

# THE PARADOXES OF TIME TRAVEL

David Lewis

**T**IME travel, I maintain, is possible. The paradoxes of time travel are oddities, not impossibilities. They prove only this much, which few would have doubted: that a possible world where time travel took place would be a most strange world, different in fundamental ways from the world we think is ours.

I shall be concerned here with the sort of time travel that is recounted in science fiction. Not all science fiction writers are clear-headed, to be sure, and inconsistent time travel stories have often been written. But some writers have thought the problems through with great care, and their stories are perfectly consistent.<sup>1</sup>

If I can defend the consistency of some science fiction stories of time travel, then I suppose parallel defenses might be given of some controversial physical hypotheses, such as the hypothesis that time is circular or the hypothesis that there are particles that travel faster than light. But I shall not explore these parallels here.

What is time travel? Inevitably, it involves a discrepancy between time and time. Any traveler departs and then arrives at his destination; the time elapsed from departure to arrival (positive, or perhaps zero) is the duration of the journey. But if he is a time traveler, the separation in time between departure and arrival does not equal the duration of his journey. He departs; he travels for an hour, let us say; then he arrives. The time he reaches is not the time one hour after his departure. It is later, if he has traveled toward the future; earlier, if he has traveled toward the past. If he has traveled far toward the past, it is earlier even than his departure. How can it be that the same two events, his departure and his arrival, are separated by two unequal amounts of time?

It is tempting to reply that there must be two independent time dimensions; that for time travel to be possible, time must be not a line but a plane.<sup>2</sup> Then a pair

of events may have two unequal separations if they are separated more in one of the time dimensions than in the other. The lives of common people occupy straight diagonal lines across the plane of time, sloping at a rate of exactly one hour of time<sub>1</sub> per hour of time<sub>2</sub>. The life of the time traveler occupies a bent path, of varying slope.

On closer inspection, however, this account seems not to give us time travel as we know it from the stories. When the traveler revisits the days of his childhood, will his playmates be there to meet him? No; he has not reached the part of the plane of time where they are. He is no longer separated from them along one of the two dimensions of time, but he is still separated from them along the other. I do not say that two-dimensional time is impossible, or that there is no way to square it with the usual conception of what time travel would be like. Nevertheless I shall say no more about two-dimensional time. Let us set it aside, and see how time travel is possible even in one-dimensional time.

The world—the time traveler's world, or ours—is a four-dimensional manifold of events. Time is one dimension of the four, like the spatial dimensions except that the prevailing laws of nature discriminate between time and the others—or rather, perhaps, between various timelike dimensions and various spacelike dimensions. (Time remains one-dimensional, since no two timelike dimensions are orthogonal.) Enduring things are timelike streaks: wholes composed of temporal parts, or *stages*, located at various times and places. Change is qualitative difference between different stages—different temporal parts—of some enduring thing, just as a “change” in scenery from east to west is a qualitative difference between the eastern and western spatial parts of the landscape. If this paper should change your mind about the

possibility of time travel, there will be a difference of opinion between two different temporal parts of you, the stage that started reading and the subsequent stage that finishes.

If change is qualitative difference between temporal parts of something, then what doesn't have temporal parts can't change. For instance, numbers can't change; nor can the events of any moment of time, since they cannot be subdivided into dissimilar temporal parts. (We have set aside the case of two-dimensional time, and hence the possibility that an event might be momentary along one time dimension but divisible along the other.) It is essential to distinguish change from "Cambridge change," which can befall anything. Even a number can "change" from being to not being the rate of exchange between pounds and dollars. Even a momentary event can "change" from being a year ago to being a year and a day ago, or from being forgotten to being remembered. But these are not genuine changes. Not just any old reversal in truth value of a time-sensitive sentence about something makes a change in the thing itself.

A time traveler, like anyone else, is a streak through the manifold of space-time, a whole composed of stages located at various times and places. But he is not a streak like other streaks. If he travels toward the past he is a zig-zag streak, doubling back on himself. If he travels toward the future, he is a stretched-out streak. And if he travels either way instantaneously, so that there are no intermediate stages between the stage that departs and the stage that arrives and his journey has zero duration, then he is a broken streak.

I asked how it could be that the same two events were separated by two unequal amounts of time, and I set aside the reply that time might have two independent dimensions. Instead I reply by distinguishing time itself, *external time* as I shall also call it, from the *personal time* of a particular time traveler: roughly, that which is measured by his wristwatch. His journey takes an hour of his personal time, let us say; his wristwatch reads an hour later at arrival than at departure. But the arrival is more than an hour after the departure in external time, if he travels toward the future; or the arrival is before the departure in external time (or less than an hour after), if he travels toward the past.

That is only rough. I do not wish to define personal time operationally, making wristwatches infallible by definition. That which is measured by my own wristwatch often disagrees with external time, yet I am no time traveler; what my misregulated wristwatch measures is neither time itself nor my personal time. Instead of an operational definition, we need a functional definition of personal time; it is that which occupies a certain role in the pattern of events that comprise the time traveler's life. If you take the stages of a common person, they manifest certain regularities with respect to external time. Properties change continuously as you go along, for the most part, and in familiar ways. First come in-

fantile stages. Last come senile ones. Memories accumulate. Food digests. Hair grows. Wristwatch hands move. If you take the stages of a time traveler instead, they do not manifest the common regularities with respect to external time. But there is one way to assign coordinates to the time traveler's stages, and one way only (apart from the arbitrary choice of a zero point), so that the regularities that hold with respect to this assignment match those that commonly hold with respect to external time. With respect to the correct assignment properties change continuously as you go along, for the most part, and in familiar ways. First come infantile stages. Last come senile ones. Memories accumulate. Food digests. Hair grows. Wristwatch hands move. The assignment of coordinates that yields this match is the time traveler's personal time. It isn't really time, but it plays the role in his life that time plays in the life of a common person. It's enough like time so that we can—with due caution—transplant our temporal vocabulary to it in discussing his affairs. We can say without contradiction, as the time traveler prepares to set out, "Soon he will be in the past." We mean that a stage of him is slightly later in his personal time, but much earlier in external time, than the stage of him that is present as we say the sentence.

We may assign locations in the time traveler's personal time not only to his stages themselves but also to the events that go on around him. Soon Caesar will die, long ago; that is, a stage slightly later in the time traveler's personal time than his present stage, but long ago in external time, is simultaneous with Caesar's death. We could even extend the assignment of personal time to events that are not part of the time traveler's life, and not simultaneous with any of his stages. If his funeral in ancient Egypt is separated from his death by three days of external time and his death is separated from his birth by three score years and ten of his personal time, then we may add the two intervals and say that his funeral follows his birth by three score years and ten and three days of *extended personal time*. Likewise a bystander might truly say, three years after the last departure of another famous time traveler, that "he may even now—if I may use the phrase—be wandering on some plesiosaurus-haunted oolitic coral reef, or beside the lonely saline seas of the Triassic Age."<sup>3</sup> If the time traveler does wander on an oolitic coral reef three years after his departure in his personal time, then it is no mistake to say with respect to his extended personal time that the wandering is taking place "even now".

We may liken intervals of external time to distances as the crow flies, and intervals of personal time to distances along a winding path. The time traveler's life is like a mountain railway. The place two miles due east of here may also be nine miles down the line, in the westbound direction. Clearly we are not dealing here with two independent dimensions. Just as distance along the railway is not a fourth spatial dimension, so a time traveler's personal time is not a second dimension of

time. How far down the line some place is depends on its location in three-dimensional space, and likewise the locations of events in personal time depend on their locations in one-dimensional external time.

Five miles down the line from here is a place where the line goes under a trestle; two miles further is a place where the line goes over a trestle; these places are one and the same. The trestle by which the line crosses over itself has two different locations along the line, five miles down from here and also seven. In the same way, an event in a time traveler's life may have more than one location in his personal time. If he doubles back toward the past, but not too far, he may be able to talk to himself. The conversation involves two of his stages, separated in his personal time but simultaneous in external time. The location of the conversation in personal time should be the location of the stage involved in it. But there are two such stages; to share the locations of both, the conversation must be assigned two different locations in personal time.

The more we extend the assignment of personal time outwards from the time traveler's stages to the surrounding events, the more will such events acquire multiple locations. It may happen also, as we have already seen, that events that are not simultaneous in external time will be assigned the same location in personal time—or rather, that at least one of the locations of one will be the same as at least one of the locations of the other. So extension must not be carried too far, lest the location of events in extended personal time lose its utility as a means of keeping track of their roles in the time traveler's history.

A time traveler who talks to himself, on the telephone perhaps, looks for all the world like two different people talking to each other. It isn't quite right to say that the whole of him is in two places at once, since neither of the two stages involved in the conversation is the whole of him, or even the whole of the part of him that is located at the (external) time of the conversation. What's true is that he, unlike the rest of us, has two different complete stages located at the same time at different places. What reason have I, then, to regard him as one person and not two? What unites his stages, including the simultaneous ones, into a single person? The problem of personal identity is especially acute if he is the sort of time traveler whose journeys are instantaneous, a broken streak consisting of several unconnected segments. Then the natural way to regard him as more than one person is to take each segment as a different person. No one of them is a time traveler, and the peculiarity of the situation comes to this: all but one of these several people vanish into thin air, all but another one appear out of thin air, and there are remarkable resemblances between one at his appearance and another at his vanishing. Why isn't that at least as good a description as the one I gave, on which the several segments are all parts of one time traveler?

I answer that what unites the stages (or segments) of a time traveler is the same sort of mental, or mostly mental, continuity and connectedness that unites anyone else. The only difference is that whereas a common person is connected and continuous with respect to external time, the time traveler is connected and continuous only with respect to his own personal time. Taking the stages in order, mental (and bodily) change is mostly gradual rather than sudden, and at no point is there sudden change in too many different respects all at once. (We can include position in external time among the respects we keep track of, if we like. It may change discontinuously with respect to personal time if not too much else changes discontinuously along with it.) Moreover, there is not too much change altogether. Plenty of traits and traces last a lifetime. Finally, the connectedness and the continuity are not accidental. They are explicable; and further, they are explained by the fact that the properties of each stage depend causally on those of the stages just before in personal time, the dependence being such as tends to keep things the same.<sup>4</sup>

To see the purpose of my final requirement of causal continuity, let us see how it excludes a case of counterfeit time travel. Fred was created out of thin air, as if in the midst of life; he lived a while, then died. He was created by a demon, and the demon had chosen at random what Fred was to be like at the moment of his creation. Much later someone else, Sam, came to resemble Fred as he was when first created. At the very moment when the resemblance became perfect, the demon destroyed Sam. Fred and Sam together are very much like a single person: a time traveler whose personal time starts at Sam's birth, goes on to Sam's destruction and Fred's creation, and goes on from there to Fred's death. Taken in this order, the stages of Fred-cum-Sam have the proper connectedness and continuity. But they lack causal continuity, so Fred-cum-Sam is not one person and not a time traveler. Perhaps it was pure coincidence that Fred at his creation and Sam at his destruction were exactly alike; then the connectedness and continuity of Fred-cum-Sam across the crucial point are accidental. Perhaps instead the demon remembered what Fred was like, guided Sam toward perfect resemblance, watched his progress, and destroyed him at the right moment. Then the connectedness and continuity of Fred-cum-Sam has a causal explanation, but of the wrong sort. Either way, Fred's first stages do not depend causally for their properties on Sam's last stages. So the case of Fred and Sam is rightly disqualified as a case of personal identity and as a case of time travel.

We might expect that when a time traveler visits the past there will be reversals of causation. You may punch his face before he leaves, causing his eye to blacken centuries ago. Indeed, travel into the past necessarily involves reversed causation. For time travel requires personal identity—he who arrives must be the same person who departed. That requires causal continuity, in

which causation runs from earlier to later stages in the order of personal time. But the orders of personal and external time disagree at some point, and there we have causation that runs from later to earlier stages in the order of external time. Elsewhere I have given an analysis of causation in terms of chains of counterfactual dependence, and I took care that my analysis would not rule out casual reversal *a priori*.<sup>5</sup> I think I can argue (but not here) that under my analysis the direction of counterfactual dependence and causation is governed by the direction of other *de facto* asymmetries of time. If so, then reversed causation and time travel are not excluded altogether, but can occur only where there are local exceptions to these asymmetries. As I said at the outset, the time traveler's world would be a most strange one.

Stranger still, if there are local—but only local—causal reversals, then there may also be causal loops: closed causal chains in which some of the causal links are normal in direction and others are reversed. (Perhaps there must be loops if there is reversal: I am not sure.) Each event on the loop has a causal explanation, being caused by events elsewhere on the loop. That is not to say that the loop as a whole is caused or explicable. It may not be. Its inexplicability is especially remarkable if it is made up of the sort of causal processes that transmit information. Recall the time traveler who talked to himself. He talked to himself about time travel, and in the course of the conversation his older self told his younger self how to build a time machine. That information was available in no other way. His older self knew how because his younger self had been told and the information had been preserved by the causal processes that constitute recording, storage, and retrieval of memory traces. His younger self knew, after the conversation, because his older self had known and the information had been preserved by the causal processes that constitute telling. But where did the information come from in the first place? Why did the whole affair happen? There is simply no answer. The parts of the loop are explicable, the whole of it is not. Strange! But not impossible, and not too different from inexplicabilities we are already inured to. Almost everyone agrees that God, or the Big Bang, or the entire infinite past of the universe, or the decay of a tritium atom, is uncaused and inexplicable. Then if these are possible, why not also the inexplicable causal loops that arise in the time travel?

I have committed a circularity in order not to talk about too much at once, and this is a good place to set it right. In explaining personal time, I presupposed that we were entitled to regard certain stages as comprising a single person. Then in explaining what united the stages into a single person, I presupposed that we were given a personal time order for them. The proper way to proceed is to define personhood and personal time simultaneously, as follows. Suppose given a pair of an aggregate of persona-stages, regarded as a candidate for personhood, and an assignment of coordinates to those

stages, regarded as a candidate for his personal time. If the stages satisfy the conditions given in my circular explanation with respect to the assignment of coordinates, then both candidates succeed: the stages do comprise a person and the assignment is his personal time.

I have argued so far that what goes on in a time travel story may be a possible pattern of events in four-dimensional space-time with no extra time dimension; that it may be correct to regard the scattered stages of the alleged time traveler as comprising a single person; and that we may legitimately assign to those stages and their surroundings a personal time order that disagrees sometimes with their order in external time. Some might concede all this, but protest that the impossibility of time travel is revealed after all when we ask not what the time traveler *does*, but what he *could do*. Could a time traveler change the past? It seems not: the events of a past moment could no more change than numbers could. Yet it seems that he would be as able as anyone to do things that would change the past if he did them. If a time traveler visiting the past both could and couldn't do something that would change it, then there cannot possibly be such a time traveler.

Consider Tim. He detests his grandfather, whose success in the munitions trade built the family fortune that paid for Tim's time machine. Tim would like nothing so much as to kill Grandfather, but alas he is too late. Grandfather died in his bed in 1957, while Tim was a young boy. But when Tim has built his time machine and traveled to 1920, suddenly he realizes that he is not too late after all. He buys a rifle; he spends long hours in target practice; he shadows Grandfather to learn the route of his daily walk to the munitions works; he rents a room along the route; and there he lurks, one winter day in 1921, rifle loaded, hate in his heart, as Grandfather walks closer, closer, . . .

Tim can kill Grandfather. He has what it takes. Conditions are perfect in every way: the best rifle money could buy, Grandfather an easy target only twenty yards away, not a breeze, door securely locked against intruders. Tim a good shot to begin with and now at the peak of training, and so on. What's to stop him? The forces of logic will not stay his hand! No powerful chaperone stands by to defend the past from interference. (To imagine such a chaperone, as some authors do, is a boring evasion, not needed to make Tim's story consistent.) In short, Tim is as much able to kill Grandfather as anyone ever is to kill anyone. Suppose that down the street another sniper, Tom, lurks waiting for another victim, Grandfather's partner. Tom is not a time traveler, but otherwise he is just like Tim: same make of rifle, same murderous intent, same everything. We can even suppose that Tom, like Tim, believes himself to be a time traveler. Someone has gone to a lot of trouble to deceive Tom into thinking so. There's no doubt that Tom can kill his victim; and Tim has everything going for him that Tom

does. By any ordinary standards of ability, Tim can kill Grandfather.

Tim cannot kill Grandfather. Grandfather lived, so to kill him would be to change the past. But the events of a past moment are not subdivisible into temporal parts and therefore cannot change. Either the events of 1921 timelessly do include Tim's killing of Grandfather, or else they timelessly don't. We may be tempted to speak of the "original" 1921 that lies in Tim's personal past, many years before his birth, in which Grandfather lived; and of the "new" 1921 in which Tim now finds himself waiting in ambush to kill Grandfather. But if we do speak so, we merely confer two names on one thing. The events of 1921 are doubly located in Tim's (extended) personal time, like the trestle on the railway, but the "original" 1921 and the "new" 1921 are one and the same. If Tim did not kill Grandfather in the "original" 1921, then if he does kill Grandfather in the "new" 1921, he must both kill and not kill Grandfather in 1921—in the one and only 1921, which is both the "new" and the "original" 1921. It is logically impossible that Tim should change the past by killing Grandfather in 1921. So Tim cannot kill Grandfather.

Not that past moments are special; no more can anyone change the present or the future. Present and future momentary events no more have temporal parts than past ones do. You cannot change a present or future event from what it was originally to what it is after you change it. What you *can* do is to change the present or the future from the unactualized way they would have been without some action of yours to the way they actually are. But that is not an actual change: not a difference between two successive actualities. And Tim can certainly do as much; he changes the past from the unactualized way it would have been without him to the one and only way it actually is. To "change" the past in this way, Tim need not do anything momentous; it is enough just to be there, however unobtrusively.

You know, of course, roughly how the story of Tim must go on if it is to be consistent: he somehow fails. Since Tim didn't kill Grandfather in the "original" 1921, consistency demands that neither does he kill Grandfather in the "new" 1921. Why not? For some commonplace reason. Perhaps some noise distracts him at the last moment, perhaps he misses despite all his target practice, perhaps his nerve fails, perhaps he even feels a pang of unaccustomed mercy. His failure by no means proves that he was not really able to kill Grandfather. We often try and fail to do what we are able to do. Success at some tasks requires not only ability but also luck, and lack of luck is not a temporary lack of ability. Suppose our other sniper, Tom, fails to kill Grandfather's partner for the same reason, whatever it is, that Tim fails to kill Grandfather. It does not follow that Tom was unable to. No more does it follow in Tim's case that he was unable to do what he did not succeed in doing.

We have this seeming contradiction: "*Tim doesn't, but can, because he has what it takes*" versus "*Tim doesn't, and*

*can't, because it's logically impossible to change the past.*" I reply that there is no contradiction. Both conclusions are true, and for the reasons given. They are compatible because "can" is equivocal.

To say that something can happen means that its happening is compossible with certain facts. *Which* facts? That is determined, but sometimes not determined well enough, by context. An ape can't speak a human language—say, Finnish—but I can. Facts about the anatomy and operation of the ape's larynx and nervous system are not compossible with his speaking Finnish. The corresponding facts about my larynx and nervous system are compossible with my speaking Finnish. But don't take me along to Helsinki as your interpreter: I can't speak Finnish. My speaking Finnish is compossible with the facts considered so far, but not with further facts about my lack of training. What I can do, relative to one set of facts, I cannot do, relative to another, more inclusive, set. Whenever the context leaves it open which facts are to count as relevant, it is possible to equivocate about whether I can speak Finnish. It is likewise possible to equivocate about whether it is possible for me to speak Finnish, or whether I am able to, or whether I have the ability or capacity or power or potentiality to. Our many words for much the same thing are little help since they do not seem to correspond to different fixed delineations of the relevant facts.

Tim's killing Grandfather that day in 1921 is compossible with a fairly rich set of facts: the facts about his rifle, his skill and training, the unobstructed line of fire, the locked door and the absence of any chaperone to defend the past, and so on. Indeed it is compossible with all the facts of the sorts we would ordinarily count as relevant is saying what someone can do. It is compossible with all the facts corresponding to those we deem relevant in Tom's case. Relative to these facts, Tim can kill Grandfather. But his killing Grandfather is not compossible with another, more inclusive set of facts. There is the simple fact that Grandfather was not killed. Also there are various other facts about Grandfather's doings after 1921 and their effects: Grandfather begat Father in 1922 and Father begat Tim in 1949. Relative to these facts, Tim cannot kill Grandfather. He can and he can't, but under different delineations of the relevant facts. You can reasonably choose the narrower delineation, and say that he can; or the wider delineation, and say that he can't. But choose. What you mustn't do is waver, say in the same breath that he both can and can't, and then claim that this contradiction proves that time travel is impossible.

Exactly the same goes for Tom's parallel failure. For Tom to kill Grandfather's partner also is compossible with all facts of the sorts we ordinarily count as relevant, but not compossible with a larger set including, for instance, the fact that the intended victim lived until 1934. In Tom's case we are not puzzled. We say without hesitation that he can do it, because we see at once that the

facts that are not compossible with his success are facts about the future of the time in question and therefore not the sort of facts we count as relevant in saying what Tom can do.

In Tim's case it is harder to keep track of which facts are relevant. We are accustomed to exclude facts about the future of the time in question, but to include some facts about its past. Our standards do not apply unequivocally to the crucial facts in this special case: Tim's failure, Grandfather's survival, and his subsequent doings. If we have foremost in mind that they lie in the external future of that moment in 1921 when Tim is almost ready to shoot, then we exclude them just as we exclude the parallel facts in Tom's case. But if we have foremost in mind that they precede that moment in Tim's extended personal time, then we tend to include them. To make the latter be foremost in your mind, I chose to tell Tim's story in the order of his personal time, rather than in the order of external time. The fact of Grandfather's survival until 1957 had already been told before I got to the part of the story about Tim lurking in ambush to kill him in 1921. We must decide, if we can, whether to treat these personally past and externally future facts as if they were straightforwardly past or as if they were straightforwardly future.

Fatalists—the best of them—are philosophers who take facts we count as irrelevant in saying what someone can do, disguise them somehow as facts of a different sort that we count as relevant, and thereby argue that we can do less than we think—indeed, that there is nothing at all that we don't do but can. I am not going to vote Republican next fall. The fatalist argues that, strange to say, I not only won't but can't; for my voting Republican is not compossible with the fact that it was true already in the year 1548 that I was not going to vote Republican 428 years later. My rejoinder is that this is a fact, sure enough; however, it is an irrelevant fact about the future masquerading as a relevant fact about the past, and so should be left out of account in saying what, in any ordinary sense, I can do. We are unlikely to be fooled by the fatalist's methods of disguise in this case, or other ordinary cases. But in cases of time travel, precognition, or the like, we're on less familiar ground, so it may take less of a disguise to fool us. Also, new methods of disguise are available, thanks to the device of personal time.

Here's another bit of fatalist trickery. Tim, as he lurks, already knows that he will fail. At least he has the wherewithal to know it if he thinks, he knows it implicitly. For he remembers that Grandfather was alive when he was a boy, he knows that those who are killed are thereafter not alive, he knows (let us suppose) that he is a time traveler who has reached the same 1921 that lies in his personal past, and he ought to understand—as we do—why a time traveler cannot change the past. What is known cannot be false. So his success is not only not compossible with facts that belong to the external future

and his personal past, but also is not compossible with the present fact of his knowledge that he will fail. I reply that the fact of his foreknowledge, at the moment while he waits to shoot, is not a fact entirely about that moment. It may be divided into two parts. There is the fact that he then believes (perhaps only implicitly) that he will fail; and there is the further fact that his belief is correct, and correct not at all by accident, and hence qualifies as an item of knowledge. It is only the latter fact that is not compossible with his success, but it is only the former that is entirely about the moment in question. In calling Tim's state at that moment knowledge, not just belief, facts about personally earlier but externally later moments were smuggled into consideration.

I have argued that Tim's case and Tom's are alike, except that in Tim's case we are more tempted than usual—and with reason—to opt for a semi-fatalist mode of speech. But perhaps they differ in another way. In Tom's case, we can expect a perfectly consistent answer to the counterfactual question: what if Tom had killed Grandfather's partner? Tim's case is more difficult. If Tim had killed Grandfather, it seems offhand that contradictions would have been true. The killing both would and wouldn't have occurred. No Grandfather, no Father; no Father, no Tim; no Tim, no killing. And for good measure: no Grandfather, no family fortune; no fortune, no time machine; no time machine, no killing. So the supposition that Tim killed Grandfather seems impossible in more than the semi-fatalistic sense already granted.

If you suppose Tim to kill Grandfather and hold all the rest of his story fixed, of course you get a contradiction. But likewise if you suppose Tom to kill Grandfather's partner and hold the rest of his story fixed—including the part that told of his failure—you get a contradiction. If you make any counterfactual supposition and hold all else fixed you get a contradiction. The thing to do is rather to make the counterfactual supposition and hold all else as close to fixed as you consistently can. That procedure will yield perfectly consistent answers to the question: what if Tim had not killed Grandfather? In that case, some of the story I told would not have been true. Perhaps Tim might have been the time-traveling grandson of someone else. Perhaps he might have been the grandson of a man killed in 1921 and miraculously resurrected. Perhaps he might have been not a time traveler at all, but rather someone created out of nothing in 1920 equipped with false memories of a personal past that never was. It is hard to say what is the least revision of Tim's story to make it true that Tim kills Grandfather, but certainly the contradictory story in which the killing both does and doesn't occur is not the least revision. Hence it is false (according to the unrevised story) that if Tim had killed Grandfather then contradictions would have been true.

What difference would it make if Tim travels in branching time? Suppose that at the possible world of Tim's story the space-time manifold branches; the

branches are separated not in time, and not in space, but in some other way. Tim travels not only in time but also from one branch to another. In one branch Tim is absent from the events of 1921; Grandfather lives; Tim is born, grows up, and vanishes in his time machine. The other branch diverges from the first when Tim turns up in 1920; there Tim kills Grandfather and Grandfather leaves no descendants and no fortune; the events of the two branches differ more and more from that time on. Certainly this is a consistent story; it is a story in which Grandfather both is and isn't killed in 1921 (in the different branches); and it is a story in which Tim, by killing Grandfather, succeeds in preventing his own birth (in one of the branches). But it is not a story in which Tim's killing of Grandfather both does occur and doesn't: it simply does, though it is located in one branch and not the other. And it is not a story in which Tim changes the past. 1921 and later years contain the events of both branches, coexisting somehow without interaction. It remains true at all the personal times of Tim's life, even after the killing, that Grandfather lives in one branch and dies in the other.<sup>6</sup>

*Princeton University*  
*Received September 4, 1975*

## Notes

1. I have particularly in mind two of the time travel stories of Robert A. Heinlein: "By His Bootstraps" in R. A. Heinlein, *The Menace from Earth* (Hicksville, N.Y., 1959), and "—All You Zombies—" in R. A. Heinlein, *The Unpleasant Profession of Jonathan Hoag* (Hicksville, N.Y., 1959).
2. Account of time travel in two-dimensional time are found in Jack W. Meiland, "A Two-Dimensional Passage Model of Time for Time Travel," *Philosophical Studies*, vol. 26 (1974), pp. 153–173; and in the initial chapters of Isaac Asimov, *The End of Eternity* (Garden City, N.Y., 1955). Asimov's denouement, however, seems to require some different conception of time travel.
3. H. G. Wells, *The Time Machine, An Invention* (London, 1895), epilogue. The passage is criticized as contradictory in Donald C. Williams, "The Myth of Passage," *The Journal of Philosophy*, vol. 48 (1951), p. 463.
4. I discuss the relation between personal identity and mental connectedness and continuity at greater length in "Survival and Identity" in *The Identity of Persons*, ed. by Amelie Rorty (forthcoming).
5. "Causation," *The Journal of Philosophy*, vol. 70 (1973), pp. 556–567; the analysis relies on the analysis of counterfactuals given in my *Counterfactuals* (Oxford, 1973).
6. The present paper summarizes a series of lectures of the same title, given as the Gavin David Young Lectures in Philosophy at the University of Adelaide in July, 1971. I thank the Australian-American Educational Foundation and the American Council of Learned Societies for research support. I am grateful to many friends for comments on earlier versions of this paper; especially Philip Kitcher, William Newton-Smith, J. J. C. Smart, and Donald Williams.

**Phil 420: *Metaphysics*  
Spring 2008**

[Handout 22]

**David Lewis: *Paradox of Time Travel***

**Professor JeeLoo Liu**

**§ Main Theses:**

1. Time travel is possible under a four-dimensional space-time view.
2. The paradoxes of time travel are oddities, but not impossibilities.
3. We should regard the scattered stages of the alleged time traveler as comprising a single person.
4. We could separate personal time from external time, and we can legitimately assign to those stages and their surroundings a personal time order that disagrees sometimes with their order in external time.

\* What is time travel?

\_\_\_ Inevitably, it involves discrepancy between time and time. Any traveler departs and then arrives at his destination; the time elapsed from departure to arrival is the duration of the journey. But if he is a time traveler, the separation in time between departure and arrival does not equal the duration of his journey.

*Q*: How can it be that the same two events, his departure and his arrival, are separated by two unequal amounts of time?

**Hypothesis #1: Time as a two-dimensional plane**

1. There are two independent time dimensions; time is not a line but a plane. A pair of events may have two unequal separations if they are separated more in one of the time dimensions than in the other.
2. The lives of ordinary people occupy straight diagonal lines across the plane of time, sloping at a rate of exactly one hour of time<sub>1</sub> per hour of time<sub>2</sub>. The life of the time traveler occupies a bent path,

**Objection:**

\_\_\_ This account is inconsistent with the stories of time travel, because the time traveler would not meet his playmates when he revisits the days of his childhood. He would not have reached the part of the plane where they are.

**Hypothesis #2: Time as one-dimensional line**

**[Four-Dimensionalism]:**

1. The world is a four-dimensional manifold of events. Time is one dimension of the four.



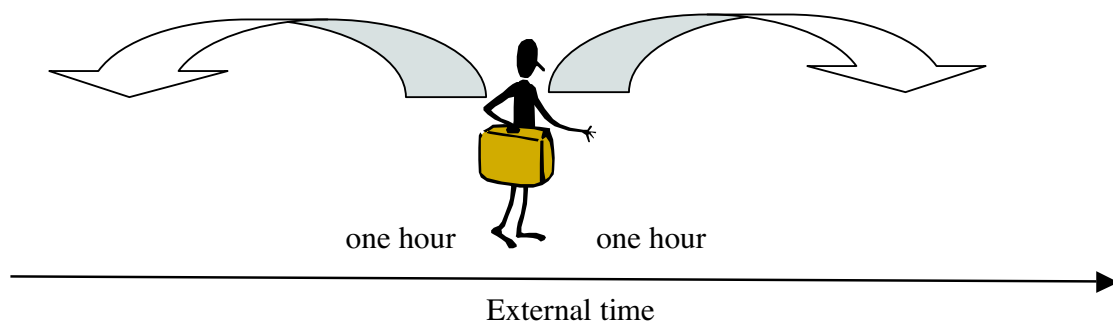
2. Enduring things are timelike streaks: wholes composed of temporal parts, or stages, located at various times and places.
3. Change is qualitative difference between different stages – different temporal parts – of some enduring thing, just as a “change” in scenery from east to west is a qualitative difference between the eastern and western spatial parts of the landscape.
4. What doesn't have temporal parts can't change. (e.g. numbers can't change; nor can the events of any moment of time.)

A time traveler, is a streak through the manifold of space-time, a whole composed of stages located at various times and places. But he is not a streak like other streaks. If he travels toward the past he is a zig-zag streak, doubling back on himself. If he travels toward the future, he is a stretched-out streak. And if he travels either way instantaneously, so that there is no intermediate stages between the stage that departs and the stage that arrives and his journey has zero duration, then he is a broken streak.

**\* external time versus personal time**

The personal time is measured by the time traveler's wristwatch. His journey takes an hour of his personal time. But the arrival time is more than an hour after the departure in external time, if he travels toward the future; or the arrival is before the departure in external time, if he travels toward the past.

We may say “soon he will be in the past,” and what we mean is that a stage of him is slightly later in his personal time, but much earlier in external time, than the stage of him that is present as we say the sentence.



If you take the stages of a common person, they manifest certain regularities with respect to external time. Properties change continuously as you go along, for the most part, and in familiar ways. But if you take the stages of a time traveler, they do not manifest the common regularities with respect to external time.

**\* The stage theory: A person has different stages located in time, but it is the same person who persists through time. Her personal identity is preserved.**

If a time traveler talks to his former self, it isn't quite right to say that the whole of him is in two places at once, since neither of the two stages involved in the conversation is the whole of him, or even the whole of the part of him is located at the (external) time of the conversation. What is true is that he, unlike the rest of us, has two different complete stages located at the same time at different places.

**Q: Is he one person with different stages, or different person continuing existence successively?**

The natural way to regard him as more than one person is to take each segment as a different person. No one of them is a time traveler, and the peculiarity of the situation comes to this: all but one of these several people vanish into thin air, all but another one appear out of thin air, and there are remarkable resemblances between one at his appearance and another at his vanishing.

Lewis's defense of the personal identity view:

\_\_\_ [He is the same person with different stages because] what unites the stages of a time traveler is the same sort of mental, or mostly mental, continuity and connectedness that unites anyone else.

\_\_\_ There is not too much change altogether. Plenty of traits and traces last a lifetime.

\_\_\_ The connectedness and the continuity are not accidental. They are explained by the fact that the properties of each stage depend causally on those of the stages just before in personal time, and this dependence tends to keep things the same.

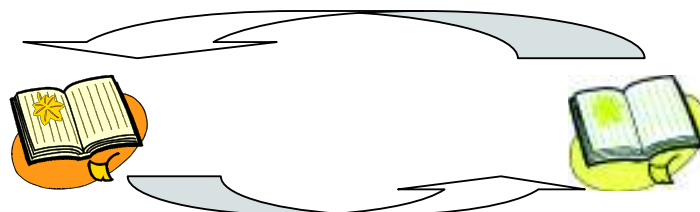
**§ Time travel, causal reversal, and causal loops**

1. Time travel requires personal identity – he who arrives must be the same person who departed.
2. Personal identity requires causal continuity, in which causation runs from earlier to later stages in the order of personal time.
3. But the orders of personal and external time disagree at some point, and there we have causation that runs from later to earlier stages in the order of external time.
4. Therefore, reversed causation and time travel are not excluded altogether.
5. But if there are local causal reversals, then there may also be causal loops: closed causal chains in which some of the causal links are normal in direction and others are reversed.

Examples:

(i) the way to build a time machine

(ii) the notebook



## § The Paradox of the Grandfather

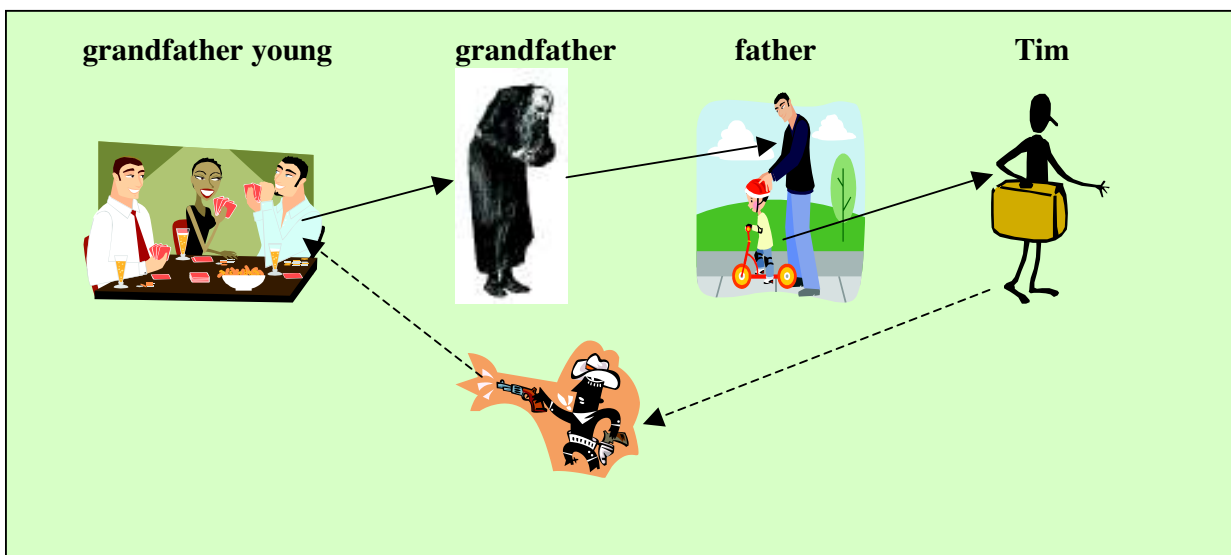
### Q: Could a time traveler change the past?

It seems not: the events of the past could not change.

It seems yes: the time traveler would be able as anyone to do things that would change the past if he did them.

If a time traveler visiting the past both could and couldn't do something that would change it, then there cannot possibly be such a time traveler.

### \* The paradox of the Grandfather:



1. By ordinary standard of ability, Tim can kill grandfather.
2. But Tim cannot kill grandfather. Grandfather lived, so to kill him would be to change the past. It is logically impossible to change the past.
3. The above is a contradiction.

### \*Lewis' diagnosis:

\_\_\_ There is no contradiction. Both conclusions are true. They are compatible because "can" is equivocal.

### [compossible]: being able to coincide with other things, consistent with

1. To say that something can happen means that its happening is *compossible with* certain facts – which facts is determined by context.
2. Whenever the context leaves it open which facts are to count as relevant, it is possible to equivocate about whether *x* can do *A*.

**Tim's case:**

\_\_\_ Time's killing Grandfather that day in 1921 is compossible with a fairly rich set of facts: the facts about his rifle, his skill and training, .... Relative to these facts, Tim *can* kill Grandfather. But his killing Grandfather is not compossible with another, more inclusive set of facts. There is a simple fact that Grandfather was not killed. Also there are various other facts about Grandfather's doings after 1921 and their effects: Grandfather begat Father in 1922 and Father begat Tim in 1949. Relative to these facts, Tim cannot kill Grandfather. He can and he can't, but under different delineations of relevant facts.

You can reasonably choose the narrower delineation, and say that he can; or the wider delineation, and say that he can't. But choose. What you mustn't do is waver, say in the same breath that he both can and he can't, and then claim that this contradiction proves that time travel is impossible.

**Q: What is Lewis' argument here?**

1. What counts as "can" = what is compossible with a set of given facts.
2. 'Tim *can* kill Grandfather' is compossible with a set of facts A.
3. 'Tim *cannot* kill Grandfather' is compossible with another set of facts B.
4.  $A \neq B$ .
5. Hence, there is no contradiction involved in Tim's traveling back in time.

**Q: Is this a good argument?****§ Conclusion**

If you suppose Tim to kill grandfather and hold all the rest of his story fixed, of course you get a contradiction. If you make *any* counterfactual supposition and hold all else fixed you get a contradiction. The thing to do is rather to make the counterfactual supposition and hold all else as close to fixed as you consistently can. That procedure will yield perfectly consistent answer to the question: what if Tim had not killed Grandfather? In that case, some of the story I told would not have been true. .... Hence it is false (according to the unrevised story) that if Tim had killed Grandfather then contradictions would have been true.