Chapter 20

Delayed Choice Paradox

20.1 Statement of the Paradox

Consider the Mach-Zehnder interferometer shown in Fig. 20.1. The second beam splitter can either be at its regular position B_{in} where the beams from the two mirrors intersect, as in (a), or moved out of the way to a position B_{out}, as in (b). When the beam splitter is in place, interference effects mean that a photon which enters the interferometer through channel a will always emerge in channel f to be measured by a detector F. On the other hand, when the beam splitter is out of the way, the probability is 1/2 that the photon will be detected by detector E, and 1/2 that it will be detected by detector F.

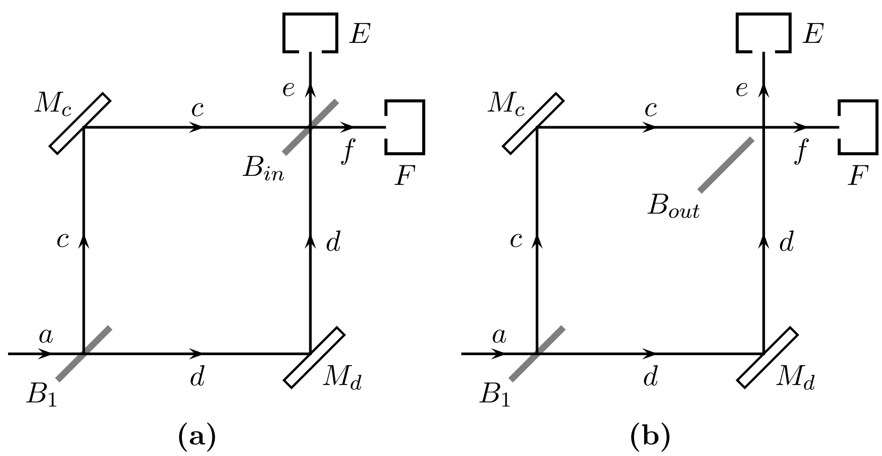

Figure 20.1: Mach-Zehnder interferometer with the second beamsplitter (a) in place, (b) moved out of the way.

The paradox is constructed in the following way. Suppose that the beam splitter is out of the way, Fig. 20.1(b), and the photon is detected in E. Then it seems plausible that the photon was earlier in the d arm of the interferometer. For example, were the mirror M_d to be removed, no photons would arrive at E; if the length of the path in the d arm were doubled by using additional mirrors, the photon would arrive at E with a time delay, etc. On the other hand, when the beam splitter is in place, we understand the fact that the photon always arrives at F as due to an interference effect arising from a coherent superposition state of photon wave packets in both arms c and d. That this is the correct explanation can be supported by placing phase shifters in the two arms, Sec. 13.2, and observing that the phase *difference* must be kept constant in order for the photon to always be detected in F. Similarly, removing either of the mirrors will spoil the interference effect.

Suppose, however, that the beam splitter is in place until just before the photon reaches it, and is then suddenly moved out of the way. What will happen? Since the photon does not interact with the beam splitter, we conclude that the situation is the same as if the beam splitter had been absent all along. If the photon arrives at E, then it was earlier in the d arm of the interferometer. But this seems strange, because if the beam splitter had been left in place, the photon would surely have been detected by F, which requires, as noted above, that inside the interferometer it is in a superposition state between the two arms. Hence it would seem that a later event, the position or absence of the beam splitter as decided at the very last moment before its arrival must somehow influence the earlier state of the photon, when it was in the interferometer far away from the beam splitter, and determine whether it is in one of the individual arms or in a superposition state. How can this be? Can the future influence the past?

The reader may be concerned that given the dimensions of a typical laboratory Mach-Zehnder interferometer and a photon moving with the speed of light, it would be physically impossible to shift the beam splitter out of the way while the photon is inside the interferometer. But we could imagine a very large interferometer constructed some place out in space so as to allow time for the mechanical motion. Also, modified forms of the delayed choice experiment can be constructed in the laboratory using tricks involving photon polarization and fast electronic devices.

It is possible to state the paradox in counterfactual terms. Suppose the beam splitter is not in place and the photon is detected by E, indicating that it was earlier in the d arm of the interferometer. What *would* have occurred *if* the beam splitter had been in place? On the one hand, it seems reasonable to argue that the photon would certainly have been detected by F; after all, it is always detected by F when the beam splitter is in place. On the other hand, experience shows that if a photon arrives in the d channel and encounters the beam splitter, it has a probability of 1/2 of emerging in either of the two exit channels. This second conclusion is hard to reconcile with the first.

20.2 Unitary Dynamics

Let $|0a\rangle$ be the photon state at t_0 when the photon is in channel a, Fig. 20.1, just before entering the interferometer through the first (immovable) beam splitter, and let the unitary evolution up to a time t_1 be given by

$$|0a\rangle \mapsto |1\bar{a}\rangle := \bigl(|1c\rangle + |1d\rangle\bigr)/\sqrt{2}, \tag{20.1}$$

where $|1c\rangle$ and $|1d\rangle$ are photon wave packets in the c and d arms of the interferometer. These in turn evolve unitarily,

$$|1c\rangle \mapsto |2c\rangle, \quad |1d\rangle \mapsto |2d\rangle, \tag{20.2}$$

to wave packets $|2c\rangle$ and $|2d\rangle$ in the c and d arms at a time t_2 just before the photon reaches the second (movable) beam splitter.

What happens next depends upon whether this beam splitter is in or out. If it is in, then

$$B_{in}: \quad |2c\rangle \mapsto |3\bar{c}\rangle, \quad |2d\rangle \mapsto |3\bar{d}\rangle, \tag{20.3}$$

where

$$|3\bar{c}\rangle := \big(|3e\rangle + |3f\rangle\big)/\sqrt{2}, \quad |3\bar{d}\rangle := \big(-|3e\rangle + |3f\rangle\big)/\sqrt{2}, \tag{20.4}$$

and $|3e\rangle$ and $|3f\rangle$ are photon wave packets at time t_3 in the e and the f output channels. If the beam splitter is out, the behavior is rather simple:

$$B_{out}: \quad |2c\rangle \mapsto |3f\rangle, \quad |2d\rangle \mapsto |3e\rangle. \tag{20.5}$$

Finally, the detection of the photon during the time interval from t_3 to t_4 is described by

$$|3e\rangle|E^\circ\rangle \mapsto |E^*\rangle, \quad |3f\rangle|F^\circ\rangle \mapsto |F^*\rangle. \tag{20.6}$$

Here $|E^\circ\rangle$ and $|F^\circ\rangle$ are the ready states of the two detectors, and $|E^*\rangle$ and $|F^*\rangle$ the states in which a photon has been detected.

The overall time development starting with an initial state

$$|\Psi_0\rangle = |0a\rangle|E^\circ\rangle|F^\circ\rangle \tag{20.7}$$

at time t_0 leads to a succession of states $|\Psi_j\rangle$ at time t_j. These can be worked out by putting together the different transformations indicated in (20.1) to (20.6), assuming the detectors do not change except for the processes indicated in (20.6). For $j \geq 2$ the result depends upon whether the (second) beam splitter is in or out. At t_4 with the beam splitter in one finds

$$B_{in}: \quad |\Psi_4\rangle = |E^\circ\rangle|F^*\rangle, \tag{20.8}$$

whereas if the beam splitter is out, the result is a macroscopic quantum superposition (MQS) state

$$B_{out}: \quad |\Psi_4\rangle = |S^+\rangle := (|E^*\rangle|F^\circ\rangle + |E^\circ\rangle|F^*\rangle)/\sqrt{2}. \tag{20.9}$$

A second MQS state

$$|S^-\rangle := (-|E^*\rangle|F^\circ\rangle + |E^\circ\rangle|F^*\rangle)/\sqrt{2}, \tag{20.10}$$

orthogonal to $|S^+\rangle$, will be needed later.

20.3 Some Consistent Families

Let us first consider the case in which the beam splitter is out. Unitary evolution leading to the MQS state $|S^+\rangle$, (20.9), at t_4 obviously does not provide a satisfactory way to describe the outcome of the final measurement. Consequently, we begin by considering the consistent family whose support consists of the two histories

$$B_{out}: \quad \Psi_0 \odot [1\bar{a}] \odot [2\bar{a}] \odot [3\bar{c}] \odot \{E^*, F^*\} \tag{20.11}$$

at the times $t_0 < t_1 < t_2 < t_3 < t_4$. Here and later we use symbols without square brackets for projectors corresponding to macroscopic properties; see the remarks in Sec. 19.2 following (19.4). This family resembles ones used for wave function collapse, Sec. 18.2, in that there is unitary time evolution preceding the measurement outcomes. For this reason, however, it does not allow us to make the inference required in the statement of the paradox in Sec. 20.1, that if the photon is detected by E (final state E^*), it was earlier in the d arm of the interferometer. Such an assertion at t_1 or t_2 is incompatible with $[1\bar{a}]$ or $[2\bar{a}]$, as these projectors do not commute with the projectors C, D for the photon to be in the c or the d arm. (For toy versions of C and D, see (12.9) in Sec. 12.1.) In order to translate the paradox into quantum mechanical terms, we need to use a different consistent family, such as the one with support

$$B_{out}: \quad \Psi_0 \odot \begin{cases} [1c] \odot [2c] \odot [3f] \odot F^*, \\ [1d] \odot [2d] \odot [3e] \odot E^*. \end{cases} \tag{20.12}$$

Each of these histories has weight 1/2, and using this family one can infer that

$$B_{out}: \quad \Pr([1d]_1 \,|\, E_4^*) = \Pr([2d]_2 \,|\, E_4^*) = 1, \tag{20.13}$$

where, as usual, subscripts indicate the times of events. That is, if the photon is detected by E with the beam splitter out, then it was earlier in the d and not in the c arm of the interferometer. Note, however, that using the consistent family (20.11) leads to the equally valid result

$$B_{out}: \quad \Pr([1\bar{a}]_1 \,|\, E_4^*) = \Pr([2\bar{a}]_2 \,|\, E_4^*) = 1. \tag{20.14}$$

The single framework rule prevents one from combining (20.13) and (20.14), because the families (20.11) and (20.12) are mutually incompatible.

Next, consider the situation in which the beam splitter is in place. In this case the unitary history

$$B_{in}: \quad \Psi_0 \odot [1\bar{a}] \odot [2\bar{a}] \odot [3f] \odot F^* \tag{20.15}$$

allows one to discuss the outcome of the final measurement. It describes the photon using coherent superpositions of wave packets in the two arms at times t_1 and t_2, as suggested by the statement of the paradox. Based upon it one can conclude that

$$B_{in}: \quad \Pr([1\bar{a}]_1 \,|\, F_4^*) = \Pr([2\bar{a}]_2 \,|\, F_4^*) = 1, \tag{20.16}$$

which is the analog of (20.14). (While (20.14) and (20.16) are correct as written, one should note that the conditions E^* and F^* at t_4 are not necessary, and the probabilities are still equal to 1 if one omits the final detector states from the condition. It is helpful to think of Ψ_0 as always present as a condition, even though it is not explicitly indicated in the notation.) On the other hand, it is also possible to construct the counterpart of (20.12) in which the photon is in a definite arm at t_1 and t_2, using the family with support

$$B_{in}: \quad \Psi_0 \odot \begin{cases} [1c] \odot [2c] \odot [3\bar{c}] \odot S^+, \\ [1d] \odot [2d] \odot [3\bar{d}] \odot S^-, \end{cases} \tag{20.17}$$

where S^+ and S^- are projectors onto the MQS states defined in (20.9) and (20.10). Note that the MQS states in (20.17) cannot be replaced with pairs of pointer states $\{E^*, F^*\}$ as in (20.11), since the four histories would then form an inconsistent family. See the toy model example in Sec. 13.3.

It is worth emphasizing the fact that there is nothing "wrong" with MQS states from the viewpoint of fundamental quantum theory. If one supposes that the usual Hilbert space structure of quantum mechanics is the appropriate sort of mathematics for describing the world, then MQS states will be present in the theory, because the Hilbert space is a linear vector space, so that if it contains the states $|E^*\rangle|F^\circ\rangle$ and $|E^\circ\rangle|F^*\rangle$, it must also contain their linear combinations. However, if one is interested in discussing a situation in which a photon is detected by a detector, (20.17) is not appropriate, as within this framework the notion that one detector or the other has detected the photon makes no sense.

Let us summarize the results of our analysis as it bears upon the paradox stated in Sec. 20.1. No consistent families were actually specified in the initial statement of the paradox, and we have used four different families in an effort to analyze it: two with the beam splitter out, (20.11) and (20.12), and two with the beam splitter in, (20.15) and (20.17). In a sense, the paradox is based upon using only two of these families, (20.15) with B_{in} and the photon in a superposition state inside the interferometer, and (20.12) with B_{out} and the photon in a definite arm of the interferometer. By focusing on only these two families—they are, of course, only specified implicitly in the statement of the paradox—one can get the misleading impression that the difference between the photon states inside the interferometer in the two cases is somehow caused by the presence or absence of the beam splitter at a later time when the photon leaves the interferometer. But by introducing the other two families, we see that it is quite possible to have the photon either in a superposition state or in a definite arm of the interferometer both when the beam splitter is in place and when it is out of the way. Thus the difference in the type of photon state employed at t_1 and t_2 is not determined or caused by the location of the beam splitter; rather, it is a consequence of a choice of a particular type of quantum description for use in analyzing the problem.

One can, to be sure, object that (20.17) with the detectors in MQS states at t_4 is hardly a very satisfactory description of a situation in which one is interested in which detector detected the photon. True enough: if one wants a description in which no MQS states appear, then (20.15) is to be preferred to (20.17). But notice that what the physicist does in employing this altogether reasonable criterion is somewhat analogous to what a writer writing a novel does when changing the plot in order to have the ending work out in a particular way. The physicist is selecting histories which at t_4 will be useful for addressing the question of which detector detected the photon, and not whether the detector system will end up in S^+ or S^-, and for this purpose (20.15), not (20.17) is appropriate. Were the physicist interested in whether the final state was S^+ or S^-, as could conceivably be the case—e.g., when trying to design some apparatus to measure such superpositions—then (20.17), not (20.15), would be the appropriate choice. Quantum mechanics as a fundamental theory allows either possibility, and does not determine the type of questions the physicist is allowed to ask.

If one does not insist that MQS states be left out of the discussion, then a comparison of the histories in (20.12) and (20.17), which are identical up to time t_2 while the photon is still inside the interferometer, and differ only at later times, shows the beam splitter having an ordinary causal effect upon the photon: events at a later time depend upon whether the beam splitter is or is not in place, and those at an earlier time do not. The relationship between these two families is then similar to that between (20.11) and (20.15), where again the presence or absence of the beam splitter when the photon leaves the interferometer can be said to be the cause of different behavior at later times. Causality is actually a rather subtle concept, which philosophers have been arguing about for a long time, and it seems unlikely that quantum theory by itself will contribute much to this discussion. However, the possibility of viewing the presence or absence of the beam splitter as influencing later events should at the very least make one suspicious of the alternative claim that its location influences earlier events.

20.4 Quantum Coin Toss and Counterfactual Paradox

Thus far we have worked out various consistent families for two quite distinct situations: the beam splitter in place, or moved out of the way. One can, however, include both possibilities in a single framework in which a quantum coin is tossed while the photon is still inside the interferometer, with the outcome of the toss fed to a servomechanism which moves the beam splitter out of the way or leaves it in place at the time when the photon leaves the interferometer. This makes it possible to examine the counterfactual formulation of the delayed choice paradox found at the end of Sec. 20.1.

The use of a quantum coin for moving a beam splitter was discussed in Sec. 19.2, and we shall use a simplified notation similar to (19.7). Let $|B_0\rangle$ be the state of the quantum coin, servomechanism, and beam splitter prior to the time t_1 when the photon is already inside the interferometer, and suppose that during the time interval from t_1 to t_2 the quantum coin toss occurs, leading to a unitary evolution

$$|B_0\rangle \mapsto (|B_{in}\rangle + |B_{out}\rangle)/\sqrt{2}, \tag{20.18}$$

with the states $|B_{in}\rangle$ and $|B_{out}\rangle$ corresponding to the beam splitter in place or removed from the path of the photon. The unitary time development of the photon from t_2 to t_3, in agreement with (20.3) and (20.5), is given by the expressions

$$\begin{aligned} |2c\rangle|B_{in}\rangle &\mapsto |3\bar{c}\rangle|B_{in}\rangle, \quad |2d\rangle|B_{in}\rangle \mapsto |3\bar{d}\rangle|B_{in}\rangle, \\ |2c\rangle|B_{out}\rangle &\mapsto |3f\rangle|B_{out}\rangle, \quad |2d\rangle|B_{out}\rangle \mapsto |3e\rangle|B_{out}\rangle. \end{aligned} \tag{20.19}$$

The unitary time development of the initial state

$$|\Omega_0\rangle = |0a\rangle|B_0\rangle|E^\circ\rangle|F^\circ\rangle \tag{20.20}$$

can be worked out using the formulas in Sec. 20.2 combined with (20.18) and (20.19). In order to keep the notation simple, we assume that the apparatus states $|B_0\rangle$, $|B_{in}\rangle$, $|B_{out}\rangle$ do not change except during the time interval from t_1 to t_2, when the change is given by (20.18). The reader may find it helpful to work out $|\Omega_j\rangle = T(t_j, t_0)|\Omega_0\rangle$ at different times. At t_4, when the photon has been detected, it is given by

$$|\Omega_4\rangle = \Big(|B_{in}\rangle|E^\circ\rangle|F^*\rangle + |B_{out}\rangle|S^+\rangle\Big)/\sqrt{2}. \tag{20.21}$$

Suppose the quantum coin toss results in the beam splitter being out of the way at the moment when the photon leaves the interferometer, and that the photon is detected by E. What would have occurred if the coin toss had, instead, left the beam splitter in place? As noted in Sec. 19.4, to address such a counterfactual question we need to use a particular consistent family, and specify a pivot. The answers to counterfactual questions are in general not unique, since one can employ more than one family, and more than one pivot within a single family.

Consider the family whose support consists of the three histories

$$\Omega_0 \odot [1\bar{a}] \odot \begin{cases} B_{in} \odot [3f] \odot F^*, \\ B_{out} \odot [3\bar{c}] \odot \{E^*, F^*\} \end{cases} \tag{20.22}$$

at the times $t_0 < t_1 < t_2 < t_3 < t_4$. Note that B_{out} and E^* occur on the lower line, and we can trace this history back to $[1\bar{a}]$ at t_1 as the pivot, and then forwards again along the upper line corresponding to B_{in}, to conclude that if the beam splitter had been in place, the photon would have been detected by F. This is not surprising and certainly not paradoxical. (Note that having the E detector detect the photon when the beam splitter is absent is quite consistent with the photon having been in a superposition state until just before the time of its detection; this corresponds to (20.11) in Sec. 20.3.) To construct a paradox we need to be able to infer from E^* at t_4 that the photon was earlier in the d arm of the interferometer. This suggests using the consistent family whose support is

$$\Omega_0 \odot \begin{cases} [1\bar{a}] \odot B_{in} \odot [3f] \odot F^*, \\ [1c] \odot B_{out} \odot [3f] \odot F^*, \\ [1d] \odot B_{out} \odot [3e] \odot E^*, \end{cases} \tag{20.23}$$

rather than (20.22). (The consistency of (20.23) follows from noting that one of the two histories which ends in F^* is associated with B_{in} and the other with B_{out}, and these two states are mutually orthogonal, since they are macroscopically distinct.) The events at t_1 are contextual in the sense of Ch. 14, with $[1\bar{a}]$ dependent upon B_{in}, while $[1c]$ and $[1d]$ depend on B_{out}.

The family (20.23) does allow one to infer that the photon was earlier in the d arm if it was later detected by E, since E^* occurs only in the third history, preceded by $[1d]$ at t_1. However, since this event precedes B_{out} but not B_{in}, it cannot serve as a pivot for answering a question in which the actual B_{out} is replaced by the counterfactual B_{in}. The only event in (20.23) which can be used for this purpose is Ω_0. Using Ω_0 as a pivot, we conclude that had the beam splitter been in, the photon would surely have arrived at detector F, which is a sensible result. However, the null counterfactual question, "What would have happened if the beam splitter had been out of the way (as in fact it was)?", receives a rather indefinite, probabilistic answer. Either the photon would have been in the d arm and detected by E, or it would have been in the c arm and detected by F. Thus using Ω_0 as the pivot means, in effect, answering the counterfactual question after erasing the information that the photon was detected by E rather than by F, or that it was in the d arm rather than the c arm. Hence if we use the family (20.23) with Ω_0 as the pivot, the original counterfactual paradox, with its assumption that detection by E implied that the photon was earlier in d, and then asking what would have occurred if this photon had encountered the beam splitter, seems to have disappeared, or at least it has become rather vague.

To be sure, one might argue that there is something paradoxical in that the superposition state $[1\bar{a}]$ in (20.23) is present in the B_{in} history, whereas non-superposition states $[1c]$ and $[1d]$ precede B_{out}. Could this be a sign of the future influencing the past? That is not very plausible, for, as noted in Ch. 14, the sort of contextuality we have here, with the earlier photon state depending on the later B_{in} and B_{out}, reflects the way in which the quantum description has been constructed. If there is an influence of the future on the past, it is rather like the influence of the end of a novel on its beginning, as noted in the previous section. Or, to put it in somewhat different terms, this influence manifests itself in the theoretical physicist's notebook rather than in the experimental physicist's laboratory.

What might come closer to representing the basic idea behind the delayed choice paradox is a family in which $[1d]$ at t_1 can serve as a pivot for a counterfactual argument, rather than having to rely on Ω_0 at t_0. Here is such a family:

$$\Omega_0 \odot \begin{cases} [1c] \odot \begin{cases} B_{in} \odot [3\bar{c}] \odot S^+, \\ B_{out} \odot [3f] \odot F^*, \end{cases} \\ [1d] \odot \begin{cases} B_{in} \odot [3\bar{d}] \odot S^-, \\ B_{out} \odot [3e] \odot E^*. \end{cases} \end{cases} \tag{20.24}$$

If we use $[1d]$ at t_1 as the pivot for a case in which the beam splitter is out and the photon is detected in E, it gives a precise answer to the null counterfactual question of what would have happened had the beam splitter been out (as it actually was): the photon would have been detected by E and not by F. But now when we ask what would have happened had the beam splitter been left in place, the answer is that the system of detectors would later have been in the MQS state S^-. In the same way, if the photon is detected in F when the beam splitter is out, a counterfactual argument using $[1c]$ at t_1 as the pivot leads to the conclusion that had the beam splitter been in, the detectors would later have been in the MQS state S^+, which is orthogonal to, and hence quite distinct from S^-. Thus detection in F rather than E when the beam splitter is out leads to a different counterfactual conclusion, in contrast with what we found earlier when using Ω_0 as the pivot. That the answers to our counterfactual questions involve MQS states is hardly surprising, given the discussion in Sec. 20.3. And, as in the case of (20.17), the MQS states in (20.24) cannot be replaced with ordinary pointer states (as defined at the end of Sec. 9.5) E^* and F^* of the detectors, for doing so would result in an inconsistent family. Also note the analogy with the situation considered in Sec. 19.4, where looking for a framework which could give a more precise answer to a counterfactual question involving a spin measurement led to a family (19.12) containing MQS states.

Let us summarize the results obtained by using a quantum coin and studying various consistent families related to the counterfactual statement of the delayed choice paradox. We have looked at three different frameworks, (20.22), (20.23), and (20.24), and found that they give somewhat different answers to the question of what would have happened if the beam splitter had been left in place, when what actually happened was that the photon was detected in E with the beam splitter out. (Such a multiplicity of answers is typical of quantum and—to a lesser degree—classical stochastic counterfactual questions; see Sec. 19.4.) In the end, none of the frameworks supports the original paradox, but each framework evades it for a somewhat different reason. Thus (20.22) does not have photon states localized in the arms of the interferometer, (20.23) has such states, but they cannot be used as a pivot for the counterfactual argument, and remedying this last problem by using (20.24) results in the counterfactual question being answered in terms of MQS states, which were certainly not in view in the original statement of the paradox.

20.5 Conclusion

The analysis of the delayed choice paradox given above provides some useful lessons on how to analyze quantum paradoxes of this general sort. Perhaps the first and most important lesson is that a paradox must be turned into an explicit quantum mechanical model, complete with a set of unitary time transformations. The model should be kept as simple as possible: there is no point in using long expressions and extensive calculations when the essential elements of the paradox and the ideas for untangling it can be represented in a simple way. Indeed, the simpler the representation, the easier it will be to spot the problematic reasoning underlying the paradox. In the interests of simplicity we used single states, rather than macroscopic projectors or density matrices, for the measuring apparatus, and for discussing the outcomes of a quantum coin toss. A more sophisticated approach is available, see Sec. 17.4, but it leads to the same conclusions.

A second lesson is that in order to discuss a paradox, it is necessary to introduce an appropriate framework, which will be a consistent family if the paradox involves time development. There will, typically, be more than one possible framework, and it is a good idea to look at several, since different frameworks allow one to investigate different aspects of a situation.

A third lesson has to do with MQS states. These are usually not taken into account when stating a paradox, and this is not surprising: most physicists do not have any intuitive idea as to what they mean. Nevertheless, families containing MQS states may be very useful for understanding where the reasoning underlying a paradox has gone astray. For example, a process of implicitly (and thus unconsciously) choosing families which contain no MQS states, and then inferring from this that the future influences the past, or that there are mysterious nonlocal influences, lies behind a number of paradoxes. This becomes evident when one works out various alternative families of histories and sees what is needed in order to satisfy the consistency conditions.