Can quantum theory not consistently describe the use of itself?

A comment on an article of Frauchiger and Renner

Gerold Gründler [1]

In 2018, Daniela Frauchiger and Renato Renner published an article [1] titled "Quantum theory cannot consistently describe the use of itself".[2] Their argument is wrong, for reasons explicated in this comment.[3]

1. Overview

By definition of the experimental setup, the isolated labs of $\overline{F}$ and F will — in view of the outside world — not decohere. Instead their evolution, starting from state $|\text{init}\rangle = (1)$ to their entangled state $|\overline{L}\,\&\,L\rangle = (4)$ is unitary, consequently *deterministic*. And the measurement results of $\overline{W}$ and W are exclusively determined by (4). Therefore the measurement results of the superposed components of the friends $\overline{F}$ and F have *no impact at all* on the probable results obtained by $\overline{W}$ and W. All four agents know that. Consequently will $\overline{F}$ and F not start nonsense speculations about whatever impact of their results onto the results of $\overline{W}$ and W. Knowing quantum theory, they *know* that there is no such impact. The argument of Frauchiger and Renner does not reveal an inconsistency within quantum theory; instead it is based onto an inconsistent application of quantum theory.

[1] mailto:gerold.gruendler@astrophys-neunhof.de
[2] A preliminary version with many differences versus the final version has been posted in arXiv already in April 2016.
[3] This comment gained much from the clear presentation by D. V. Tausk [2].

2. The probabilities of the various measurement results

The generator prepares at the begin of each experimental run a quantum coin in the state

$$|\text{init}\rangle = \sqrt{\tfrac{1}{3}}\,|h\rangle + \sqrt{\tfrac{2}{3}}\,|t\rangle\ . \tag{1}$$

h codes for heads, t codes for tails. The friend $\overline{F}$ applies to this coin a measurement instrument with the orthogonal eigenvectors $|h\rangle$ and $|t\rangle$. With probability $P(h) = \frac{1}{3}$ the measurement result is heads, and with probability $P(t) = \frac{2}{3}$ the result is tails. Thereby the state $|\text{init}\rangle$ collapses (in $\overline{F}$'s point of view, but not as seen from the world outside her isolated lab) to $|h\rangle$ or $|t\rangle$, respectively.

Seen from the world outside the isolated lab $\overline{L}$, there is no collapse. Instead the lab $\overline{L}$ is — after $\overline{F}$ has noted her measurement result, but before she forwards the polarized electron to F — described by the coherent superposition state function

$$|\overline{L}\rangle = \sqrt{\tfrac{1}{3}}\,|h\rangle \otimes |\overline{L}_h\rangle + \sqrt{\tfrac{2}{3}}\,|t\rangle \otimes |\overline{L}_t\rangle\ . \tag{2}$$

Here $|\overline{L}_h\rangle$ resp. $|\overline{L}_t\rangle$ is describing the instrument which displays the result h resp. t, and the friend $\overline{F}$ and her brain, in which the result h resp. t is stored. This is a feature which we will encounter repeatedly in the sequel: If the unitary evolution is not interrupted by a collapse, then a measurement does not produce a unique result. Instead the observed object (in this case the quantum coin) and the measurement instrument get entangled, in this case into the entangled state (2).

It seems most unlikely that scientists will ever be able to keep a living human being in a well-controlled coherent superposition state like (2). But the friends $\overline{F}$ and F could be replaced by much simpler objects. Remarkably, in a recently reported [3] beautiful experiment, a quite similar setup as in the Frauchiger-Renner gedankenexperiment has been actually realized in the lab. In this experiment, the friends $\overline{F}$ and F were replaced by photons, and the memory states of their brains were replaced by the polarization states of those photons. Hence the Frauchiger-Renner gedankenexperiment is not as far away from reality as one might guess in the first moment, and a careful discussion of it's implications is certainly sensible.

If $\overline{F}$ gets the result h, she forwards to F an electron with polarization state $|\downarrow\rangle$. If she gets t, she forwards to F an electron with polarization state $|\rightarrow\rangle = \sqrt{\frac{1}{2}}\Big(|\downarrow\rangle + |\uparrow\rangle\Big)$. In perspective of the outside world, $\overline{F}$ does *never* get this or that result. Instead she gets *deterministically* in *each* run of the experiment the result $\sqrt{\frac{1}{3}}\,|h\rangle + \sqrt{\frac{2}{3}}\,|t\rangle$, and therefore forwards to F in each run of the experiment deterministically an electron with polarization state $\sqrt{\frac{1}{3}}\,|\downarrow\rangle + \sqrt{\frac{2}{3}}\,|\rightarrow\rangle$.

Next F measures the polarization of the electron, which she received from $\overline{F}$. She applies a measurement instrument with the two orthogonal eigenvectors $|\downarrow\rangle$ and $|\uparrow\rangle$.

$$P(\downarrow|h) = 1 \tag{3a}$$

is the conditional probability that F will observe the result $\downarrow$, given that $\overline{F}$ got the result h.

$$P(\uparrow|h) = 0 \tag{3b}$$

is the conditional probability that F will observe the result $\uparrow$, given that $\overline{F}$ got the result h.

$$P(\downarrow|t) = \tfrac{1}{2} \tag{3c}$$

is the conditional probability that F will observe the result $\downarrow$, given that $\overline{F}$ got the result t. And

$$P(\uparrow|t) = \tfrac{1}{2} \tag{3d}$$

is the conditional probability that F will observe the result $\uparrow$, given that $\overline{F}$ got the result t.

In her point of view, but not as seen from the world outside her isolated lab, F collapses the polarization state of the electron to $|\downarrow\rangle$ resp. to $|\uparrow\rangle$, depending on her measurement result. In perspective of the outside world, the unitary evolution is not interrupted. Instead F's measurement instrument, and F herself, including her brain, get *deterministically* in *each* run of the experiment entangled with the electron, which has the coherent superposition polarization state $\sqrt{\frac{1}{3}}\,|\downarrow\rangle + \sqrt{\frac{2}{3}}\,|\rightarrow\rangle = \sqrt{\frac{2}{3}}\,|\downarrow\rangle + \sqrt{\frac{1}{3}}\,|\uparrow\rangle$.

Furthermore, due to the transfer of the polarized electron from lab $\overline{L}$ to lab L, these both labs get entangled. Thus, after F has completed her measurement, the state of the entangled labs as seen from the outside world is

$$\begin{aligned}|\overline{L}\,\&\,L\rangle = \sqrt{\tfrac{1}{3}}\,|h\rangle|\overline{L}_h\rangle \otimes |\downarrow\rangle|L_\downarrow\rangle + \\ + \sqrt{\tfrac{2}{3}}\,|t\rangle|\overline{L}_t\rangle \otimes \sqrt{\tfrac{1}{2}}\Big(|\downarrow\rangle|L_\downarrow\rangle + |\uparrow\rangle|L_\uparrow\rangle\Big)\ .\end{aligned} \tag{4}$$

$|L_\downarrow\rangle$ resp. $|L_\uparrow\rangle$ is the state function describing F's measurement instrument and F herself, including her brain, after she observed a $\downarrow$ resp. an $\uparrow$ result. Note that I suppressed for simplicity the $\otimes$ signs between the Hilbert spaces related to the same lab, and kept them explicit only for products between $\overline{L}$ and L.

It is characteristic for an entangled state like (4), that it can not be factorized, i. e. it can not be written as a product $|\overline{L}\rangle \otimes |L\rangle$. Quantum theory does not assign a state function to $\overline{L}$ nor to L. It only assigns the well-defined state function $|\overline{L}\,\&\,L\rangle = (4)$ to the overall system $\overline{L}\,\&\,L$.

Frauchiger and Renner overlooked the entanglement of the labs $\overline{L}$ and L. Instead they obviously assumed that from the outside perspective of $\overline{W}$ and W the lab $\overline{L}$ of $\overline{F}$ is, after F has completed her measurement, in the coherent superposition state

$$|\overline{L}\rangle = \sqrt{\tfrac{1}{3}}\,|h\rangle|\overline{L}_h\rangle + \sqrt{\tfrac{2}{3}}\,|t\rangle|\overline{L}_t\rangle \ , \tag{5a}$$

and that the lab L of F is in the coherent superposition state

$$|L\rangle = \sqrt{\tfrac{2}{3}}\,|{\downarrow}\rangle|L_{\downarrow}\rangle + \sqrt{\tfrac{1}{3}}\,|{\uparrow}\rangle|L_{\uparrow}\rangle \ . \tag{5b}$$

This error, however, is only of minor importance. Their argument is wrong anyway, no matter whether (5) or (4) is correct. I will come back to this point in the next section. Upfront I continue the consideration with the correct state function (4).

Next $\overline{W}$ applies to $\overline{L}$ a measurement instrument with the orthogonal eigenvectors

$$|\overline{\text{OK}}\rangle = \sqrt{\tfrac{1}{2}}\Big(\,|h\rangle|\overline{L}_h\rangle - |t\rangle|\overline{L}_t\rangle\Big) \tag{6a}$$

$$|\overline{\text{fail}}\rangle = \sqrt{\tfrac{1}{2}}\Big(\,|h\rangle|\overline{L}_h\rangle + |t\rangle|\overline{L}_t\rangle\Big) \ . \tag{6b}$$

We expand the state $|\overline{L}\,\&\,L\rangle = (4)$ in this basis:

$$|\overline{L}\,\&\,L\rangle = \overbrace{|\overline{\text{OK}}\rangle\langle\overline{\text{OK}}}^{\pi_{\overline{\text{OK}}}}\underbrace{\langle\overline{\text{OK}}\,|\,\overline{L}\,\&\,L\rangle}_{a_{\overline{\text{OK}}}} + \overbrace{|\overline{\text{fail}}\rangle\langle\overline{\text{fail}}}^{\pi_{\overline{\text{fail}}}}\underbrace{\langle\overline{\text{fail}}\,|\,\overline{L}\,\&\,L\rangle}_{a_{\overline{\text{fail}}}} =$$

$$= |\overline{\text{OK}}\rangle\sqrt{\tfrac{1}{2}}\Big(\langle h|\langle\overline{L}_h| - \langle t|\langle\overline{L}_t|\Big)\,\Big[\sqrt{\tfrac{1}{3}}\,|h\rangle|\overline{L}_h\rangle \otimes |{\downarrow}\rangle|L_{\downarrow}\rangle +$$

$$+ \sqrt{\tfrac{2}{3}}\,|t\rangle|\overline{L}_t\rangle \otimes \sqrt{\tfrac{1}{2}}\Big(|{\downarrow}\rangle|L_{\downarrow}\rangle + |{\uparrow}\rangle|L_{\uparrow}\rangle\Big)\,\Big] +$$

$$
\begin{aligned}
&+ |\overline{\text{fail}}\rangle \sqrt{\tfrac{1}{2}}\Big(\langle h|\langle \overline{L}_h| + \langle t|\langle \overline{L}_t|\Big) \left[\sqrt{\tfrac{1}{3}}\,|h\rangle|\overline{L}_h\rangle \otimes |{\downarrow}\rangle|L_{\downarrow}\rangle + \right.\\
&\quad \left. + \sqrt{\tfrac{2}{3}}\,|t\rangle|\overline{L}_t\rangle \otimes \sqrt{\tfrac{1}{2}}\Big(|{\downarrow}\rangle|L_{\downarrow}\rangle + |{\uparrow}\rangle|L_{\uparrow}\rangle\Big)\right] = \\
&= \underbrace{\sqrt{\tfrac{1}{6}}}_{a_{\overline{\text{OK}}}}\,|\overline{\text{OK}}\rangle \otimes \Big[-|{\uparrow}\rangle|L_{\uparrow}\rangle \Big] + \\
&+ \underbrace{\sqrt{\tfrac{5}{6}}}_{a_{\overline{\text{fail}}}}\,|\overline{\text{fail}}\rangle \otimes \sqrt{\tfrac{1}{5}}\Big[\,2\,|{\downarrow}\rangle|L_{\downarrow}\rangle + |{\uparrow}\rangle|L_{\uparrow}\rangle\Big]
\end{aligned}
$$

Thus with probability

$$P(\overline{\text{OK}}) = |\, a_{\overline{\text{OK}}}\,|^2 = \tfrac{1}{6} \tag{7a}$$

$\overline{W}$ will measure the result $\overline{\text{OK}}$, thereby disentangle the isolated labs, and prepare them in the states

$$|\overline{L}_{\overline{\text{OK}}}\rangle = \sqrt{\tfrac{1}{2}}\Big(\,|h\rangle|\overline{L}_h\rangle - |t\rangle|\overline{L}_t\rangle\Big) \tag{7b}$$

$$|L_{\overline{\text{OK}}}\rangle = |{\uparrow}\rangle|L_{\uparrow}\rangle\ . \tag{7c}$$

Here the irrelevant phase factor (-1) has been dropped. With probability

$$P(\overline{\text{fail}}) = |\, a_{\overline{\text{fail}}}\,|^2 = \tfrac{5}{6} \tag{8a}$$

$\overline{W}$ will measure the result $\overline{\text{fail}}$, thereby disentangle the isolated labs, and prepare them in the states

$$|\overline{L}_{\overline{\text{fail}}}\rangle = \sqrt{\tfrac{1}{2}}\Big(\,|h\rangle|\overline{L}_h\rangle + |t\rangle|\overline{L}_t\rangle\Big) \tag{8b}$$

$$|L_{\overline{\text{fail}}}\rangle = \sqrt{\tfrac{4}{5}}\,|{\downarrow}\rangle|L_{\downarrow}\rangle + \sqrt{\tfrac{1}{5}}\,|{\uparrow}\rangle|L_{\uparrow}\rangle\ . \tag{8c}$$

After $\overline{W}$ has completed his measurement, W measures the lab L with a measurement instrument with the orthogonal eigenvectors

$$|\text{OK}\rangle = \sqrt{\tfrac{1}{2}}\Big(|{\downarrow}\rangle|L_{\downarrow}\rangle - |{\uparrow}\rangle|L_{\uparrow}\rangle\Big) \tag{9a}$$

$$|\text{fail}\rangle = \sqrt{\tfrac{1}{2}}\Big(|{\downarrow}\rangle|L_{\downarrow}\rangle + |{\uparrow}\rangle|L_{\uparrow}\rangle\Big)\ . \tag{9b}$$

We expand the states $|L_{\overline{\text{OK}}}\rangle =$ (7c) and $|L_{\overline{\text{fail}}}\rangle =$ (8c) of L in this basis:

$$\begin{aligned}
|L_{\overline{\text{OK}}}\rangle &= |\text{OK}\rangle\langle\text{OK}|L_{\overline{\text{OK}}}\rangle + |\text{fail}\rangle\langle\text{fail}|L_{\overline{\text{OK}}}\rangle = \\
&= |\text{OK}\rangle\sqrt{\tfrac{1}{2}}\Big(\langle{\downarrow}|\langle L_{\downarrow}| - \langle{\uparrow}|\langle L_{\uparrow}|\Big)\,|{\uparrow}\rangle|L_{\uparrow}\rangle + \\
&\quad + |\text{fail}\rangle\sqrt{\tfrac{1}{2}}\Big(\langle{\downarrow}|\langle L_{\downarrow}| + \langle{\uparrow}|\langle L_{\uparrow}|\Big)\Big(\sqrt{\tfrac{4}{5}}\,|{\downarrow}\rangle|L_{\downarrow}\rangle + \sqrt{\tfrac{1}{5}}\,|{\uparrow}\rangle|L_{\uparrow}\rangle\Big) = \\
&= \underbrace{-\sqrt{\tfrac{1}{2}}}_{a_{\text{OK},\overline{\text{OK}}}}\,|\text{OK}\rangle \underbrace{+\sqrt{\tfrac{1}{2}}}_{a_{\text{fail},\overline{\text{OK}}}}\,|\text{fail}\rangle
\end{aligned} \tag{10a}$$

$$\begin{aligned}
|L_{\overline{\text{fail}}}\rangle &= |\text{OK}\rangle\langle\text{OK}|L_{\overline{\text{fail}}}\rangle + |\text{fail}\rangle\langle\text{fail}|L_{\overline{\text{fail}}}\rangle = \\
&= |\text{OK}\rangle\sqrt{\tfrac{1}{2}}\Big(\langle{\downarrow}|\langle L_{\downarrow}| - \langle{\uparrow}|\langle L_{\uparrow}|\Big)\Big(\sqrt{\tfrac{4}{5}}\,|{\downarrow}\rangle|L_{\downarrow}\rangle + \sqrt{\tfrac{1}{5}}\,|{\uparrow}\rangle|L_{\uparrow}\rangle\Big) + \\
&\quad + |\text{fail}\rangle\sqrt{\tfrac{1}{2}}\Big(\langle{\downarrow}|\langle L_{\downarrow}| + \langle{\uparrow}|\langle L_{\uparrow}|\Big)\Big(\sqrt{\tfrac{4}{5}}\,|{\downarrow}\rangle|L_{\downarrow}\rangle + \sqrt{\tfrac{1}{5}}\,|{\uparrow}\rangle|L_{\uparrow}\rangle\Big) = \\
&= \underbrace{+\sqrt{\tfrac{1}{10}}}_{a_{\text{OK},\overline{\text{fail}}}}\,|\text{OK}\rangle \underbrace{+\sqrt{\tfrac{9}{10}}}_{a_{\text{fail},\overline{\text{fail}}}}\,|\text{fail}\rangle
\end{aligned} \tag{10b}$$

Thus the conditional probabilities for the results measured by W, given the results measured by $\overline{W}$, are

$$\begin{aligned}
P(\text{OK}|\overline{\text{OK}}) &= \tfrac{1}{2} & \qquad P(\text{fail}|\overline{\text{OK}}) &= \tfrac{1}{2} \\
P(\text{OK}|\overline{\text{fail}}) &= \tfrac{1}{10} & \qquad P(\text{fail}|\overline{\text{fail}}) &= \tfrac{9}{10}\ ,
\end{aligned} \tag{11}$$

from which with (7a) and (8a) the absolute probabilities

$$P(\text{OK}) = \tfrac{1}{6} \qquad\qquad P(\text{fail}) = \tfrac{5}{6} \tag{12}$$

result. With probability

$$P(\text{OK}, \overline{\text{OK}}) = P(\overline{\text{OK}}) \cdot P(\text{OK}|\overline{\text{OK}}) \overset{(7a),(11)}{=} \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12} \tag{13}$$

the combined result $\{\text{OK}, \overline{\text{OK}}\}$ will turn up, and the experiment will stop.

Thus far we encountered no inconsistencies or contradictions.

3. The alleged contradiction

The probabilities for the various results of $\overline{W}$'s and W's measurement results depend *exclusively* on the state function $|\overline{L}\,\&\,L\rangle = (4)$, and on *nothing else*. And that state function is *perfectly deterministic*, i. e. it is *perfectly identical* in each run of the experiment. $|\overline{L}\,\&\,L\rangle = (4)$ is in particular *absolutely independent* of the measurement results, which the superposed components of $\overline{F}$ and F observe in their measurements.

Even if — as assumed in error by Frauchiger and Renner — not the state function (4) but the state functions (5) would describe the the isolated labs of the friends, before $\overline{W}$ and W start their measurements, the state of the isolated labs — as seen from the outside by $\overline{W}$ and W — would be *perfectly deterministic*, i. e. *perfectly identical* in each run of the experiment, hence *absolutely independent* of the measurement results, which the superposed components of $\overline{F}$ and F observe in their measurements.

To construct an alleged inconsistency, Frauchiger and Renner let the components of the friends $\overline{F}$ and F, which have amplitudes $\sqrt{1/3}$ or $\sqrt{2/3}$ according to (4), start speculations about the impacts of their measurement results onto the probable results which $\overline{W}$ and W will obtain. This is not consistent with the setup of the gedankenexperiment.

The friends $\overline{F}$ and F are assumed to be informed about the setup of the experiment, and they are assumed to have a reasonable knowledge of textbook quantum theory. Hence they *know* that the state function $|\overline{L}\,\&\,L\rangle = (4)$, which is the only one which affects the results of $\overline{W}$ and W, is fully deterministic. (Alternatively, if the friends would adopt Frauchiger's and Renner's error and believe (5) to be correct, they would still *know* that those state functions are fully deterministic.) Why, then, should they all with a sudden forget all their knowledge of quantum theory, and start such nonsense speculations?

We do not know what human beings would feel if they could be forced into a coherent superposition state. I guess that they would not survive the experiment, or that this experiment would at least drive them crazy. But the basic assumption of this gedankenexperiment is, that the friends $\overline{F}$ and F neither are killed nor get crazy, but are very well able to perform measurements and note the results. Consequently they should not have lost their knowledge of quantum theory.

If $\overline{F}$ registers an h result and notes that observation into her lab-book, she *knows* (from her textbook on quantum theory and from her knowledge of the experimental setup) that she is only one component, with amplitude $\sqrt{1/3}$, of a coherent superposition, and she *knows* (from her textbook on quantum theory and from her knowledge of the experimental setup) that the other $\sqrt{2/3}$ component of herself observes at the same time a t result, and notes that observation into her $\sqrt{2/3}$ component of the same lab-book. Hence she will definitively not construct whatever contradiction between her "componential" observation and the likely observations of the other agents.

After $\overline{W}$ and W have completed their measurements, both friends $\overline{F}$ and F are still in coherent superposition states. We can collapse them to unique states by simply opening the doors of their labs, and let the molecules of the ambient air and the air molecules within the previously isolated labs interact uncontrolled. Then the air molecules decohere the friends within an unmeasurable short fraction of a second.[4]

For example, if $\overline{W}$ has measured $\overline{\text{OK}}$, then the friend $\overline{F}$ and her lab are at the end of the experiment in the superposition state

$$|\overline{L}_{\overline{\text{OK}}}\rangle \overset{(7c)}{=} \sqrt{\tfrac{1}{2}}\Big(\,|h\rangle|\overline{L}_h\rangle - |t\rangle|\overline{L}_t\rangle\Big)\ . \qquad (14)$$

Due to uncontrolled interaction with the ambient air molecules, this state collapses with probability $P = 1/2$ to $|h\rangle|\overline{L}_h\rangle$, and with probability $P = 1/2$ to $|t\rangle|\overline{L}_t\rangle$.

Let's assume that the state of $\overline{F}$ and her lab collapsed to $|t\rangle|\overline{L}_t\rangle$. If we would ask $\overline{F}$ after the collapse what she remembers, she would answer that she remembers to have observed t, and that she forwarded to F a $\rightarrow$ polarized electron. That would be confirmed by a clear and unique entry in her lab-book, documenting the t observation. This is not surprising, because due to the collapse by chance the amplitude of the $|t\rangle|\overline{L}_t\rangle$ component of $|\overline{L}_{\overline{\text{OK}}}\rangle = (14)$ increased from $\sqrt{1/2}$ to 1, while the amplitude of the $|h\rangle|\overline{L}_h\rangle$ component decreased from $\sqrt{1/2}$ to 0. We can not ask the $|h\rangle|\overline{L}_h\rangle$ component of $\overline{F}$ what she remembers, and we can not read in her lab-book, because the amplitude of both is zero. But $\overline{W}$'s $\overline{\text{OK}}$ result proves beyond doubt — if we assume that quantum theory is correct — that $\overline{F}$'s $|h\rangle|\overline{L}_h\rangle$ component indeed had the amplitude $\sqrt{1/2}$ before the collapse.

Thus there is absolutely no inconsistency in this experiment. As all four agents apply textbook quantum theory correctly, they agree on all computations listed in the previous section. The alleged inconsistency asserted by Frauchiger and Renner is merely an artifact, not caused by an inconsistency of quantum theory, but caused by an inconsistent application of quantum theory.

[4] See e. g. [4] for an easy introduction to the decoherence stuff.

References

[1] D. Frauchiger, R. Renner : *Quantum theory cannot consistently describe the use of itself*, Nature Com. **9**, 3711 (2018), 10.1038/s41467-018-05739-8 arXiv: 1604.07422 (2016) https://arxiv.org/abs/1604.07422

[2] D. V. Tausk : *A brief introduction to the Foundations of Quantum Theory and an analysis of the Frauchiger-Renner paradox*, arXiv: 1812.11140 (2018) https://arxiv.org/abs/1812.11140

[3] M. Proietti, A. Pickston, F. Graffitti, P. Barrow, D. Kundys, C. Branciard, M. Ringbauer, A. Fedrizzi : *Experimental rejection of observer-independence in the quantum world*, arXiv: 1902.05080 (2019) https://arxiv.org/abs/1902.05080

[4] *Decoherence*, APIN Circular se91319 (2013) https://astrophys-neunhof.de/mtlg/se91319.pdf