

A New Grandfather Paradox?

THEODORE SIDER

Philosophy and Phenomenological Research 57 (1997): 139–144

In an article in *Scientific American* (March 1994, pp. 68–74) entitled “The Quantum Physics of Time Travel”, Oxford physicist David Deutsch and Oxford philosopher Michael Lockwood give a defense of the physical possibility of time travel based on the “Many Worlds” interpretation of quantum mechanics. This positive view of theirs is not my concern, however—I want to quarrel with their argument that time travel cannot be accommodated in any other way.¹

The best way to spell out the traditional “grandfather paradox” that appears to threaten the possibility of time travel involves the notion of ability, or personal possibility, or free will. An example of David Lewis’s: Tim travels back in time with the intent to kill his grandfather.² Let us fix the case as one in which Tim in fact will not kill Grandfather; still, it seems that he *can* kill Grandfather because he is a good shot, has a gun, and is alone with Grandfather at close range. As Lewis says, Tim “has what it takes” to kill Grandfather. However, it is also compelling that Tim *cannot* kill Grandfather, because if Grandfather had been killed in his youth, Tim would not have existed to kill him.

It is important to realize that the paradox essentially involves the notion of ability. No inconsistency results from supposing that Tim *does not* kill Grandfather. As for the case in which Tim *does* kill Grandfather, there are various possibilities. We could tell a consistent time travel story in which Tim kills Grandfather, but Grandfather is miraculously resurrected. Or one in which Tim kills Grandfather, but in which Grandfather has already had a child. Or one in which Tim kills Grandfather permanently, before Grandfather has any children, but in which Tim’s grandfather is someone other than Grandfather. As for the story in which Tim both kills Grandfather permanently in such a way that Grandfather has no children, and also is descended from Grandfather, this is an inconsistent time travel story; but of course the existence of *some*

¹Deutsch and Lockwood actually argue about whether “classical physics” is consistent with the possibility of time travel. It is by no means clear that “classical physics” means Newtonian physics, for they say on p. 72 that “when we refer to quantum mechanics, we mean its so-called many-universes interpretation”. Moreover, the only feature of classical physics that they use in the argument involving the “autonomy principle” (see below) is that, as they put it, “Classical physics says there is only one history” (p. 71).

²See David Lewis, “The Paradoxes of Time Travel”, in his *Philosophical Papers, Vol. 2* (Oxford, 1986): pp. 67–80.

inconsistent time travel stories does nothing to establish that time travel is impossible.

Deutsch and Lockwood dismiss the importance of this “free will” version of the grandfather paradox in the case of “classical physics” without closed timelike curves (CTCs):

...classical physics in the absence of CTCs is deterministic...This determinism alone is often held to be incompatible with free will. So time travel poses no more of a threat to free will than does classical physics itself. (p. 71)

This reason for dismissing the traditional paradox is odd. The argument that Tim cannot kill Grandfather is not simply the old argument that determinism is incompatible with free will—its motivation is quite different. Compatibilists, those who think free will is compatible with determinism, do not agree that “classical physics in the absence of CTCs” is incompatible with free will; but compatibilists may very well be persuaded that Tim cannot travel back in time and kill his Grandfather, given that in fact Tim is descended from Grandfather. And the argument is no less persuasive in the case of indeterminism: even if past events don’t determine current events, since Tim is in fact descended from Grandfather, it seems on its face that Tim cannot kill him.

In place of the free will version of the paradox, Deutsch and Lockwood have their own version, which appeals to an “autonomy principle”:

The real core of the grandfather paradox is not the violation of free will but of ...the autonomy principle. According to this principle, it is possible to create in our immediate environment any configuration of matter that the laws of physics permit locally, without reference to what the rest of the universe may be doing. (p. 71)

In their example, a time traveller, Sonia, goes back in time (say, to 1970) to meet her earlier self. She recalls in her youth having met a time traveller who claimed to be her later self, and now she is determined to go back in time to visit her earlier self and utter different things than she uttered “the first time around”:

Classical physics says there is only one history, so try as she might to do other than what history dictates, consistency requires Sonia to act out her part in it...Must we suppose, absurdly, that she is gripped by an irresistible

compulsion to utter the original words, contrary to her prior intentions to do otherwise? ... Within classical physics, the answer is yes. Something must prevent Sonia ... from deviating from what has already happened ... *according to classical physics, consistency requires the autonomy principle to fail.* (p. 71, my emphasis)

The argument seems to be the following. The configuration of matter which would involve Sonia's uttering words different from "the original words" is locally permitted by the laws of physics. So if the autonomy principle were true, it would be possible to create this configuration. But this is not possible within classical physics, and indeed in any physics without the many worlds interpretation, for in such a physics there is a single 1970, a single Sonia, and a single utterance. Thus, time travel without the many worlds interpretation of quantum mechanics contradicts the autonomy principle.

In fact, this is nothing more than the original free will version of the paradox in disguise. Recall Deutsch and Lockwood's statement of the autonomy principle:

it is *possible* to create in our immediate environment any configuration of matter that the laws of physics permit locally (p. 71, my emphasis)

Let us ask: what does 'possible' mean here? On one weak interpretation of the principle, it means 'physically possible', which we may gloss as 'true in some physically possible world'. This version is not what Deutsch and Lockwood need for their argument, for thus interpreted the autonomy principle is apparently a trivial analytic truth—"the laws of physics permit configuration X locally" presumably meaning that configuration X is embedded in some physically possible world. That this principle is not violated by the case of Sonia may be seen as follows. Let the "original" string of words Sonia uttered (i.e., the one and only string of words Sonia uttered at the relevant time) be S, and let S' be some other string of words. The laws of physics permit locally a configuration of matter consisting of Sonia uttering S', but this configuration of matter is also possible in the current sense, for there are physically possible worlds in which Sonia utters S'—worlds in which Sonia never uttered S but rather uttered S' instead.

A stronger notion of physical possibility would have the principle require that if the laws of physics permit a certain configuration locally, then it is physically possible to have that configuration coexist with the rest of the world the way it actually is. This might be thought to rule out the appeal in the

previous paragraph to possibilities where Sonia never uttered string S in the first place, for those possibilities presumably differ from the actual world at parts of spacetime other than the local area of the utterance. Sonya's intentions when entering the time machine, for example, may be different in the world where she utters S' than they are in the actual world. This reading of the principle may seem to be suggested by the end of the long quotation above: "according to the principle, it is possible to create in our immediate environment any configuration of matter that the laws of physics permit locally, *without reference to what the rest of the universe may be doing.*" But this principle has no plausibility. No one supposes that the result of patching together physically possible partial configurations of the world need always result in a physically possible total configuration. Given the actual past, the laws of nature may preclude certain outcomes while permitting others; and in time travel worlds the laws plus the future may constrain the past.

Clearly, neither interpretation is what Deutsch and Lockwood had in mind by the principle of autonomy. To get closer to what they might have meant, we should look to some things they say by way of motivation for the principle:³

When we strike a match, we do not have to worry that we might be thwarted because the configuration of the planets, say, might be inconsistent with the match being lit. Autonomy is a logical property that is highly desirable for the laws of physics to possess. For it underpins all experimental science: we typically take for granted that we *can* set up our apparatus in any configuration allowed by physical law and that the rest of the universe will take care of itself. (p. 71, my emphasis)

They say that *we can* set up our experimental apparatus in various ways allowed by physical law, and that we will not be thwarted when we try to strike a match—this suggests that 'possible' in the autonomy principle expresses some notion of ability, or possibility-for-persons, or **free will**. And the original statement of the principle, recall, reads: "it is possible to *create* in our immediate environment...". The idea, pretty clearly, is that the **ability** of humans to affect their local surroundings is in some sense independent of what goes on in remote parts of the universe.

³The match example in this quotation suggests a spatial version of the autonomy principle—that physically possible spatially local configurations can be patched together at will to result in a spatially larger configuration that is physically possible. Whether or not this principle is acceptable, it isn't what Deutsch and Lockwood need for the argument.

The principle thus interpreted needs certain qualifications. Brainwashing, injury, technological limitations, etc. set boundaries on what physically possible configurations we can bring about. I think we can circumvent these difficulties while preserving the intent of the principle as follows:

The autonomy principle restated if a person in local circumstances, *C*, is able to bring about some configuration of matter, *M*, then any other intrinsically similar person in physically possible circumstances *C'* that are *intrinsically similar* to *C* is able to bring about configuration *M*, regardless of what goes on in the universe outside of circumstances *C'*

The idea is that ability, in the sense relevant to the autonomy principle, is a purely local matter—constrained only by the “local” circumstances.

The precise meaning of this principle would depend on how “intrinsically similar” is spelled out, and how broadly we construe “local circumstances”, but the point I want to make is independent of those issues. That point, quite simply, is that the restated autonomy principle simply codifies the argument in the original free-will grandfather paradox that Tim is able to kill Grandfather because he “has what it takes”. Why is it so convincing that Tim can kill Grandfather? Because the circumstances are *intrinsically similar* to other circumstances (those in cases that don’t involve time travel, for example) in which we clearly grant Tim this ability—after all, he has a gun, he is alone with Grandfather, he is a good shot, etc. Why is it so convincing that Sonia can utter string *S'* of words? Because the circumstances are intrinsically similar to those cases where we clearly grant her this ability—after all, she has no constraints on her speech, there is nothing abnormal about her psychology (“Must we suppose, absurdly, that she is gripped by an irresistible compulsion to utter the original words, contrary to her prior intentions to do otherwise?”), etc.

The new, “real core” of the grandfather paradox, then, is nothing other than the original paradox in disguise. And this paradox has at least one satisfying resolution: that presented by David Lewis in his “The Paradoxes of Time Travel”. Lewis’s idea is that a statement attributing ability, like “Tim can kill Grandfather”, is ambiguous. The statement means “Tim’s killing Grandfather is compossible with a certain set of facts”, but the relevant set of facts can vary from one context of utterance to another. When we say that Tim can kill Grandfather because he has what it takes, we mean that his killing Grandfather is compossible with a certain set of facts that includes only relatively “local” facts

about the killing situation; when we say that Tim can't kill Grandfather because Grandfather is Tim's grandfather, we mean that Tim's killing Grandfather isn't compossible with a more inclusive set of facts that includes the fact that Grandfather survived his youth and helped produce Tim. The assertions that Tim can't, and that Tim can kill Grandfather are compatible, because the relevant facts in each case are different. Applied to Deutsch and Lockwood's version of the paradox, the idea would be that the autonomy principle (revised version) holds for one sense of 'ability', namely, a sense in which the relevant facts are all "purely local" facts, whereas the autonomy principle does not hold for another sense of 'ability' under which Sonia's inability to rebel against history and utter sequence S' of words follows from the fact that she in fact uttered sequence S. Thus, if some solution like Lewis's adequately dissolves the Grandfather paradox, we have been given no motivation to consider the more exotic interpretations of quantum mechanics to which Deutsch and Lockwood appeal.

References