and necessarily follows upon a condition of that type' (1966: 138)). However, this interpretation is belied by Kant's ship example, where A (the ship upstream) is clearly not to be taken as the cause of B (the ship downstream) (see Allison 1983: 229-231, 234). What falls within the scope of a causal account is the alteration of some object from state A to state B. This single 'event, as the conditioned, thus affords reliable evidence of some condition, and this condition is what determines the event' (A194/B239). *Pace* Strawson, Kant's arguments are aimed to show that in experiencing this event we must think it determined by some condition from which if follows in accordance with some universal rule.

> [T]here is an order in our representations in which the present, so far as it has come to be, refers us to some preceding state as a correlate of the event which is given; and though this correlate is, indeed, indeterminate, it none the less stands in a determining relation to the event as its consequence, connecting the event in necessary relation with itself in the time series. (A198-9/B244)

It does not follow that we actually perceived the correlate, though it is of the realm of appearances — since it is a formal condition of all perceptions that the preceding time in this sense determines the succeeding, it is also a law of 'empirical representation' that 'appearances of past time determine all existences in the succeeding time' (A199/B244). Thus specific causal laws are the appropriate subject for scientific enquiry — and in many cases we presumably can both identify the cause and formulate the rule. Consider a cue ball qua Newtonian rigid body striking an eight ball, the event of interest the acceleration of the latter; both cause (being subjected to a force F by the cue ball) and rule (Newton's second law, F=ma) can be identified.

Friedman argues (1992: 170ff.) that the strong separation of empirical causal laws from the transcendental principle of causality which Allison effects to avoid Strawson's non sequitur objection is in tension with much that Kant says in the Transcendental Analytic. Admittedly, Kant takes (some) empirical causal laws to be 'necessary and universally valid' (Kant 1783 §29: 312) and thus, as Friedman argues, of a more than merely inductive status (1992: 173). However, on my interpretation above, the separation of empirical causal laws from the transcendental principle is neither strong nor problematic. That is, the principle establishes the necessity of some empirical causal law l holding of each event b, such that l with cause a yields b. *Pace* Strawson, it is not intended to establish that a

must be whatever appearance we actually notice as preceding the event (such an appearance may be an epiphenomenon, a co-effect of the same cause, merely a preceding state, etc.), nor whether *a* and *b* under *l* are of realistically repeatable types rather than maximally specific descriptions.

Having explicated what I believe Kant is trying to establish, we can now examine to what extent he succeeds. That is, does Kant succeed in establishing that every event presupposes some cause (or other) upon which it follows according to a rule? This is best answered by beginning with what I think Kant does *not* succeed in establishing.

Walsh (1969) argues that the Analogies concern the necessary conditions for the possibility of our possessing a unitary time system (a complete ordering of events — see A176-7). An event by its nature points forward to its effect and back to its cause, tying the succession into a determinate, determined, temporal order, with any exceptions threatening the continuity of time (there would be no tying them into the objective order and thus no telling them from mere illusion). Whether this is a good interpretation of Kant or no, I do not think it is now the most charitable. For on pain of paradox given the light postulate, we do not have in this sense a unitary time system: we now know that as a matter of fact the ordering is complete only with reference to the inertial reference frame.

One way an argument concerning the necessary conditions for the possibility of (an actual) x can go wrong is by misrepresenting x. Further, even with x properly represented, it is often not possible to know with certainty what is necessary for its possibility. In the Aesthetic, Kant suggests that Euclidean geometry and intuitive arithmetic hold: that is, are true in respect of all possible objects of our experience. These two branches of mathematics are, he argues, given, or established, parts of any tenable understanding of things. Kant is wrong — while any successors must explain much that is right and richly successful with Euclidean geometry and intuitive arithmetic, our understanding of things is more flexible than Kant realised, and we have had to grow and change our mathematics. Even under the most charitable interpretation of the Second Analogy, I believe Kant goes wrong in at least this latter way. For while there is much that is right and richly successful with a principle of universal causal determinism, our tying together of appearances is more flexible than Kant realised, and we may grow away from it. We may be led to accept uncaused events (some quantum events may serve as actual examples), creation and annihilation (virtual particle pairs, the first and last moments of time, and so-called 'inflationary' periods of spacetime expansion may serve as actual examples), or backwards causation (if, for example, mysteriously correlated events are constantly followed in time

by a third event linked to each of the first two, or if we accept Feynman's theory of positrons). Kant's unitary system of time requires some one constant substance whose alterations are all fully determined — we might visualise this as being like a long single variously-coloured strand of string. It seems credible that a unitary system of time (spacetime) could instead be subserved by a variety of semipermanent pseudo-substances undergoing largely-rule-governed transformations — a rope of many fibres, no single strand necessarily running the whole length, with each fibre beginning and ending outside the rope. In other words, it could be a brute fact about the world that some events are not fully systematisable in a mathematical physics — the physics forever incomplete, our predictive and explanatory power limited in certain respects.

To apply these thoughts to Kant's argument, I believe that the truth of 'there is an object here' does imply, but only contingently, the truth of a particular raft of counterfactuals (governed by a particular set of rules, as per the discussion above of cups). We can, I believe coherently, imagine a world (perhaps, as above, it is the actual) wherein objects require the holding of only some of the counterfactuals, of only statistical truths or approximations, of only the majority of some particular set of rules, of some set of rules only under certain conditions, etc. That is, Kant claims that (entirely) substantial (completely) deterministic objects are necessary conditions for possible experience; but while perhaps our experience is in fact underwritten by such objects (object-concepts), this is not a necessary condition for possible experience.

This is not to conclude that Kant establishes nothing of value. The Second Analogy establishes that much of what happens (at macro-level, locally in spacetime — and perhaps with further qualifications) presupposes something upon which it follows according to a rule encoding a high degree of determinism. If there were in general only chaos in stuff, we would not be able to differentiate between subjective and objective orderings. This is a special case of the general argument in the Transcendental Deduction for the transcendental affinity of the manifold as a necessary condition for the possibility of experience (A112ff., A121ff.). For example:

> [i]f cinnabar were sometimes red, sometimes black, sometimes light, sometimes heavy, if a man changed sometimes into this and sometimes into that animal form, my empirical imagination would never find opportunity when representing red colour to bring to mind heavy cinnabar. (A100-101.)

However, *pace* Kant, the insight that the world must be structured enough for the possibility of cognition does not imply universal causal determinism. We can expect there to be much that is right and richly successful with the principle of universal causal determinism: that the everyday objects we interact with alter in accordance with rules encoding high degrees of determinism, and that any successor theory will explain this (just as we can expect any successor account of physical geometry to explain what is successful about Euclidean geometry at everyday scales). Still, a high degree of determinism in everyday objects does not entail that '[a]ll alterations take place in conformity with the law of the connection of cause and effect' (B232). If the world is fully determined, this is an empirical scientific feature and not a transcendental necessity.

That said, correcting Kant's argument by noting that causality could be in fact be probabilistic (etc.) feels a little like burying the lead. The insight that the subjective and the objective are distinguished not through being different in our experience, but rather via connection to expected agreement between idealized cognizers (A68/B93, B233-4, B141-2) is tremendously powerful and admits of application in a range of fields. As it happens, Kant is over-confident about the universal applicability of Euclidean geometry, intuitive arithmetic, and causally deterministic accounts of physical systems, but this does not damage his insight regarding objectivity. Even idealized cognizers will presumably agree that there is much that is (objectively) right about Euclidean geometry, albeit within limits. So where Kant argues that all possible cognizers must agree with some judgement, we are now bound to add qualifications about relative motion or scale, and to be open to the possibility of more qualifications. Kant was insightful on the distinction between subjectivity and objectivity, and right that there is at least much that is right with our intuitive heuristics about geometry, mathematics, and causation; but the world is stranger, and our understanding more flexible, than Kant allowed.

References

Allison, H.E. 1983. *Kant's Transcendental Idealism: an interpretation and defense* (New Haven: Yale University Press).

Friedman, M. 1992. 'Causal laws and the foundations of natural science' in Guyer, P. (ed.) *The Cambridge Companion to Kant* (Cambridge: Cambridge University Press): 161-199.

Kant, I. [1781, 1787] 1929. *Critique of Pure Reason,* Norman Kemp Smith (trans.) (London: MacMillan).

Kant, I. [1783] 1977. *Prolegomena To Any Future Metaphysics That Will Be Able To Come Forward As Science,* James W. Ellington (trans.) (Indianapolis, Hackett).

Strawson, P.F. 1966. *The Bounds of Sense: An Essay on Kant's Critique of Pure Reason* (London: Methuen).

Walsh, W.H. 1969. 'Kant on the perception of Time' in Penelhum, T. & MacIntosh, J.J. (eds.) *The First Critique: Reflections on Kant's Critique of Pure Reason* (California: Wadsworth): 70-88.