

Delayed Quantum Suicide: Assuming you will survive quantum suicide may be simplistic.

Paul Almond

<http://www.paul-almond.com> info@paul-almond.com

19 July 2011

One view of the quantum suicide thought experiment is that you should be certain of your subjective survival if the many-worlds interpretation of quantum mechanics is correct, because you never observe the branches in which you do not survive. A more general version of the thought experiment, the delayed quantum suicide thought experiment, shows problems with such a simplistic view. In this thought experiment, there is a variable time delay between the quantum event and your death occurring, ranging from a very long time delay to a negligibly short time delay, with many intermediate time delays being possible. If you think that, given the many-worlds interpretation being correct, your subjective survival is guaranteed in conventional quantum suicide, to know how short the time delay needs to be for your subjective survival to be assured, you need a way of deciding what duration of existence your brain must have for this to be counted as “observation” of your situation. This is a problem, because there is no obvious way of deciding where to draw the line – where the time delay is so short that you cannot “observe”. The simplistic view used to justify the expectation of survival in quantum suicide is inadequate for this purpose. This simplistic view was also shown to have a problem with a previous argument about “quantum brain damage” and the problems raised in each case are really special cases of the same general problem. Both the delayed quantum suicide thought experiment and the previously described “quantum brain damage” thought experiment show that, if the idea of expected survival in quantum suicide is to be viewed as viable in any way at all, a more sophisticated approach is needed for this more general case. Any expectation of survival in conventional quantum suicide would merely be a special case in such a more general view, and it is far from certain that the special case will be one of guaranteed survival. In fact, given that any general approach will have to examine the details of the situation, we might reasonably expect the details of the situation to affect the subjective probabilities of survival. We should therefore be sceptical of the idea that, given the many-worlds interpretation being correct, survival is guaranteed in the quantum suicide thought experiment, though an argument might be made that the true situation approximates this in some way, but with probabilities being influenced by various details of the situation.

1 INTRODUCTION

Quantum suicide is a thought experiment originally proposed by Hans Moravec [1] and Bruno Marchal [2], and further developed by Max Tegmark [3] that relates to the many-worlds interpretation of quantum mechanics (MWI) [4,5]. In the thought experiment, you participate in an experiment in which a sequence of iterations occurs. In each iteration, a quantum event occurs with two outcomes and a 0.5 probability of each. If one outcome occurs, the mechanism kills you. If the other outcome occurs, the mechanism does nothing: you remain alive.

As most people would understand the matter, each iteration has a 0.5 probability of killing you, so as the number of iterations increases, the chance of you surviving to the end decreases.

An alternative view, however, is that things are different if MWI is correct. According to this view, whenever the quantum event occurs, the world splits and both possible outcomes occur in different worlds: there is always a world in which you survive. That would seem to follow from MWI, but some people take a stronger position in which your subjective survival is guaranteed. In this view, if MWI is true, in every iteration, from your point of view, there is always a future in which you survive, which is experienced by you, and a future in which you die, which is not experienced by you. As you never experience the future in which you die, you will never make an observation of that outcome, and so it can be eliminated from your possible futures. You should therefore always expect to survive.

This is a controversial position, and many people think that MWI being true would not necessarily imply that you should expect to survive the quantum suicide experiment: they think that even if MWI is true, you should still think you have only a 0.5 chance of surviving in each iteration.

This article will argue that there is a possible problem with the idea that you should expect to survive quantum suicide. The *delayed quantum suicide* thought experiment will be described, which will be similar to the quantum suicide experiment, but with a difference: there is a variable time delay between the quantum-random event occurring and the mechanism killing you when the relevant outcome has occurred. The idea that you can eliminate a future if there is no conscious observer will be shown to be too simplistic in this scenario, which will in turn suggest that it is too simplistic for the “conventional” quantum suicide scenario. As was shown in our previous discussion on *quantum brain damage* [6], a more sophisticated way of approaching personal continuity is needed, and it is not guaranteed that this will assure your survival in the quantum suicide thought experiment.

2 THE DELAYED QUANTUM SUICIDE THOUGHT EXPERIMENT

The delayed quantum suicide thought experiment has some similarity with the conventional quantum suicide thought experiment.

As with conventional quantum suicide, a sequence of iterations occurs. In each iteration, one of two random outcomes is generated by some quantum device, and these outcomes will be called “1” and “0”. If the outcome is “1”, a signal is sent to some execution device which automatically kills you on receipt of this signal.

With conventional quantum suicide, the execution device is supposed to kill you as quickly as possible after the relevant quantum outcome has occurred. Here, however, there is a difference. After an outcome of “1” there is a time delay before the signal is sent to the execution device to tell it to kill you. The time delay is adjustable up or down by a dial. You can make the time delay very short, or you can make it very long: you can have a time delay of a nanosecond if you like, or you can have a time

delay of one year – if you are prepared to wait that long. You can never have a time delay which is *absolutely* zero, because any system that could be made would actually take time to work. We will assume, however, that the time delay can be made as short as you want as long it is not actually zero.

There is one complication. Once the signal has been sent to the execution device, telling it to kill you, this signal will take time to propagate to the execution device. The execution device will take further time to kill you. For example, if the execution device is a gun, aimed at your head, and activated by an electrical signal sent along a cable, there will be time taken for the signal to travel along the cable, further time for the firing pin to strike a cartridge’s percussion cap, further time for burning to spread through the gunpowder, further time for the bullet to accelerate down the barrel, further time for it to reach your head and further time for it to penetrate your skull and do enough damage to mean that you are not an observer. For simplicity, we will assume that all these events take a negligible amount of time: once the signal has been sent to the execution device, after the time delay set by the dial has expired, we will assume that you are killed, for all practical purposes, immediately.

That is what happens in a single iteration. After each iteration, a similar process occurs in the next one, and so on.

With the conventional quantum suicide thought experiment, the question posed is as follows:

If MWI is correct, do you subjectively expect to survive?

Here, the question with the delayed quantum suicide scenario:

If MWI is correct, for which settings of the time-delay dial, if any, do you subjectively expect to survive?

3 TRYING TO ANSWER THE THOUGHT EXPERIMENT'S QUESTION

3.1 If you do *not* expect to survive conventional quantum suicide

An obvious view to take, for many people, is that you should not expect to survive quantum suicide, even if MWI is true – that you should view yourself as having a 0.5 probability of dying. If you take this view, the delayed quantum suicide thought experiment should cause you little difficulty. You can simply say that one outcome is as likely as any other, irrespective of the time delay. There is a 0.5 probability of an outcome of “1”, meaning you ultimately die – no matter how long it takes – and a 0.5 probability of an outcome of “0”, meaning you survive.

3.2 If you *do* expect to survive conventional quantum suicide

If you expect to survive quantum suicide if MWI is correct, according to the simple idea that you only experience those branches in which you continue to exist as a conscious observer, the thought experiment causes you more of a problem. Most people who take this view seem to imagine a version of quantum suicide in which, if the quantum outcome is the one in which you die, you die very quickly indeed. You can never die instantly – but you die as quickly as possible. The idea seems to be that you need to die before you have any chance to “make an observation” that you are on the branch in which you are going to die. How quickly does death need to occur, though?

This is the issue raised by the delayed quantum suicide thought experiment. You need to work out what settings of the time-delay dial satisfy the condition of causing your death quickly enough. How do you do this?

You cannot reasonably say that there needs to be no time delay at all – and the time-delay dial in the thought experiment cannot be set to zero. To have any expectation of surviving in conventional quantum suicide, you must already be accepting that *some* delay – even if it has to be a very small one – is acceptable. Suppose you start with the time-delay dial set with a *very* small value, say 10^{-20} seconds. Most people who

would expect to survive conventional quantum suicide would probably find that time delay acceptable. Let us now imagine moving the dial to increase the time delay. How long can the time delay become before the rules of the game say that “it is too long” and that the branch on which you die has become a future which you can experience as a conscious observer? Could the time delay be 10^{-9} seconds? 0.01 seconds? 1 second? 10 seconds? Could it be increased to an hour? If you set the dial to any time delay and think that that is acceptable, the case can always be made for saying that the time delay can be increased just a bit further without substantially changing the situation. We could therefore end up with a situation in which any time delay – no matter how long – even one of 10 years – is acceptable.

Similarly, you can start at the other end of the scale, with the time delay set to its maximum value. We have not specified what the maximum time delay setting is, but you can assume it can be any large value you want. Let us imagine that it is 24 hours. With a time delay of 24 hours, few people would think there was any assurance of subjective survival: they would think that there was a 0.5 chance of finding themselves in the branch where the execution mechanism is going to activate in 24 hours, and a 0.5 chance of finding themselves in the other branch. But what about a time delay of 23 hours, 59 minutes and 59.9999 seconds? It should be hard to see how we cannot make a decrease of 0.0001 seconds in the time delay without substantially changing things. However, if we extend such reasoning to its limit we end up with the result that any time delay at all – no matter how short – means that your survival is not subjectively assured.

There does not seem to be any point at which an obvious line is crossed and the time delay is short enough to mean that you do not subjectively enter a branch. Why? What would be so special about this particular time delay? The simple idea that only the branches where you continue as a conscious observer should be counted is inadequate for dealing with this situation.

4 THE PROBLEM

4.1 There may not be an answer to be had.

The thought experiment causes a problem if you think that you should subjectively expect to survive quantum suicide. The conventional quantum suicide thought experiment presents you with two supposedly distinct possible outcomes – one in which there is a conscious observer and one in which there is not – with the distinction being the supposedly clear-cut one between your normal self and someone who is turned into a corpse in such a short space of time that you do not need to take account of his experiences.

The *delayed quantum suicide* thought experiment described here, though, removes this supposedly clear distinction between outcomes. There are still two outcomes, but you are presented with a *scale* of possible degrees of difference between the two outcomes, with no obvious way of deciding when to count the outcome in which you die as containing a conscious observer. It is clear that when the time delay is very long the relevant outcome must be counted and that when the time delay is negligibly short the outcome cannot be counted. It is less obvious what should be done about the time delays in-between: how long does the time delay need to be for a branch to be counted? If you cannot resolve this issue then you cannot answer the question: *If MWI is correct, for which settings of the time-delay dial, if any, do you subjectively expect to survive?*

It may seem that, if you think that you are assured of survival in conventional quantum suicide, the problem with the delayed quantum suicide scenario is that you do not know the answer. The problem, however, runs deeper than that: as was the case with the quantum brain damage scenario [6], *the scenario brings into question the very idea that there is an answer to be had with this way of thinking.*

4.2 The Problem with an Abrupt Transition

The idea that there is a transition point between the time delay being too long and being too short and that a small increase or decrease in the time delay should change the situation from one to the other should appear implausible. The events that happen on either side of the transition point –

wherever it is supposed to be – are essentially the same, so a profound change in the high-level nature of a sequence of events is happening with an insignificant change to the actual underlying sequence of events that happens. If we make a given change, someone might object that the change is not really very small, but we could respond by making the steps still smaller: we could even reduce the size of the change in the time delay until it is nudging up against the uncertainty principle.

The idea of a profound change in the high-level nature of a sequence of events being caused by a trivially small change in the underlying sequence of events is not something we normally associate with physical systems and their properties, and it is so unusual that, as well as being implausible, it suggests that minds are being treated in a radically different way to everything else, and this should be suggestive of dualism.

The problem also has similarities with a problem that some people regard as existing with the “consciousness causes collapse” interpretation of quantum mechanics: how do you draw the line between conscious observers and everything else? Many people who think that MWI is likely to be true might have problems with this issue when they see it in that context, so the same issue should suggest problems to them when they meet it in the context of delayed quantum suicide.

Although it presents problems of plausibility and is suggestive of dualism, let us suppose that there is some point along the scale from a very short time delay to a very long time delay at which an “observation” starts to be made in a branch – or something equivalent starts to happen: any version of the thought experiment with a time delay at least this long is going to have the outcome in which you are killed subjectively counted as a possibility for your future. What is it that is supposed to determine where that point is? We should demand that whatever rules govern this are justifiable and relate to the underlying physical system – the underlying physical events involving the physical substrate – but it should be hard to imagine how such rules are going to deliver the abruptness of transition that is needed when nothing else about the system seems to have this abruptness. If the rules do *not* relate to underlying physical events in the underlying physical system then we seem to be

in the realm of some “metaphysical rules” about what constitutes a conscious observation and what does not that work in some mysterious, unknowable way, and which have some kind of existence apart from things like brains. This, again, should be suggestive of dualism and should be implausible.

So, there must be some rules that determine when a sequence of events implies an observation and when it does not – or something equivalent that indicates when a path should be counted in the statistics of future expectations – whether these rules relate to the physical events in the underlying physical system or whether they are “metaphysical” in some sense – and they give the abrupt transition from conscious observation to absence of it, or something equivalent, that is needed to declare some branches irrelevant. You do not know what these rules are. Nobody does. That being the case, how can you be sure that the rules are even going to give the abrupt transition from observation by a conscious observer to absence of it, or behave in an equivalent way with regard to telling you which paths to count, as a single small change is made? A serious problem has crept in here. The idea that you should expect subjectively to survive quantum suicide is based on an idea that there are branches in which you are not able to make conscious observations and branches in which you are, but that idea has now been shown to rely on something else – some other way of approaching this and dealing with consciousness. The appeal to the simple idea that you can make conscious observations in some branches and not others no longer works. Instead, you are reduced to *hoping* that the underlying, unknown approach to consciousness and observation – the correct one, whatever it is – agrees with this idea and gives this abrupt transition as the time delay is slightly decreased or increased. We should be far from convinced that this is the case: with metaphysical rules that have nothing to do with your brain, we do not have any idea what they would do, and with an approach based on the events in the physical system in some way, such an approach is going to have to get involved in the actual substrate that causes human minds – the underlying mess of matter – and things may be more involved.

4.3 The Problem with a Gradual Change

You may think that there is no abrupt transition between being able to make a conscious observation and not being able to make one as the time delay is decreased or increased. What if there is, instead, a gradual decline in ability to make an observation as the time delay decreases? Maybe the observation becomes “weaker” in some way? The problem here is that it seems to leave any time delay as being associated with a branch in which you are there to observe – even the ones with very short time delays. This approach therefore seems to be unworkable.

If you think there is a gradual decline in ability to observe as the time delay is decreased, you may think that some kind of measure approach can be used – that the probability of finding yourself in a particular future is somehow related to the strength of your ability to observe it – so that as your ability to make observations in some future decreases, it is increasingly unlikely that you will find yourself in that future. You might think that the probability of an observation being made in some branch – or of something equivalent happening that establishes it as a possible future – is related to the amount of time available in it, so that a branch in which you die quickly is one in which whatever is supposed to happen to make it a “possible future” is less likely to happen. The problem with all this is that it is, again, a significant departure from the simple idea that you are there to observe in some branches and not in others, and that just the former constitute possible futures for you. Instead, you must think there are some underlying rules that describe what the probabilities are. As before, these rules must be metaphysical rules that are just there – for no reason (which is implausible and suggestive of dualism) – or they must somehow relate to events in the underlying physical system. Either case has the problem that, whatever they are, the rules may not agree with the simple assumption in the quantum suicide thought experiments: the assumption is again unsafe.

4.4 The Problem with Consciousness Only Being Relevant Over Some Durations of Time

You may think that it is a mistake even to talk about a transition, abrupt or gradual, from ability to make consciousness observations or lack of it as the available time duration is reduced – that

consciousness is only a meaningful idea when considering the behaviour of a system over some reasonable period of time, and that when this is reduced, conscious observations, rather than ceasing to exist or becoming “weaker” in some way, simply become irrelevant. With this view, the question of where to draw the line between ability to make conscious observations and lack of it may not be a meaningful one: the concept would just become less “useful” as the time delay was reduced. The problem with this view is that it leaves no way of saying which time delays should be associated with possible futures and which ones should not be.

4.5 The Problem with Saying that it is “Something Else” that Determines which Branches are Possible Futures

You may think that your ability to make conscious observations ceases to exist as the time delay decreases, or declines gradually, or becomes less relevant, but that this ability to make conscious observations is not the same as what is needed to say whether you are “in” a particular branch. Maybe something else is needed to say that you are in a particular branch, and maybe this is no longer there when the time delay decreases below a certain level? Such an idea will be of no help. This merely involves positing some kind of “secondary observation ability” – whatever it is that you are supposed to have that puts you in a branch – and we have the existing problem of how it is supposed to change between being there and not being there with a small change in the time duration. However this is dealt with, it must be done using an approach that uses general, underlying rules, with the problems that have previously been identified for these.

We return now to the question asked in the delayed quantum suicide thought experiment:

If MWI is correct, for which settings of the time-delay dial, if any, do you subjectively expect to survive?

If you think that your subjective survival is assured in the conventional quantum suicide experiment, the problem is that the simple idea used to justify that idea – that you can eliminate branches in which you are “not there to observe” from your set of possible futures – has been shown to be inadequate – whether you think that

the ability to make an observation abruptly ceases to exist when the available time to make it falls below a certain amount, that there is a gradual reduction in ability to make observations as the time available to do it decreases, that the making of conscious observations is only a meaningful idea over some period of time or that it is not directly the making of observations but “something else” associated with events in your brain that determines which branches should be viewed as possible futures. The idea that you should be sure of your survival in the quantum suicide scenario should be therefore be treated with some degree of scepticism.

5 CONCLUSION

The quantum suicide thought experiment involves a situation with two possible futures, one of which involves your survival and one of which involves your death, with the future that happens being determined by some quantum event. If the many-worlds interpretation of quantum mechanics (MWI) is true, both futures happen on different branches. One view of this is that you can eliminate the branch in which you die as a possible future for you, subjectively, and that therefore you should view your subjective survival as guaranteed. This idea has been questioned by the delayed quantum suicide thought experiment.

In the delayed quantum suicide thought experiment, there is no clear distinction between futures in which die and futures in which you live. Instead, death is not immediate in either future, with one future in which you die later, of some indeterminate cause (such as old age) and one future in which you die after some time delay, ranging from some large delay to hardly any delay at all, with many possible time delays in-between in different versions of the experiment. The thought experiment suggests the following question:

If MWI is correct, for which settings of the time-delay dial, if any, do you subjectively expect to survive?

If you think your subjective survival is assured in conventional quantum suicide, the problem here is one of deciding how short the time interval between the quantum event and death needs to be for a future to be discounted as a possible future with delayed quantum suicide. You can

clearly discount the branch in which you die when the time delay is extremely short – this is just like conventional quantum suicide – and you should clearly count the branch in which you die when the time delay is very long. It is the intermediate time delays that are the problem. You need to decide which of them allow an “observation” to be made after the quantum event and before your death – or something equivalent which will allow the relevant branch to be counted as a possible future. The idea that you count branches where you are there to observe, and do not count branches where you are not there to observe is too simplistic to deal with this. Instead, there is a need for a more general approach. Such an approach would seem to need some general, underlying rules that indicate when branches should be counted and when they should not.

We might imagine that the general, underlying rules are “metaphysical”, having nothing to do with the physical nature of your brain and the events occurring in it, but instead dealing with the issue in a different way. This should seem to be dualism and should seem implausible. If such metaphysical rules exist, your survival in conventional quantum suicide would only be assured if that occurred as a special case of these rules, but there should be little reason to think that that would be the case: we really have no idea of how such rules would work.

We might imagine, and this is more plausible, that the general, underlying rules take account of the physical situation, and that what happens is implied by the details of the physical situation in some way and can be justified just by examining the physical situation. Such rules would have to come out of some more general approach – some way of analyzing the situation that indicates which branches should be counted and which should not. Such rules do not necessarily need to deal with things in terms of “observations”. It may be that this is a concept that only makes sense in the simplified view used with the special case of conventional quantum suicide, but the rules must, somehow, do something equivalent in indicating which branches to count. If you think that you would survive conventional quantum suicide, this would need to follow as a special case of such general rules. However, we do not know what such rules are, so we cannot be sure about what any special cases of those rules would be like. Such general rules might

need to involve many things not typically considered in the conventional quantum suicide experiment, such as the method of death, how quickly it occurs, other aspects of the experimental setup and the way the human brain works. In fact, we might have every reason to think that a general approach *would* have to take account of such things. The very thing that the approach would have to do – deal with a continuum of different experiments, each slightly different from its neighbours in its fine details – would seem to require that whatever approach is used takes into account the fine details of the situation, and that should seem to make it likely that details of the situation could affect the outcome. This should give us reason to doubt that the simple reasoning behind the idea that your survival in conventional quantum suicide is guaranteed gives an accurate representation of the situation.

One way of answering this is simply to reject the entire idea that any branches can be eliminated in quantum suicide scenarios, and many people will already think that this is the correct approach.

The issue that has been raised here is similar to the issue raised by the quantum brain damage thought experiment, which we discussed previously [6]. Both delayed quantum suicide and quantum brain damage can be considered to be showing issues that are special cases of the same general issue: that it is not obvious where to draw the line between something that is an “observer” and something that is not, that any approach that deals with the issue of which branches to count may not necessarily give results that agree exactly with the simple reasoning used to justify an expectation of survival in conventional quantum suicide – which is just a special case – and, further, that any conceivable approach that addressed this issue would need to be a general approach that took into account the physical details of the situation, and we should therefore have every reason to expect that the details of the situation would affect the probabilities, making it unlikely that the simple reasoning used to justify an expectation of survival in quantum suicide was completely accurate.

What has been said here, however, does not mean that the idea that you should expect to survive quantum suicide must be viewed as *completely* wrong. Conventional quantum

suicide is an extreme of the kind of situation being considered here, and we might think that it tells us *something*. If, however, we accept that the idea has any validity at all, we should at least accept that the reasoning used to support this idea is a special case of a more general approach and that it may only approximate the actual

situation. In this view, you should possibly think that your chances of surviving quantum suicide are based on various things, and your survival, while possibly considered more likely due to some general version of the quantum suicide argument, may be far from assured.

REFERENCES

- [1] Moravec, H., 1988. The Doomsday Device. *Mind Children: The Future of Robot and Human Intelligence*. Harvard: Harvard University Press. p.188. (Also available at: <http://books.google.com/books?id=56mb7XuSx3QC&pg=PP188&pg=PA188> [Accessed 10 December 2010].)
- [2] Marchal, B., 1988. Mechanism and Personal Identity. *Proceedings of WOCFAI 91* (Paris. Angkor), pp.335-345. (Also available at: http://iridia.ulb.ac.be/~marchal/publications/M&PI_15-MAI-91.pdf [Accessed 10 December 2010].)
- [3] Tegmark, M., 1997. *The Interpretation of Quantum Mechanics: Many Worlds or Many Words?*. [Online] arXiv:quant-ph/9709032. Available at: <http://arxiv.org/abs/quant-ph/9709032> [Accessed 8 December 2010]. (Also available at: http://xxx.lanl.gov/PS_cache/quant-ph/pdf/9709/9709032v1.pdf [Accessed 12 December 2010].)
- [4] Everett, H., 1957. Relative State Formulation of Quantum Mechanics. *Reviews of Modern Physics*, 29, pp.454-462.
- [5] Price, M. C., 1995. *The Many-Worlds FAQ*. [Online] The Anthropic Principle. Available at: <http://www.anthropic-principle.com/preprints/manyworlds.htm> [Accessed 12 December 2010]. (Also available at: <http://www.hedweb.com/everett/everett.htm> [Accessed 12 December 2010] and <http://kuoi.com/~kamikaze/doc/many-worlds-faq.html> [Accessed 29 June 2011].)
- [6] Almond, P., 2011. *Quantum Brain Damage: Assuming you will survive quantum suicide may be simplistic*. [Online] paul-almond.com. Available at: http://www.paul-almond.com/Quantum_BrainDamage.pdf or http://www.paul-almond.com/Quantum_BrainDamage.doc [Accessed 18 July 2011].