The Misleading Equivalence of Decoherence and Branching

Guido Bacciagaluppi

MCMP, 14 April 2015

Introduction

Ken Wharton's short story 'Aloha' (in *Analog Science Fiction and Fact*, June 2003) is a love story between two people with opposite time arrows meeting around the mid-point of a universe with both past and future low-entropy boundary conditions.

The funny thing about free will, Felix thought, was that it always maintained its own illusion.

Aloha (for that was what he had named her) had implied that they had met several times in her past—in Felix's future—so now he knew something about the choices he would soon make. In a universe where free will reigned supreme, it would be a simple matter to create a paradox. Felix merely had to choose not to see Aloha again, and the universe would be inconsistent.

He laughed out loud at the idea. As if he were more powerful than the universe.

No, paradox-prevention had turned out to be a major underpinning of reality, the lynchpin to explaining why quantum mechanics worked the way that it did. He couldn't force a paradox, no matter how he tried.

We are familiar with (maybe not quite as) odd things happening through the time-reversal and time-recurrence objections in the foundations of statistical mechanics.

I suggest we should be worrying about the same kind of problems also in the context of decoherence.

Standard models of decoherence take unentangled states to *extremely highly* entangled states, and thereby very effectively diagonalise the reduced state of the system in the eigenbasis of the decohering variables.

Schematically,

$$|\psi\rangle \otimes |E_0\rangle \rightarrow \sum_i c_i(t)|\psi_i\rangle \otimes |E_i(t)\rangle$$

We may imagine (and in finite dimensions we know we have) *recurrence* after a suitably long time:

$$|\psi\rangle \otimes |E_0\rangle \to \sum_i c_i(t)|\psi_i\rangle \otimes |E_i(t)\rangle \to |\psi\rangle \otimes |E_0\rangle$$

What is happening in the very long time in-between? And given that (except when we artificially isolate a system) decoherence goes on all the time, should we not expect that whatever is in fact happening is a fairly *generic* kind of behaviour?

Here is a possible intuition: the system is in a dynamically stationary state of entanglement with the environment, in that the interaction involves *as many* particles getting entangled with the system as getting disentangled from it.

We shall give a very elementary look at this possibility, from two points of view:

- decoherence and branching
- dynamical consistency

Decoherence and branching

What seems to tell *against* this possibility: decoherence induces branching of the quantum state, but we cannot have branching in both directions, or only trivially so.

And even if it makes sense, we can *see* that the universal wave function is non-trivially branching all the time.

(Or rather, the evidence is naturally interpreted in terms of constant branching, at least if we adopt a no-collapse approach to quantum mechanics — e.g. Everett.)

We shall need a few definitions.

A *history* is a time-ordered sequence of (Heisenberg-picture) projectors:

$$P_{i_1}(t_1),\ P_{i_2}(t_2),\ \ldots,\ P_{i_n}(t_n)$$

The associated *history operator* is:

$$C_{\alpha(n)} := P_{i_n}(t_n)\ldots P_{i_1}(t_1)$$

If for all k the $P_{i_k}(t_k)$ are mutually orthogonal and sum to the identity we have a *history space*.

If we form sums of the projections in $\{C_{\alpha(n)}\}$, we obtain a *coarse-graining* of the history space.

Take a quantum state $|\Psi\rangle$, and define the *decoherence functional*

$$\mathcal{D}(C_{\alpha(n)}, C_{\beta(n)}) := \mathrm{Tr}\Big(C_{\alpha(n)}|\Psi\rangle\langle\Psi|C^*_{\beta(n)}\Big)$$

The positive number $\mathcal{D}(C_{\alpha(n)}, C_{\alpha(n)})$ is called the *weight* of the history $C_{\alpha(n)}$.

Weak decoherence condition: for distinct histories,

$$\mathrm{Re}\,\mathcal{D}(C_{\alpha(n)}, C_{\beta(n)}) = 0$$

Weak decoherence is equivalent to the disappearance of interference terms between histories, i. e. we may sum weights when coarse-graining (the weights ‘behave like probabilities’).

[As Diósi points out, weak decoherence does not respect composition: the real part of the decoherence functional of a product is not the product of the real part of the decoherence functionals of the factors!]

Decoherence condition: for distinct histories,

$$\mathcal{D}(C_{\alpha(n)}, C_{\beta(n)}) = 0$$

Decoherence is equivalent to orthogonality of the vectors $C_{\alpha(m)}(t_m)|\Psi\rangle$, thus to existence of mutually orthogonal projections $R_{\alpha(m)}(t_m)$ summing to identity, with

$$R_{\alpha(m)}(t_m)|\Psi\rangle = C_{\alpha(m)}(t_m)|\Psi\rangle \tag{1}$$

('permanent records').

In general, these projections might just be given by

$$R_{\alpha(m)}(t_m) = C_{\alpha(m)}(t_m)|\Psi\rangle\langle\Psi|C^*_{\alpha(m)}(t_m)$$

('generalised records').

We shall focus on the familiar case of environmentally-induced decoherence, with records in the environment, e.g.

$$R_{\alpha(m)}(t_m) = E^1_{i_1}(t_m) \otimes E^2_{i_2}(t_m) \otimes \ldots \otimes E^m_{i_m}(t_m) \quad (2)$$

The (schematic) picture is that at each time t_k, the system interacts with some degree of freedom in the environment such that

$$|\psi_i\rangle \otimes |e_0^k\rangle \rightarrow |\psi_i\rangle \otimes |e_i^k\rangle \tag{3}$$

and each such environmental degree of freedom then evolves separately to $|e_i^k(t_m)\rangle$.

Each projection (2) thus picks out exactly one component in the state

$$|\psi(t_m)\rangle = \sum_{i_m} c_{i_m}(t_m)|\psi_{i_m}\rangle|e^1_{i_1}(t_m)\rangle|e^2_{i_2}(t_m)\rangle \ldots |e^m_{i_m}(t_m)\rangle$$

and is a permanent record in the sense of (1).

Note that if we define a fine-graining of the original history space by

$$P_{j_1}(t_1) \otimes R_{\alpha(1)}(t_1) \ldots, \; P_{j_n}(t_n) \otimes R_{\alpha(n)}(t_n) \qquad (4)$$

this space decoheres, and for $t_m \geq t_k$ the 'conditional weight' of $\sum_{\alpha(m)|i_k=j_k} R_{\alpha(m)}(t_m)$ given $P_{j_k}(t_k)$ is 1.

The space (4) is an example of a history space that is *branching* (past-deterministic, backwards deterministic), i. e. any two histories of non-zero weight that coincide at any time t_j, coincide also at all previous times t_i.

One can prove the *branching-decoherence theorem*: if a history space is branching, then it is decoherent; and if a history space is decoherent, then it is a coarse-graining of a branching history space.

Note that analogous definitions and results can be derived if we consider instead the so-called backwards decoherence functional

$$\mathcal{D}(C^*_{\alpha(n)}, C^*_{\beta(n)}) = \mathrm{Tr}\Big(C^*_{\alpha(n)}|\Psi\rangle\langle\Psi|C_{\beta(n)}\Big)$$

Indeed, imagine that at each t_k, the system interacts with some degree of freedom in the environment such that

$$|\psi_i\rangle \otimes |f_i^k\rangle \leftarrow |\psi_i\rangle \otimes |f_0^k\rangle \tag{5}$$

backwards in time.

Then the projections

$$S_{\tilde{\alpha}(m)}(t_m) = F^m_{i_m}(t_m) \otimes F^{m+1}_{i_{m+1}}(t_m) \otimes \ldots \otimes F^n_{i_n}(t_m)$$

(with $\tilde{\alpha}(m)$ the multi-index $i_m, i_{m+1}, \ldots, i_n$) are records in the environment of *later* events (one might call them 'antirecords', or 'prophecies').

It follows that the histories

$$P_{j_1}(t_1) \otimes S_{\tilde{\alpha}(1)}(t_1) \ldots, \; P_{j_n}(t_n) \otimes S_{\tilde{\alpha}(n)}(t_n) \qquad (6)$$

will define a history space that is *antibranching* (forward deterministic, future-deterministic).

There is also a 'two-state' version of the formalism that uses the time-symmetric decoherence functional

$$\mathcal{D}^{\text{sym}}(\alpha(n), \beta(n)) = \text{Tr}\Big(\rho_f C_{\alpha(n)} \rho_i C^*_{\beta(n)}\Big)$$

Note, however, that if both initial and final state are pure then

$$\mathcal{D}^{\text{sym}}(\alpha(n), \beta(n)) = \langle \Psi_f | C_{\alpha(n)} | \Psi_i \rangle \langle \Psi_i | C_{\beta(n)} | \Psi_f \rangle$$

which will vanish for any distinct histories iff only *one* history has non-zero probability.

Note also that forwards and backwards decoherence can be non-trivially satisfied simultaneously.

If that is the case, it follows that forwards and backwards weights coincide.

(See GB, ‘Probability, Arrow of Time and Decoherence’, `http://philsci-archive.pitt.edu/3157/`, where the main example, however, is not physically interesting.)

Now for the sake of argument assume that (3) and (5) are dynamically consistent, at least during the given interval $[t_1, t_n]$ (we return to this assumption in the next section).

Then our history space

$$P_{i_1}(t_1),\ P_{i_2}(t_2),\ \ldots,\ P_{i_n}(t_n)$$

will indeed decohere both forwards and backwards.

Further, the history spaces (4) and (6) will be fine-grainings that are respectively branching and antibranching, and one would expect that their common refining

$$P_{\jmath_1} \otimes R_{\alpha(1)} \otimes S_{\tilde{\beta}(1)}, \ldots , P_{\jmath_n} \otimes R_{\alpha(n)} \otimes S_{\tilde{\beta}(n)}$$

would be in fact both branching (past-deterministic) and antibranching (future-deterministic).

That is, a history space satisfying both (3) and (5) would be a coarse-graining of a *deterministic* history space.

It this by itself a fatal objection to the existence of both records of past and of future events?

Not necessarily: branching becomes *perspectival*.

If one coarse-grains over the records of the future, one recovers the familiar branching structure of decoherence.

And if one coarse-grains over the records of the past, one obtains an analogous antibranching structure.

The real question is that of dynamical consistency of (3) and (5).

Dynamical consistency

We shall be looking at a toy example. Take $n = 2$, and let the state at t_1^- be a superposition or mixture of the four components

$$|\psi_1\rangle|e_0^1\rangle|e_0^2\rangle|f_1^1\rangle|f_1^2\rangle$$
$$|\psi_1\rangle|e_0^1\rangle|e_0^2\rangle|f_1^1\rangle|f_2^2\rangle$$
$$|\psi_2\rangle|e_0^1\rangle|e_0^2\rangle|f_2^1\rangle|f_1^2\rangle$$
$$|\psi_2\rangle|e_0^1\rangle|e_0^2\rangle|f_1^1\rangle|f_2^2\rangle$$

(no records of past events, but records of future events).

The state evolves between t_1^- and t_1^+ to the corresponding superposition or mixture of

$$
\begin{gathered}
|\psi_1\rangle|e_1^1\rangle|e_0^2\rangle|f_0^1\rangle|f_1^2\rangle \\
|\psi_1\rangle|e_1^1\rangle|e_0^2\rangle|f_0^1\rangle|f_2^2\rangle \\
|\psi_2\rangle|e_2^1\rangle|e_0^2\rangle|f_0^1\rangle|f_1^2\rangle \\
|\psi_2\rangle|e_2^1\rangle|e_0^2\rangle|f_0^1\rangle|f_2^2\rangle
\end{gathered}
$$

consistently with (3) and (5).

Now assume the system evolves freely between t_1^+ and t_2^- according to some unitary evolution. We obtain the corresponding superposition or mixture of:

$$
\begin{gathered}
(a|\psi_1\rangle + b^*|\psi_2\rangle)\ |e_1^1\rangle|e_0^2\rangle|f_0^1\rangle|f_1^2\rangle \\
(a|\psi_1\rangle + b^*|\psi_2\rangle)\ |e_1^1\rangle|e_0^2\rangle|f_0^1\rangle|f_2^2\rangle \\
(b|\psi_1\rangle - a^*|\psi_2\rangle)\ |e_2^1\rangle|e_0^2\rangle|f_0^1\rangle|f_1^2\rangle \\
(b|\psi_1\rangle - a^*|\psi_2\rangle)\ |e_2^1\rangle|e_0^2\rangle|f_0^1\rangle|f_2^2\rangle
\end{gathered}
$$

(they are orthogonal, so there is no interference—and no terms cancel out—even with an initial superposition).

But now we are in trouble, because $|\psi_i\rangle$ does not always match up with $|f_i^2\rangle$, e. g. the first component at t_2^- is

$$a|\psi_1\rangle|e_1^1\rangle|e_1^2\rangle|f_0^1\rangle|f_1^2\rangle + b^*|\psi_2\rangle|e_1^1\rangle|e_2^2\rangle|f_0^1\rangle|f_1^2\rangle$$

By (5), since $|\psi_2\rangle|f_2^2\rangle \mapsto |\psi_2\rangle|f_0^2\rangle$, in order to preserve unitarity between t_2^- and t_2^+ this needs to evolve to some

$$a|\psi_1\rangle|e_1^1\rangle|e_1^2\rangle|f_0^1\rangle|f_0^2\rangle + b^*|\psi_2\rangle|e_1^1\rangle|e_2^2\rangle|f_0^1\rangle|f_?^2\rangle$$

with $\langle f_0^2|f_?^2\rangle = 0$, but this is now inconsistent with (5).

As with Felix in the short story, if the system evolves freely the universe is inconsistent...

As with Felix in the short story, if the system evolves freely the universe is inconsistent...

But what about paradox prevention?

There are in fact two (essentially equivalent) methods of implementing 'paradox prevention'.

First, we can let the two-state formalism *postselect* for the 'good' components, by postulating there is a final state that is a mixture just of the desired components

$$
\begin{aligned}
&|\psi_1\rangle|e_1^1\rangle|e_1^2\rangle|f_0^1\rangle|f_0^2\rangle \\
&|\psi_2\rangle|e_1^1\rangle|e_2^2\rangle|f_0^1\rangle|f_0^2\rangle \\
&|\psi_1\rangle|e_2^1\rangle|e_1^2\rangle|f_0^1\rangle|f_0^2\rangle \\
&|\psi_2\rangle|e_2^1\rangle|e_2^2\rangle|f_0^1\rangle|f_0^2\rangle
\end{aligned}
$$

(it has to be a mixed state because of (7)).

The forward-evolving state ρ_i will develop components containing states orthogonal to $|f_0^2\rangle$, and the backwards-evolving state ρ_f will develop components containing states orthogonal to $|e_0^1\rangle$, but they will all be assigned weight 0 by the symmetrised decoherence functional.

Alternatively, we can deny there is free evolution between t_1^+ and t_2^-, and postulate that each of

$$|\psi_1\rangle|e_1^1\rangle|e_0^2\rangle|f_0^1\rangle|f_1^2\rangle$$
$$|\psi_1\rangle|e_1^1\rangle|e_0^2\rangle|f_0^1\rangle|f_2^2\rangle$$
$$|\psi_2\rangle|e_2^1\rangle|e_0^2\rangle|f_0^1\rangle|f_1^2\rangle$$
$$|\psi_2\rangle|e_2^1\rangle|e_0^2\rangle|f_0^1\rangle|f_2^2\rangle$$

evolves, respectively,

to

$$
\begin{gathered}
|\psi_1\rangle|e^1_1\rangle|e^2_0\rangle|f^1_0\rangle|f^2_1\rangle \\
|\psi_2\rangle|e^1_1\rangle|e^2_0\rangle|f^1_0\rangle|f^2_2\rangle \\
|\psi_1\rangle|e^1_2\rangle|e^2_0\rangle|f^1_0\rangle|f^2_1\rangle \\
|\psi_2\rangle|e^1_2\rangle|e^2_0\rangle|f^1_0\rangle|f^2_2\rangle
\end{gathered}
$$

(possibly up to phases). Note that this step *can* be done unitarily (because the vectors are orthogonal).

Combining this with (3) and (5), we see that all histories are indeed fully *deterministic*.

E.g. we could have at t_1^+ the state

$$\frac{1}{\sqrt{2}}\Big(a|\psi_1\rangle|e_1^1\rangle|e_0^2\rangle|f_0^1\rangle|f_1^2\rangle + b^*|\psi_1\rangle|e_1^1\rangle|e_0^2\rangle|f_0^1\rangle|f_2^2\rangle +$$

$$+b\,|\psi_2\rangle|e_2^1\rangle|e_0^2\rangle|f_0^1\rangle|f_1^2\rangle - a^*|\psi_2\rangle|e_2^1\rangle|e_0^2\rangle|f_0^1\rangle|f_2^2\rangle\Big)$$

evolving at t_2^- to

$$\frac{1}{\sqrt{2}}\Big(a|\psi_1\rangle|e_1^1\rangle|e_0^2\rangle|f_0^1\rangle|f_1^2\rangle + b^*|\psi_2\rangle|e_1^1\rangle|e_0^2\rangle|f_0^1\rangle|f_2^2\rangle +$$

$$+b\,|\psi_1\rangle|e_2^1\rangle|e_0^2\rangle|f_0^1\rangle|f_1^2\rangle - a^*|\psi_2\rangle|e_2^1\rangle|e_0^2\rangle|f_0^1\rangle|f_2^2\rangle\Big)$$

N.B. this can be rewritten as

$$\frac{1}{\sqrt{2}}\Big(|\psi_1\rangle|e_1^1\rangle|e_0^2\rangle|f_0^1\rangle\Big[a|f_1^2\rangle + b^*|f_2^2\rangle\Big]+ \\ +|\psi_2\rangle|e_2^1\rangle|e_0^2\rangle|f_0^1\rangle\Big[b|f_1^2\rangle - a^*|f_2^2\rangle\Big]\Big)$$

evolving to

$$\frac{1}{\sqrt{2}}\Big(|\psi_1\rangle\Big[a|e_1^1\rangle + b|e_2^1\rangle\Big]|e_0^2\rangle|f_0^1\rangle|f_1^2\rangle+ \\ |\psi_2\rangle\Big[b^*|e_2^1\rangle - a^*|e_2^1\rangle\Big]|e_0^2\rangle|f_0^1\rangle|f_1^2\rangle\Big)$$

which is suggestive.

Indeed, while the system is in fact entangled also with the antirecords of its state at t_2, it has not yet interacted with this degree of freedom in the environment.

If we assume this entanglement is hard to detect and we neglect the antirecords, we are left with the 'illusion' of free evolution for the system, since

$$|\psi_1\rangle\Big[a|e_1^1\rangle + b|e_2^1\rangle\Big]|e_0^2\rangle + |\psi_2\rangle\Big[b^*|e_2^1\rangle - a^*|e_2^1\rangle\Big]|e_0^2\rangle =$$
$$\Big[a|\psi_1\rangle + b^*|\psi_2\rangle\Big]|e_1^1\rangle|e_0^2\rangle + \Big[b|\psi_1\rangle - a^*|\psi_2\rangle\Big]|e_2^1\rangle|e_0^2\rangle$$

Note that the two-state method yields the same histories (up to histories of zero weight), and the coefficients can be chosen so as to yield the same non-zero probabilities as with the constrained unitary evolution.

With either method we obtain, indeed, a history space that is deterministic.

Conclusion

We see that time-symmetric decoherence leads to some satisfiable but odd constraints on the dynamical evolution along histories.

Indeed, the evolution of the system between t_1^+ and t_2^- depends on both the records of the past and the future.

On the other hand, this behaviour might not be obvious to a time-directed observer.

If some such effect were generically present in decoherence interactions, the customary picture of branching through decoherence would be misleading.

Branching would be a perspectival effect of coarse-graining over the (unobserved) records of the future, and there would be no branching at a more fundamental level.

The toy model is overly simple, but it may well generalise to the case in which an interaction between different parts of the system is switched on and off, as in a measurement.

The results would thus apply also to situations standardly described as branching in the Everett theory, and one would not have splitting but simply uncertainty about (present records of) the future!