Could the Universe Cause itself to Exist?

WILLIAM F. VALLICELLA

This paper rests on two main assumptions, both of which I think are quite reasonable. The first is that the universe began to exist; the second is that nothing can begin to exist without a cause. Given these two assumptions, the one from current big bang cosmology, the other from a plausible metaphysics, the question I would like to pursue is whether or not the universe could have begun to exist from an *internal* cause. An example of an *external* cause of the universe would of course be God as classically conceived. So to ask whether the universe could have had an internal cause is to ask whether it could have been caused to exist by some part or phase of itself. If we plausibly assume the universe to be nothing above and beyond the succession of its states, then the question is whether or not it is coherent to suppose that in a finitely old universe every state has a causal explanation in terms of preceding states. If this is a coherent supposition, then every version of the cosmological argument for the existence of God would appear to be in serious trouble.

Our question, then, is whether the following trio of propositions is consistent:

1. The universe began to exist.
2. The universe's beginning to exist was caused.
3. The universe's beginning to exist was internally caused.

I

We may approach this question by recalling a famous argument Hume gave in rebuttal of the cosmological argument.[1] He argued in effect that if the universe always existed, then one may coherently suppose that each state of the universe has an internal cause in that each state is caused by a preceding one. In an infinitely old universe it is clear that there will be no lack of preceding states to serve as causes for any state you care to choose. The cardinality of the set of such states will be at least aleph-zero. So if each state is caused by a preceding one, and if the universe is just the sum-total of its states, then the universe is causally self-explanatory and there is no need to

[1] David Hume, *Dialogues on Natural Religion*, Part IX.

https://doi.org/10.1017/S003181910000070X Published online by Cambridge University Press

posit an external cause such as God. What is interesting to note, however, is that a variant of this Humean argument seems to work even on the assumption that the universe had a beginning. Of course, whether either argument *really* works is a further question.

The variant argument is animated by the insight that if the universe did not always exist, it does not follow that there was a temporally first state or event which would perforce be internally uncaused. What follows is merely that there was a first or earliest temporal interval, a first hour say, or a first nanosecond. Consider any such interval, and make two assumptions. (i) Assume that the interval is half-open in the earlier direction, and (ii) assume that time is either dense or continuous as opposed to discrete. If time is dense, then its moments are packed together like the rational numbers; if continuous, then like the reals. To say that the earliest temporal interval is half-open in the earlier direction is to say that there is no moment in it corresponding to zero in the real line interval $0 < x \leq 1$. And if time is either dense or continuous, then for any two instants a, b, there is a third instant c such that c is preceded by a and followed by b. Thus between any two instants there is a third, which of course implies that between any two there are infinitely many. Equivalently, no instant has an immediately preceding, or an immediately succeeding, instant.

If no instant is such that there is an immediately preceding one, then every instant has infinitely many preceding instants. And if we make the further natural but unobvious assumption that (iii) causation is as dense or continuous as time itself, then every state of the universe will have infinitely many causal antecedents internal to the universe. This will be the case even if the universe has a finite past. Given that the universe is just the totality of its states, it follows that the universe has an internal cause of its existence. The universe, even if finitely old, will be *causa sui* in the sense that each of its states is caused by an earlier one.

Quentin Smith, who endorses this intriguing argument, sums it up as follows.

> We can characterize the universe as a continuum of successive, instantaneous states. This continuum of instantaneous states begins to exist in the sense that there is an earliest half-open interval of each length (a first hour, a first minute, a first second, etc.). The continuum's beginning to exist is caused in the sense that each instantaneous state that belongs to the continuum is caused by some earlier instantaneous states that also belong to the continuum.[2]

[2] Quentin Smith, 'The Reason the Universe Exists is that it Caused itself to Exist', *Philosophy* **74** (1999), 583–584.

https://doi.org/10.1017/S003181910000070X Published online by Cambridge University Press

The conclusion Smith would have us draw, of course, is that the universe caused itself to exist. And I believe he takes this to exclude the possibility that the universe also has an external cause: it cannot have two separate sufficient reasons for its existence on pain of causal overdetermination. If the universe caused itself to exist, then an external creator is not only not needed, but impossible. Contrapositively, if God is so much as possible, then the universe cannot be self-causing.

The overall argument, however, strikes me as fallacious, although instructively so.[3]

The main difficulty is that it appears to prove entirely too much. Granting that the universe may be characterized as a continuum of successive, instantaneous states, this is also true of such rather smaller objects as Smith's life.[4] It too is a continuum of successive states. And it too can be viewed as half-open in the earlier direction. One way to do this is to reckon the moment of Smith's transition from nonexistence to existence—call it time *t*—as the last moment of the period of his nonexistence.[5] The period that follows, that of his existence, will then necessarily be half-open in the earlier direction. This is because of the continuity of time, which excludes there being a time *t*' immediately following *t*. If one objects that this is arbitrary, and insists on reckoning the moment of transition as the first moment of Smith's existence, then I will simply invite the reader to 'subtract' the first moment from Smith's life. The remainder, call it Smith's truncated life, will then be such that its earliest interval is half-open in the earlier direction. There will be continuum-many instantaneous states in this interval each of which will have causes within it. No matter how thin you slice the earliest interval in Smith's truncated life, it will always contain plenty of instantaneous states—2-to-the-aleph-zero to be exact—such that *no* state is internally uncaused. And if no state is internally uncaused, then every state is internally caused. So doesn't the above argument show that the beginning of Smith's existence (or else Smith's trun-

[3] After working out the essentials of the critique that follows, I came across an excellent article that argues along roughly similar lines, though without reference to Smith. Cf. Alexander R. Pruss, 'The Hume-Edwards Principle and the Cosmological Argument', *International Journal for the Philosophy of Religion* 43 (1998), 149–165.

[4] Perhaps a better example would be a segment of the entire universe from say August 1952 to August 2000.

[5] By the Law of Excluded Middle, *t* must belong either to the period of Smith's nonexistence or the period of his existence. By the Law of Noncontradiction, *t* cannot belong to both periods. So either *t* is the last moment of his nonexistence, or the first moment of his existence.

https://doi.org/10.1017/S003181910000070X Published online by Cambridge University Press

cated existence) has no need of an external cause and that his life (or else his truncated life) caused itself? The better to appreciate this, substitute 'Smith's life' for 'the universe' in the above quotation and leave everything else the same. The result is as follows:

> We can characterize Smith's life as a continuum of successive, instantaneous states. This continuum of instantaneous states begins to exist in the sense that there is an earliest half-open interval of each length (a first hour, a first minute, a first second, etc.). The continuum's beginning to exist is caused in the sense that each instantaneous state that belongs to the continuum is caused by some earlier instantaneous states that also belong to the continuum.

So if the original argument is valid, the parody argument is also valid: after all, the two arguments have the same form, and validity is a matter of form. And if the original argument is sound, then so is the parody. The only difference between the two is in the first premise of each. But it seems that the initial premises are either both true, or else both false. If it is true that the universe is a continuum of successive, instantaneous states, then it is also true that Smith's life is a continuum of such states. But it is as obvious as anything that the parody argument is unsound, issuing as it does in a false conclusion: we know that the beginning of Smith's life has an external cause in the conjugal activities of his parents, and since his life's beginning has an external cause, it cannot have an internal cause or causes on pain of causal overdetermination. I conclude that the original argument is also unsound, which is to say that it is either invalid in point of logical form, or possesses one or more false premises, or both. It is perhaps not unnecessary to point out that from the unsoundness of Smith's argument one cannot infer that the universe has an external cause: for all that has been shown so far, it might have neither an internal nor an external cause.

There are of course important differences between the universe and Smith's life. One is that there were no times or events prior to the beginning of the universe, it being the beginning of space-time itself; but there were times and events prior to the beginning of Smith's life. This difference, however, obvious as it is, cannot be invoked to save Smith's argument. For the upshot of the argument is that for *any* continuum C which is half-open in the earlier direction, any times and events that there might be external to C are simply irrelevant to the causal explanation of C. This holds both for the universe and any processes within it, regardless of the fact, if it is a fact, that the universe has nothing of a concrete nature external to it. This is just another way of making the point that Smith's original argument proves too much.

https://doi.org/10.1017/S003181910000070X Published online by Cambridge University Press

Having shown that Smith's argument is fallacious, I will now try to isolate the exact fallacy Smith's argument commits. I think the fallacy consists in an equivocation on 'begins to exist' as it functions in the following two sentences which are close paraphrases of Smith's own words:

4. Continuum U begins to exist in the sense that there is an earliest half-open interval of each length.
5. That U begins to exist is caused in the sense that each instantaneous state of U is caused by earlier states of U.

There is considerable conceptual slippage between (4) and (5). If U's beginning to exist is its having an earliest half-open interval, then it is not at all clear how a beginning of existence in this sense can be caused by the fact that each instantaneous state of U is caused by earlier states of U. At most, what the latter explains is why U exists; it does not explain why U *begins* to exist. For each state of U could be caused by earlier states of U even if U had *no* earliest interval, and so was never caused to *begin* to exist. Whether the universe has a finite or an infinite past, it could still be the case that each of its states is caused by earlier states. So each state's being caused by earlier states cannot be what makes the difference between the universe's having a beginning and its having no beginning. Hence each state's being caused by earlier states cannot be invoked to explain why the universe began to exist. The question, Why did the universe begin to exist? cannot be answered by saying that each state of the universe is caused by preceding states; for this could be true even if the universe always existed.

Thus Smith equivocates on 'begins to exist' in the passage quoted. He first uses it to mean *begins*; he then uses it to mean *exists*. These senses are clearly different since a thing can exist without beginning to exist. An infinitely old universe exists without beginning to exist. The confusion is easy to succumb to, however, since a thing cannot begin without beginning to exist.

II

Having refuted the view that a finitely old universe has the resources to cause its own existence, let me now tease out an implication of this refutation. The soundness of the refutation implies the falsity of a key assumption that drives both Smith's argument, and its Humean prototype. This assumption is that a succession of states is a logically supervenient entity in that nothing more is required for its existence than the existence of its members. It is

https://doi.org/10.1017/S003181910000070X Published online by Cambridge University Press

this assumption that validates the move from 'Every state of the universe has an internal cause' to 'The universe itself has an internal cause.' To review what is wrong with this supervenience assumption, consider a second, albeit similar, example. My cat Zeno starts a ball rolling across the floor. The rolling is a causal process with continuum-many instantaneous states, and it can be viewed as having no first state. For one may reckon the moment of the ball's transition from rest to motion as the last moment of the ball's being at rest, which then entails, given the continuity of time, that there is no first state of the ball's being in motion. Arguably, we *must* view the matter in this way: otherwise the ball would suddenly be moving at some positive velocity without having accelerated to this velocity from velocity zero through continuum-many instants.

And yet there is an external cause of motion, Zeno's pushing the ball. Somehow the external cause has to 'hook onto' the rolling-ball process. How? The last instantaneous state of the pushing cannot hook onto the first state of the rolling-ball process, for there is no first state. This suggests that the last state of the pushing connects to an entire causal sequence, namely the sequence of rolling-ball states. Thus a whole sequence would somehow be the effect of a cause. If so, the sequence or succession would not be logically supervenient upon its members, an 'ontological free lunch,'[6] but something in addition to its members. We would thus have a counterexample to Hume's principle that a collection is explained when its members are. If something like this could be worked out in detail —clearly not a task for the present occasion!—then the way would be clear for the theist to argue that the entire universe is a sequence of states that is the effect of a cause.

The argument might go like this:

6. Ordinary causal processes within the universe (e.g., rolling balls) are not causally self-explanatory: they have external causes.
7. The external cause of a continuous process cannot 'operate upon' any member (any instantaneous state) thereof, since each state has a sufficient cause in a preceding state or states within the process, and there cannot be causal overdetermination of any given state.

Therefore,

8. The external cause must 'operate upon' the entire causal process or sequence, which implies that causal processes have causal properties above and beyond those of their members. This undercuts the Humean principle that to explain a collection it suffices to explain its members.

[6] To borrow a phrase from D. M. Armstrong.

https://doi.org/10.1017/S003181910000070X Published online by Cambridge University Press

Therefore,

9. The Humean principle gives us no reason to think that the universe cannot have an external cause. Note that this so whether the universe is finitely or infinitely old. For what the counterexamples show is that there are infinite collections which cannot be explained by explaining their members.

Therefore,

10. The refutation of Smith's argument also applies to Hume's original argument.

III

But not only does the refutation of Smith's argument apply to the Humean prototype, it also applies *mutatis mutandis* to two other arguments Smith gives. As before, he wants to show that a universe can begin to exist from an internal cause. He gives two 'causal loop' arguments. In the first, he has us consider a causal loop in which there is mutual simultaneous causation. At $t = 0$, the universe consists of three states, a, b, and c, where a causes b, b causes c, and c causes a, with all of this causing taking place instantaneously. Smith concludes, 'If the universe at t = 0 is a, b, and c, and a, b, and c are each caused to begin to exist by something internal to the universe, it follows that the universe is caused to begin to exist, but not by anything external to the universe.'[7]

But this follows only assuming Hume's principle whereby explanation of the members of a collection suffices to explain the collection. But it is precisely this principle that is refuted by the 'Smith's life' and 'rolling ball' counterexamples given above. For what they show is that there are collections which cannot be explained by explaining their members. In the case of the rolling ball, each instantaneous state has a cause internal to the sequence of states, but it surely does not follow that the sequence as a whole has no external cause.

The second 'causal loop' argument involves (nonsimultaneous) backward causation.[8] The details, which involve Kurt Goedel's extraction of the possibility of time-travel from Einstein's General Theory of Relativity, etc., need not detain us. The essence of the argument is that the initial state of the universe is backwardly caused by a later state. But even supposing that this is possible, one

[7] Smith, op. cit., 581.
[8] Smith, op. cit., 584–585.

https://doi.org/10.1017/S003181910000070X Published online by Cambridge University Press

cannot arrive at the conclusion that such a universe causes itself to exist without reaching for Hume's principle. If *a* causes *b*, and *b* backwardly causes *a*, then both are caused. But we cannot infer that the *pair* is caused without assuming Hume's principle. But that principle was shown to be false by the 'Smith's life' and 'rolling ball' counterexamples.

Thus all three of Smith's arguments, as well Hume's original argument, are defeated by these counterexamples.

IV

Let me conclude by returning to the question of causal overdetermination. Smith assumes that if the universe has an internal cause, in virtue of each of its state's having an internal cause, then it cannot also have an external cause on pain of causal overdetermination. But this isn't obvious, even if we acquiesce in the Humean principle that to explain a collection it suffices to explain its members. A theist could grant that the universe is such that each state has an internal cause while also maintaining that it has an external cause. There are theories that allow this without overdetermination. Suppose our theist adopts a regularity theory of causation. On such a theory, what makes an event-sequence causal is its instantiation of a regularity. As I have argued in mind-numbing detail elsewhere,[9] on no such theory does the cause *produce* or bring into existence, the effect. For on a regularity theory, what you have out in the world is simply contiguous succession, just one event following another, and thus no producing of any event by any other. On this theory, what makes e1 the cause of e2 is simply the fact that e1 and e2 are tokens of event-types F and G which are related by a certain defeasible generalization, namely, 'Whenever an F-type event occurs, it is followed by a G-type event.' Thus no such theory can yield a causal explanation of the *existence* of events or *a fortiori* of the existence of the universe. So a theist who is a regularity theorist could hold that while every state of the universe has a causal explanation in terms of earlier events, the existence of these states and thus of the universe is explained by God's creative activity. There is no causal overdetermination here since although God is a cause of existence, Humean causes are not causes of existence.

Thus if Smith is to make good his claim that a self-explanatory universe excludes an external cause, he must specify a theory of nat-

[9] William F. Vallicella, 'The Hume-Edwards Objection to the Cosmological Argument', *Journal of Philosophical Research*, vol. XXII (1997), 423–443.

https://doi.org/10.1017/S003181910000070X Published online by Cambridge University Press

ural causation according to which causes bring into existence their effects. It is invalid to infer straightaway that the universe cannot have an external cause just because each of its states has an internal cause. But if he adds to his premises a theory of natural causation that precludes the possibility of an external cause, then he begs the question against the theist, who of course will adopt a theory of natural causation that does not preclude the possibility of an external cause.[10]

V

I have made three main points. The first was that Smith's argument for a finitely old, but causally self-explanatory, universe fails from probative overkill: if sound, it also shows that all manner of paltry event-sequences are causally self-explanatory. The second point was that the refutation of Smith's argument extends to Hume's prototype, as well as to Smith's two 'causal loop' arguments. The third point was that, even if Hume's principle were true, Smith's argument could not succeed without the aid of a theory of causation according to which causation is production (causation of existence).[11]

Gold Canyon, Arizona

[10] For more on my debate with Smith over God and causation, see Quentin Smith, 'Causation and the Logical Impossibility of a Divine Cause', *Philosophical Topics* 24 (1996), 169–191, and William F. Vallicella, 'God, Causation and Occasionalism', *Religious Studies* 35 (1999), 3–18.

[11] I thank Quentin Smith for comments on the penultimate draft. Send correspondence to billvallicella@compuserve.com.

https://doi.org/10.1017/S003181910000070X Published online by Cambridge University Press