Quantum Measurement as a Final-State Interaction with a Macroscopic External System

K.-E. Eriksson
Faculty of Science and Technology
Karlstad University SE 651 88 Karlstad, Sweden
February 5, 2008

Abstract

A small quantum scattering system (the microsystem) is studied in interaction with a large system (the macrosystem) described by unknown stochastic variables. The interaction between the two systems is diagonal for the microsystem in a certain orthonormal basis, and the interaction gives an imprint on the macrosystem. Moreover, the interaction is assumed to involve only small transfers of energy and momentum between the two systems (as compared to typical energies/momenta within the microsystem). The analysis is carried out within scattering theory. Calculated in the conventional way, the transition amplitude for the whole system factorizes. The interaction taking place within the macrosystem is assumed to depend on the stochastic variables in such a way that, on the average, no particular basis vector state of the microsystem is favoured. The density matrix is studied in a formalism which includes generation of the ingoing state and absorption of the final state. Then the dependence of the final state on the conventional scattering amplitude for the microsystem is highly non-linear.

In the thermodynamic limit of the macrosystem, the density matrix of the ensemble (of microsystem plus macrosystem) develops into a final state which involves a set of macroscopically distinguishable states, each with the microsystem in one of the basis vector states and the macrosystem in an entangled state.

For an element of the ensemble, i.e., for a single measurement, the result is instead a random walk, where the microsystem ends up in one of the basis vector states (reduction of the wave packet).

Thus, the macrosystem can be interpreted as a measurement device for performing a measurement on the microsystem. The whole discussion is carried out within quantum mechanics itself without any modification or generalization.
1 Introduction: Quantum measurement as a process to be understood within quantum mechanics

It is often argued that the reduction of the wave function of a quantum system in connection with a measurement process cannot be understood within quantum mechanics itself due to the linear nature of the theory. We show here that if the measurement interaction is included in the quantum-mechanical description, then the amplitude of the quantum process itself—which plays the role of the wave function—enters in a non-linear way (see Eqs (33) and/or (105) below). If the measurement apparatus (which is unknown in detail) is represented by a large number of stochastic variables, then the reduction of the wave function results in the thermodynamic limit.

This paper is a revised and extended version of a previous paper [3]. (We refer to this paper for more references.)

A process in a microscopic quantum system \( \mu \) is described together with a related interaction with a macroscopic system \( A \) (the measurement apparatus), not known in any detail and therefore described by stochastic variables.

We assume that an observable \( R \) with non-degenerate eigenstates \( |j\rangle_\mu \) is to be measured. We assume the interaction between \( \mu \) and \( A \) to be such that the state \( |j\rangle_\mu \) of \( \mu \) makes an imprint on \( A \) without the state of \( \mu \) being changed. This leads to an entanglement of \( A \) with \( \mu \). The imprint on \( A \) by \( \mu \) is made with a small energy and momentum transfer. For each \( j \), the interaction with \( \mu \) induces \( A \) to set off along a specific succession of states with an increasing number of degrees of freedom involved. The notion of metastability of \( A \) will be more precisely defined in Section 3.

We use S-matrix theory, based on quantum field theory, to analyse the interaction within \( \mu \) and the interaction between \( \mu \) and \( A \) as a whole. The resulting transition probabilities then are non-linear in terms of the transition probabilities for a pure \( \mu \) process (without \( A \)).

Moreover, the unknown stochastic variables of \( A \) are allowed to have an enhancing or inhibiting influence on the transitions within \( A \) to a final state. Therefore, the different initial states of \( A \), described by stochastic variables, compete on an unequal basis to reach the final state, and the ensemble of final states can have a very different composition from that of the initial states.

The system \( A \) should not only be metastable; it should also be unbiased. We take this to mean that the corresponding enhancement factors and inhibition factors of \( A \) occur with the same frequency in the initial state.

In the limit of low energy and momentum transfer, the \( \mu - A \) interaction factorizes in the scattering amplitude (before normalization) and hence also in the transition rate. This factor from \( \mu - A \) interaction depends on \( \mu \) only through its final state, labelled by \( j \).

The stochastic variables of \( A \) can be introduced through a stepwise mapping procedure, thus going in steps from the situation of the microsystem \( \mu \) by itself to a situation where \( \mu \) interacts with a system \( A \) in the thermodynamic limit,
Such a mapping leads to a random walk of a kind that has been suggested earlier, with the understanding that quantum mechanics may have to be abandoned for a more general theory [4, 5]. In this paper, we consider a process that takes place within linear quantum mechanics itself but produces non-linearities. Instead of a general mapping, we have chosen here a highly simplified model for the whole stochastic dynamics. In this model, the mathematics can be carried out in detail, for a single measurement as well as for an ensemble of measurements. The procedure of increasing the number of degrees of freedom of $A$ is transparent. The result is a change in the final-state distribution over the crucial variables of $A$ from a unimodal distribution to a multimodal distribution describing the different outcomes of measurement.

The non-linear dependence of the transition probabilities for the entire process (for $\mu$ and $A$) in terms of the transition probabilities for the pure quantum process ($\mu$ without $A$) can be explained in perturbation theory. This is most easily done in a model with sources of the incoming states and sinks of outgoing states shown in Appendix A. We use there a method due to Kinoshita and applied in a similar context by Nakanishi [7], to generalise Feynman diagrams to represent the dynamics for the elements of the final state density matrix.

2 The microsystem: quantum decay or scattering

Let $|0\rangle_\mu$ be the initial state of a scattering or decay process and $|f\rangle_\mu$ a final state, assumed to be different from $|0\rangle_\mu$,

$$\langle f|0\rangle_\mu = 0. \tag{1}$$

Then in a plane-wave basis, the scattering operator has the matrix element

$$\mu\langle f|S|0\rangle_\mu = \delta^4(P_f - P_0_\mu)\mu\langle f|M|0\rangle_\mu, \quad P_0 = (m_0, 0, 0, 0), \tag{2}$$

where $\mu\langle f|M|0\rangle_\mu$ is the scattering amplitude.

If the initial state $|0\rangle_\mu$ represents an unstable system of mass $m$ and $|f\rangle_\mu$ a state of outgoing decay products, then

$$\Gamma = (2\pi)^{-1}S_f\delta^4(P_f - P_0_\mu)|\mu\langle f|M|0\rangle_\mu|^2, \tag{3}$$

is the decay rate, with $S_f$ denoting integration over $P_f$ and summation/integration over other variables of $|f\rangle_\mu$.

If instead $|0\rangle_\mu$ represents an incoming state of two colliding particles in their
centre-of-mass frame with momenta
\[ p_1 = (\epsilon_1, q, 0, 0), \]
\[ p_2 = (\epsilon_2, q, 0, 0), \]
\[ \epsilon_j = \sqrt{m_j^2 + q^2}, \quad j = 1, 2, \]
and
\[ q = \frac{1}{2} \left[ m_0^2 - (m_1 + m_2)^2 - (m_1 - m_2)^2 \left( 1 - \left( \frac{m_1 + m_2}{m_0} \right)^2 \right) \right], \quad (4) \]
then the scattering cross section (into the set of states included in the summation \( S_f \)) is
\[ \sigma = (2\pi)^2 \frac{\epsilon_1 \epsilon_2}{q m_0} S_f \delta^4(P_f - P_0)|\mu(M|0)_\mu|^2. \quad (5) \]
The density matrix for the pure initial state \( |0\>_\mu \) is
\[ \rho^{(0)} = |0\>_{\mu\mu}\langle 0|, \quad (6) \]
with
\[ \text{Tr}\rho^{(0)} = 1. \quad (7) \]
Equations (3) and (5) then take the form
\[ \Gamma = (2\pi)^{-1} S_f \delta^4(P_f - P_0)\text{Tr}(\rho^{(f)} M \rho^{(0)} M^\dagger), \]
\[ \sigma = (2\pi)^2 \frac{\epsilon_1 \epsilon_2}{q m_0} S_f \delta^4(P_f - P_0)\text{Tr}(\rho^{(f)} M \rho^{(0)} M^\dagger), \quad (8) \]
where
\[ \rho^{(f)} = |f\>_{\mu\mu}\langle f|. \quad (9) \]
Here \( \Gamma \) and \( \sigma \) are proportional to what we may call the weight of the process,
\[ w_0 = \text{Tr}(M \rho^{(0)} M^\dagger). \quad (10) \]
The corresponding final (decay or scattering) state is
\[ \rho^{(s)} = w_0^{-1} M \rho^{(0)} M^\dagger = \frac{M \rho^{(0)} M^\dagger}{\text{Tr}(M \rho^{(0)} M^\dagger)}; \quad \text{Tr}\rho^{(s)} = 1. \quad (11) \]
Thus, in general, the probabilities here, i.e., the diagonal elements of \( \rho^{(s)} \) are non-linear in the diagonal elements of \( M \rho^{(0)} M^\dagger \).
We can rewrite (8) as
\[ \Gamma = (2\pi)^{-1} S_f \delta^4(P_f - P_0)\text{Tr}(\rho^{(f)} \rho^{(s)}), \]
\[ \sigma = (2\pi)^2 \frac{\epsilon_1 \epsilon_2}{q m_0} S_f \delta^4(P_f - P_0)\text{Tr}(\rho^{(f)} \rho^{(s)}). \quad (12) \]
We introduce

$$|\psi\rangle_{\mu} = \frac{M|0\rangle_{\mu}}{\sqrt{\mu|0\rangle_{M^\dagger M|0\rangle_{\mu}}} = \sum_{j=1}^{n} \psi_{j}|j\rangle_{\mu},$$

(13)

where we use a basis of eigenstates of the observable $R$ (assuming non-degenerate eigenvalues),

$$R|j\rangle_{\mu} = r_{j}|j\rangle_{\mu}, \quad r_{j} = r_{j}^*, \quad r_{j} \neq r_{j} \text{ for } j \neq k$$

$$\mu \langle j|k\rangle_{\mu} = \delta_{j,k}, \quad \mu \langle j|0\rangle_{\mu} = 0.$$  

Thus (13) is the final state for $\mu$ in the absence of $A$, and

$$\psi_{j} = \frac{M_{j}}{\sqrt{\sum_{l=1}^{n} |M_{l}|^{2}}}, \quad M_{j} = \mu \langle j|M|0\rangle_{\mu}$$

(15)

and

$$\mu \langle \psi|\psi\rangle_{\mu} = \sum_{j=1}^{n} |\psi_{j}|^{2} = 1.$$  

(16)

The non-linearity as manifested in the expression for the density matrix (11) of the outgoing state is most easily explained in a formalism involving a source of the incoming state and a sink of the outgoing state. This is presented in Appendix A, where (11) appears as the result of a unitary time development. Then, using (13), we have

$$\rho^{(s)}_{\mu} = |\psi\rangle_{\mu}\langle\psi| = \sum_{j,k=1}^{n} \psi_{j}\psi_{k}^{*}|j\rangle_{\mu}\langle k|.$$  

(17)

### 3 The measurement apparatus

We next consider the system $\mu$ together with another system $A$ with many degrees of freedom. We shall use a set of discrete stochastic variables $e$ to characterize the initial state of $A$,

$$\rho^{(0)}_{A} = |0;0;\rangle_{AA} \langle 0;0;|.$$  

(18)

In $|0;0;\rangle_{A}$, the first zero stands for preparedness of $A$. The second zero indicates that no signal has started propagation through $A$. We shall also use two other sets of states to characterize $A$,

$$|j,0;\rangle_{A} \text{ and } |j,x;\rangle_{A}, \quad j = 1, 2, ..., n; \quad x = 1, 2, ..., 2X \gg 1.$$  

(19)
Here $j$ indicates that in the interaction between $\mu$ and $A$, the $j$th eigenstate of $R$ has made an imprint on $A$. In the first of these states, the zero indicates that signal propagation within $A$ has not started, whereas $x$ in the second set of states, indicates the state of propagation within $A$. For $x = 2X$, the signal has reached its goal in the sense that it is ready to be irreversibly recorded. (As we shall see, during this propagation from 0 to $2X$, the collapse of the wave function takes place. The reason for having a variable $x$ taking on values in this interval is to show how a quantitative change of $x$ describes a process that involves such a qualitative change.)

We assume the stochastic variables $e$ to be defined in such a way that they are constants of motion. They are assumed not to influence the copying process from $\mu$ to $A$ but (and even decisively) the signal propagation within $A$. In our model, we label according to this influence. We assume copying and signal propagation for the different $j$, to be totally independent processes but also not to introduce bias for any particular measurement result.

We choose $e$ as
\[
\begin{align*}
\mathbf{e} &= (e_1, \ldots, e_n) \in \Omega, \\
\mathbf{e}_j &= (e_{j1}, \ldots, e_{j(2X)}),
\end{align*}
\] (20)
and the set of values $\Omega$ for $\mathbf{e}$ to be
\[
e_{jx} = \pm 1; \quad j = 1, \ldots, n; \quad x = 1, \ldots, 2X.
\] (21)

We have chosen an even number $2X$ here, since this will slightly simplify the model. We assume the orthonormality conditions,
\[
\begin{align*}
A\langle 0, 0; \mathbf{e} | 0, 0; \mathbf{e}' \rangle_A &= \delta_{ee'}; \\
A\langle 0, 0; \mathbf{e} | j, x; \mathbf{e}' \rangle_A &= 0, \quad j = 1, \ldots, n; \quad x = 0, 1, 2, \ldots, 2X; \\
A\langle j, x; \mathbf{e} | k, x'; \mathbf{e}' \rangle_A &= \delta_{jk} \delta_{xx'} \delta_{ee'}.
\end{align*}
\] (22)

For measuring the observable $R$, the measuring apparatus $A$ should be classical, metastable and non-biased. We take classical and metastable to imply the following:

**Classical:**
a) the apparatus $A$ can be treated semiclassically with respect to the stochastic variables $\mathbf{e}$, in the sense that the density matrix of $A$ (218 generalised) is diagonal in $\mathbf{e}$ initially and remains diagonal in $\mathbf{e}$. Moreover, $X$ should be very large; the precise meaning of this will be made clear in the model of Section 6. Niels Bohr used to emphasize the classical nature of the measuring apparatus.

**Metastable:**
b) The interaction of $A$, originally in a state of preparedness, with $\mu$ in an eigenstate $|0\rangle_\mu$ of $R$ leads to a corresponding impact (copying) on $A$ (without changing $\mu$), involving the transition into a propagation path, specific for the
value \( j \) (see Section 4).

c) the stochastic variables \( \epsilon \) influence signal propagation within \( A \). Propagation up to the coordinate value \( x \) involves variables within \( \epsilon^{(x)} \),

\[
\epsilon^{(x)} = (\epsilon_1^{(x)}, \ldots, \epsilon_{n}^{(x)}), \quad \epsilon^{(2x)} = \epsilon.
\]  

(23)

Thus \( \epsilon^{(x)} \) influences amplitudes and transition rates/partial cross sections through final state interaction (propagation). For \( x = 2X \), we have the full process with the whole set of stochastic variables \( \epsilon \) (equations (20) and (21)) involved.

The precise meaning of a non-biased \( A \) will be introduced below in connection with the assumptions concerning signal propagation.

4 Interaction between quantum system and measurement apparatus

The initial state of the combined system of \( \mu \) and \( A \) is a product of (6) and (18),

\[
\rho^{(0)}_{\mu A}(\epsilon) = \rho^{(0)}_{\mu}(\epsilon) \otimes \rho^{(0)}_{A}(\epsilon) = |0\rangle_{\mu} \otimes |0,0,0;\epsilon\rangle_{AA} \langle 0,0,0;\epsilon|.
\]  

(24)

Here we have assumed a fixed value \( \epsilon \); the generalisation to a probability distribution over \( \epsilon \) will be introduced below.

Interaction (scattering or decay) within \( \mu \), with \( A \) staying passive, then leads to the state (with defined in (15))

\[
\rho^{(s)}_{\mu A}(\epsilon) = \frac{\rho^{(s)}_{\mu}(\epsilon) \otimes \rho^{(0)}_{A}(\epsilon)}{\sum_{j,k=1}^{n} |M_{jk}|^{2}} = \sum_{j,k=1}^{n} \psi_{j}^{*} \psi_{k}^{*} |j\rangle_{\mu} \langle k| \otimes |0,0,\epsilon\rangle_{AA} \langle 0,0,\epsilon|.
\]  

(25)

We then include the first step of the interaction between \( \mu \) and \( A \), the copying interaction, resulting in the change

\[
|j\rangle_{\mu} \otimes |0,0;\epsilon\rangle_{A} \Rightarrow |j\rangle_{\mu} \otimes |j,0;\epsilon\rangle_{A},
\]  

(26)

assumed to take place similarly in each channel. Then (25) is transformed into the state

\[
\rho^{(c)}_{\mu A}(\epsilon) = \sum_{j,k=1}^{n} \psi_{j}^{*} \psi_{k}^{*} |j\rangle_{\mu} \langle k| \otimes |j,0;\epsilon\rangle_{AA} \langle k,0;\epsilon|.
\]  

(27)
We then have to include signal propagation within $A$, involving degrees of freedom up to position $x$, let us say,

$$|j, 0; e \rangle_A \Rightarrow |j, x; e \rangle_A ,$$

which depends on the stochastic parameters $e^{(x)}$ through factors $B_{jx}(e^{(x)})$. The state at propagation position $x$, if absorption were to take place there, would be

$$\rho_{\mu A}(e) = \frac{\sum_{j=1}^{n} B_{jx}(e^{(x)}) B_{kx}(e^{(x)})^* \psi_j \psi_k^* |j \rangle_{\mu \mu} \otimes |j, x; e \rangle_A \langle k, x; e|}{\sum_{l=1}^{n} |\psi_l|^2 |B_{lx}(e^{(x)})|^2} .$$

We can think of $x$ as describing the successive involvement of new degrees of freedom into the entanglement with $\mu$. In (29), as compared to (27), $\psi_j$ has been replaced by $B_{jx}(e^{(x)}) \psi_j$, and a normalisation like that in (25) has been carried out. How this kind of normalization can come about in a linear theory with unitary time evolution, is discussed in Appendix A. The factorization in (30) is due to the small energy and momentum transfer in the copying process.

The outgoing particles of the quantum process are practically on their mass shells, and the influence of charged outgoing particles on $A$ is well approximated by the current density of a classical point particles emerging from a point-like scattering centre. This is discussed in Appendix B. (Clearly, this implies a restriction on the kind of apparatus that our discussion can apply to. We find it more of an advantage to be specific on this point rather than general, being confident that a generalisation can be done in a rather straightforward way.)

As we shall see, the non-linear dependence on $B_{jx}(e^{(x)})$ in (29) has very drastic consequences for $\rho_{\mu A}(e)$.

It is important to note that the weight analogous to (10) of the process leading to the state (29) is

$$w_x(e^{(x)}) = \sum_{k=1}^{n} |M_k|^2 |B_{kx}(e^{(x)})|^2 = w_0 \sum_{k=1}^{n} |\psi_k|^2 |B_{kx}(e^{(x)})|^2 .$$

Starting with an ensemble of initial states, the different sets $e$ of stochastic variables compete to reach a certain propagation state, because of the different weights (31). When taking the ensemble average over (29), then the sum in (31) cancels against the denominator of (29). We shall see that the density matrix of the ensemble becomes linear in $B_{jx} B_{kx}^*$. 

8
For $x = 2X$, considered to be the goal of competitive propagation, we introduce the notation

$$\rho^{(f)}_{\mu A}(e) = \rho^{(2X)}_{\mu A}(e),$$

$$B_j(e) = B_{j(2X)}(e).$$

Then

$$\rho^{(f)}_{\mu A}(e) = \sum_{j,k=1}^{n} B_j(e) B_k(e)^* \psi_j \psi_k^* \langle j | \otimes | j, 2X; e \rangle_{AA} \langle k | \otimes | k, 2X; e \rangle.$$

The weight of this process for a given set of stochastic variables $e$ is

$$w_f(e) = \sum_{k=1}^{n} |M_k|^2 |B_k(e)|^2 = w_0 \sum_{k=1}^{n} |\psi_k|^2 |B_k(e)|^2,$$

where we have used (31).

The relative weight for $e$ in the final state is

$$P(e) \sum_{k=1}^{n} |\psi_k|^2 |B_k(e)|^2$$

$$= \sum_{e'} P(e') \sum_{l=1}^{n} |\psi_l|^2 |B_l(e')|^2,$$

where $P(e)$ is the probability for $e$ in the initial state. If all $e$ are equally probable, i.e., for

$$P(e) = 2^{-2nX},$$

the relative weight is

$$q(e) = 2^{-2nX} \sum_{k=1}^{n} |\psi_k|^2 |b_k(e)|^2,$$

$$\sum_{e} q(e) = 1,$$

where

$$b_k(e) = \frac{B_k(e)}{B},$$

and

$$B = \sqrt{\langle |B_k(e)|^2 \rangle_e} = \sqrt{2^{-2nX} \sum_{e} |B_k(e)|^2}.$$
is assumed to be the same for all \( k \). We shall see that this assumption is satisfied through our interpretation of the apparatus \( A \). Inserted into (38), eq. (39) implies that
\[
\langle |b_k(\mathbf{e})|^2 \rangle = 1, \quad k = 1, \ldots, n.
\] (40)

We have discussed already the metastability of \( A \). We shall now specify \( B_jx(\mathbf{e}^{(x)}) \) of (29) and the condition that \( A \) is a non-biased measuring instrument. We define \( B_jx(\mathbf{e}^{(x)}) \) through the recursive relations
\[
B_jx(\mathbf{e}^{(x)}) = B_{j-1}(\mathbf{e}^{(x-1)})C_{jx},
\]
\[
C_{jx} = c_{jx}(1 + \frac{1}{2}\eta_{jx}e_{jx} - \frac{1}{8}\eta_{jx}^2) \prod_{k \neq j} c_{kx}(1 - \frac{1}{2}\eta_{kx}e_{kx} - \frac{1}{8}\eta_{kx}^2),
\]
\[
B_{j0} = 1,
\] (41)

where
\[
c_{jx} = c_{jx}^*, \quad |c_{jx} - 1| \ll 1;
\]
\[
\eta_{jx} = \eta_{jx}^*, \quad 0 < \eta_{\text{min}} < \eta_{jx} \ll 1.
\] (42)

In (41), a positive (negative) \( e_{jx} \) strengthens (weakens) the \( j \)th channel and weakens (strengthens) all others, because there is a mutual anticoincidence between the channels. The factors for strengthening or weakening the occurrence or non-occurrence of a certain channel for a certain propagation position \( x \) are the same; both values of (21) have the same a priori probability. This is our understanding of the non-bias of the measuring apparatus \( A \).

Using also (21) we get the following averages over the stochastic variables (to second order in the \( \eta \)'s),
\[
\langle C_{jx}^2 \rangle_{\mathbf{e}^{(x)}} = \prod_{l=1}^{n} c_{lx}^2,
\] (43)
and
\[
\langle B_{jx}^2 \rangle_{\mathbf{e}^{(x)}} = \prod_{y=1}^{x} \prod_{l=1}^{n} c_{ly}^2.
\] (44)

For \( x = 2X \) in (44), we get from (32)
\[
\langle B_{jx}^2 \rangle_{\mathbf{e}} = B^2; \quad B^2 = \prod_{x=1}^{2X} \prod_{l=1}^{n} c_{lx}^2,
\] (45)

which verifies (39).

The non-bias of \( A \) is thus manifest in the sense that \( C_{jx} \) and \( C_{kx} \) with \( k \neq j \) in (41) depend on the variables \( e_{lx} \) by factors that change into each other for
\( e_{jx} \leftrightarrow e_{kx} \), and the frequencies for the two cases in the initial state are the same according to (36).

We could have introduced independent random phase factors in \( C_{jx} \). We have not done so here because it is not needed. For the correlations between the \( C_{jx} \), we get (to order \( \eta^2 \))

\[
\langle C_{jx} C_{kx} \rangle = (1 - \frac{1}{2}(1 - \delta_{jk})(\eta_{jx}^2 + \eta_{kx}^2)) \prod_{l=1}^{n} c_{lx}^2 = e^{(-\frac{1}{2}(1 - \delta_{jk})(\eta_{jx}^2 + \eta_{kx}^2)) \prod_{l=1}^{n} c_{lx}^2}. \tag{46}
\]

For the final state, we get to the same order

\[
\langle B_j B_k \rangle = B^2 \exp \left( -\frac{1}{2}(1 - \delta_{jk}) \sum_{x=1}^{2x} (\eta_{jx}^2 + \eta_{kx}^2) \right). \tag{47}
\]

This agrees with (45) for \( j = k \), and goes to zero for \( j \neq k \) in the limit of infinite \( X \).

5  Statistical description of measurement dynamics

We shall now review the dynamics of the microsystem \( \mu \) in interaction with the macrosystem \( A \) described by the stochastic variables \( \mathbf{e} \) on an ensemble level.

We then start with the the whole ensemble of ingoing states, each state (24) entering with equal probability (36),

\[
\rho(0) = |0\rangle_{\mu} \langle 0| \otimes 2^{-2nX} \sum_{\mathbf{e}} |0, 0; \mathbf{e}\rangle_{AA} \langle 0, 0; \mathbf{e}|. \tag{48}
\]

The scattering taking place within \( \mu \), leads to the ensemble of scattering states of the type (25),

\[
\rho^{(s)} = 2^{-2nX} \sum_{\mathbf{e}} \rho^{(s)}_{\mu A}(\mathbf{e}) = \sum_{j,k=1}^{n} \psi_j^* \psi_k |j\rangle_{\mu} \langle k| \otimes 2^{-2nX} \sum_{\mathbf{e}} |0, 0; \mathbf{e}\rangle_{AA} \langle 0, 0; \mathbf{e}|, \tag{49}
\]

where \( \psi_j \) is given in terms of the scattering amplitudes by (15). So far, this is a trivial extension of the dynamics of \( \mu \).

The scattering is followed by interaction between \( \mu \) and \( A \), in the form of copying. The ensemble of copied states (27) is

\[
\rho^{(c)} = 2^{-2nX} \sum_{\mathbf{e}} \rho^{(c)}_{\mu A}(\mathbf{e}) = \sum_{j,k=1}^{n} \psi_j^* \psi_k |j\rangle_{\mu} \langle k| \otimes 2^{-2nX} \sum_{\mathbf{e}} |j, 0; \mathbf{e}\rangle_{AA} \langle k, 0; \mathbf{e}|. \tag{50}
\]
Here \( A \) has become entangled with \( \mu \). There are new non-zero components but of the same size as the corresponding states in (49). We note that whereas the restriction of \( \rho^{(\delta)}(s) \) to \( \mu \) is the full density matrix of \( \mu \), the corresponding restriction of \( \rho^{(c)}(c) \) is diagonal,

\[
\text{Tr}_A \rho^{(c)} = \sum_{j=1}^{n} |\psi_j|^2 |j\rangle_{\mu\mu} \langle j|.
\]

The next set of processes is signal propagation, i.e., the increase within \( A \) of the number of degrees of freedom taking part in the entanglement. We thus start from (29) and (31), noting that the relative weight for \( e_x(x) \) (i.e., the probability for \( e_x(x) \), if final absorption were to take place at the stage \( x \) of propagation) is

\[
q_x(e_x) = \frac{w_x(e_x)}{w_x(e_x)} = 2^{-n_x} \sum_{l=1}^{n} |\psi_l|^2 |b_{lx}(e_x)|^2.
\]

Here we have introduced, in analogy to (39) and (38),

\[
b_{lx}(e_x) = \frac{B_{lx}(e_x)}{B_x} \quad \text{and} \quad B_x = \sqrt{\langle |b_{lx}(e_x)|^2 \rangle_{e_x}}; \quad \langle |b_{lx}(e_x)|^2 \rangle_{e_x} = 1.
\]

The ensemble of the \( x \)th states of propagation (29) is then

\[
\rho^{(x)} = 2^{-n(2X-x)} \sum_{e} q_x(e_x) \rho_{\mu A}^{(x)}(e_x)
\]

with \( q_x(e_x) \) given by (52), and

\[
\rho_{\mu A}^{(x)}(e_x) = \frac{\sum_{j,k=1}^{n} b_{jx}(e_x) b_{kx}(e_x) \psi_j \psi_k^* |j\rangle_{\mu\mu} \otimes |j, x; e_x \rangle_{AA} \langle k, x; e_x |}{\sum_{l=1}^{n} |\psi_l|^2 |b_{lx}(e_x)|^2}
\]

with the notation (53).

For \( x = 2X \), we have the final ensemble (before absorption)

\[
\rho^{(f)} = \sum_{e} q(e) \rho_{\mu A}^{(f)}(e),
\]

\[
(54)
\]

\[
(55)
\]
with (see (37) and (33) with (38) and (39))

\[ q(e) = 2^{-2nX} \sum_{l=1}^{n} |\psi_l|^2 |b_l(e)|^2, \]

\[ \rho_{\mu A}^{(f)}(e) = \frac{\sum_{j,k=1}^{n} b_j(e)b_k(e)^* \psi_j \psi_k^* |j\rangle_{\mu \mu} \otimes |j, 2X; e\rangle_{A A} (k, 2X; e)}{\sum_{l=1}^{n} |\psi_l|^2 |b_l(e)|^2} \]  

(57)

The restriction of the ensemble density matrices (54) and (56) to \( \mu \) is

\[ \text{Tr}_A \rho^{(x)} = \text{Tr}_A \rho^{(f)} = \sum_{j=1}^{n} |\psi_j|^2 |j\rangle_{\mu \mu} \langle j|. \]  

(58)

The corresponding restrictions for the density matrices of the ensemble elements (55) and (57) are

\[ \text{Tr}_A \rho_{\mu A}^{(x)}(e) = \sum_{j=1}^{n} p_{jx}(e)|j\rangle_{\mu \mu} \langle j|, \quad p_{jx}(e) = \frac{|b_{jx}(e(x))|^2 |\psi_j|^2}{\sum_{l=1}^{n} |\psi_l|^2 |b_{lx}(e(x))|^2} \]  

\[ \text{Tr}_A \rho_{\mu A}^{(f)}(e) = \sum_{j=1}^{n} p_{j}^{(f)}(e)|j\rangle_{\mu \mu} \langle j|, \quad p_{j}^{(f)}(e) = \frac{|b_j(e)|^2 |\psi_j|^2}{\sum_{l=1}^{n} |\psi_l|^2 |b_l(e)|^2}. \]  

(59)

In the next section, we shall analyze the qualitative transition taking place for \( q_x(e) \) and \( \rho_{\mu A}^{(x)}(e) \) defined in (52) and (55) and appearing together in (54). We shall further simplify the model described in (41) to make it analytically soluble.

6 Simplified model for the dynamics and statistics of the quantum system in interaction with the measurement apparatus

We simplify the model of Section 4 by putting all factors \( c_{jx} \) equal to unity and by making all \( \eta_{jx} \) equal. Thus, instead of (42), we have more specifically,

\[ c_{jx} = 1, \quad \eta_{jx} = \eta, \quad 0 < \eta \ll 1, \quad \text{for} \]

\[ j = 1, ..., n, \quad x = 1, ..., 2X. \]  

(60)

This fixes \( C_{jx} \) in (41), and \( B_{jx} \) is easily determined. We go directly to the final state with \( B_j = B_{j(2X)} \) with the result that (see (39) and (38)) \( B = 1 \) and
\(b_j = B_j\). According to (38) and (39), we have
\[
|b_j(\mathbf{e})|^2 = \prod_{x=1}^{2X} \left[ (1 + \eta e_j x) \prod_{k \neq j} (1 - \eta e_k x) \right].
\] (61)

Also (47) is simplified,
\[
\langle b_j(\mathbf{e}) b_k(\mathbf{e}) \rangle = e^{-2X \eta^2 (1-\delta_{jk})}.
\] (62)

A short calculation using (60) gives the following recursive relation for \(p_{jx}\) of (59),
\[
p_{jx} = p_{j x-1} + \Delta p_{jx}
\]
\[
\Delta p_{jx} = \frac{2\eta(e_j x - \sum_{l=1}^{n} p_l x-1 e_{lx})}{1 + 2\eta \sum_{m=1}^{n} p_m x-1 e_{mx}}
\] (63)

with probability
\[
q(e_{1x}, \ldots, e_{nx}) = 2^{-n} \left( 1 + 2\eta \sum_{l=1}^{n} p_l x-1 e_{lx} \right)
\]
\[
\sum_{e_{1x}, \ldots, e_{nx}} q(e_{1x}, \ldots, e_{nx}) = 1.
\] (64)

This can be viewed as the \(x\)th step of a random walk. The relevant mean values over \(e_{jx} = \pm 1\) are [4]
\[
\langle \Delta p_{jx} \rangle = 0,
\]
\[
\langle \Delta p_{jx} \Delta p_{kx} \rangle = 4\eta^2 p_{j x-1} p_{k x-1} \left( \delta_{jk} - p_{j x-1} - p_{k x-1} + \sum_{l=1}^{n} p_l^2 x-1 \right).
\] (65)

They characterize the random walk, which has the corners of the probability simplex as its attractors. One way to see this is to look at the entropy
\[
S_x = -\sum_{j=1}^{n} p_{jx} \ln p_{jx}
\] (66)

along the random walk. For one step, we find to second order,
\[
\Delta S_x = S_x - S_{x-1} = \sum_{j=1}^{n} \frac{\partial S_{x-1}}{\partial p_{j x-1}} \Delta p_{jx} +
\]
\[
+ \frac{1}{2} \sum_{j,k=1}^{n} \frac{\partial^2 S_{x-1}}{\partial p_{j x-1} \partial p_{k x-1}} \Delta p_{jx} \Delta p_{kx} = -\sum_{j=1}^{n} \left( (\ln p_{j x-1}) \Delta p_{jx} + \frac{\Delta p_{jx}^2}{2p_{j x-1}} \right)
\] (67)
so that
\[\langle \Delta S_x \rangle = -2\eta^2 \sum_{j=1}^{n} p_{j x-1} \left( 1 - 2p_{j x-1} + \sum_{l=1}^{n} p_{l x-1} \right) = \]
\[= -2\eta^2 \sum_{j=1}^{n} p_{j x-1} \left( (1 - p_{j x-1})^2 + \sum_{l \neq j} p_{l x-1} \right) \leq 0.\] (68)

Thus the entropy decreases until one of the corners of the probability simplex is reached. Since the expectation value of $p_{j x}$, with increasing $x$, stays at its initial value $|\psi_j|^2$, the probability of approaching the $j$th corner is $|\psi_j|^2$.

Let us now go to the ensemble of random walks (with $2X$ steps), which is a diffusion process. Then what is important in (61) and hence also in (57) and (59) is how many $e_{jx}$ are positive or negative for each $j$. We assume $X \pm X_j$ cases of $e_{jx} = \pm 1$. We collect the values $X_j$ in vector notation,

\[X = (X_1, \ldots, X_n); \quad X_j = \frac{1}{2} \sum_{x=1}^{2X} e_{jx}, \quad -X \leq X_j \leq X.\] (69)

There are
\[
\prod_{l=1}^{n} \frac{(2X)!}{(X + X_j)!((X - X_j)!)}
\]
values of $e$ in the set $\Omega_X$, characterized by (69). We then define the following $n$ distributions over $X$,

\[P_j(X) = P(X_j) \prod_{k \neq j} P(-X_k); \quad \sum_{X} P_j(X) = 1,\] (71)

where
\[
P(Y) = \frac{(2X)!}{(X + Y)!((X - Y)!)} \left( \frac{1 + \eta}{2} \right)^X \left( \frac{1 - \eta}{2} \right)^{X - Y}; \quad \sum_{Y=-X}^{X} P(Y) = 1,
\]
\[
\sum_{Y=-X}^{X} P(Y)Y = X\eta, \quad \sum_{Y=-X}^{X} P(Y)Y^2 - X^2\eta^2 = \frac{1}{2}X.
\] (72)

Let $Q(X)$ be the distribution over $X$ corresponding to the distribution $q(e)$ for outgoing states in (56) and (57). Then using (61), (70) and (71), we have

\[Q(X) = \sum_{j=1}^{n} P_j(X) |\psi_j|^2.\] (73)
Similarly, since $p_j^{(f)}(e)$ of (59) depends on $e$ only through $X$, we have

$$ p_j(X) = \sum_{e \in \Omega_X} q(e) p_j^{(f)}(e) \quad \frac{1}{\sum_{e' \in \Omega_X} q(e')} = \frac{P_j(X)|\psi_j|^2}{\sum_{l=1}^n P_l(X)|\psi_l|^2}. \quad (74) $$

For $\sqrt{X/2} \ll X\eta$, i.e., for $X\eta^2 \gg 1$, $P(Y)$ overlaps very little with $P(-Y)$, and hence for $j \neq k$, $P_j(X)$ and $P_k(X)$ defined in (71), are also almost without overlap. The result is that $Q(X)$ in (73) is multimodal.

It is convenient to change into renormalized and continuous variables,

$$ z = \frac{1}{\sqrt{\pi}} \cdot \frac{X}{\sqrt{X}} = X^{\frac{1}{4}} \eta^{-\frac{1}{2}} X, $$

$$ Z = \eta\sqrt{X}, \quad (75) $$

and to approximate $P(Y)$ by

$$ P(Y) = \frac{1}{\sqrt{\pi X}} e^{-(Y-X\eta)^2/X}, \quad (76) $$

or, with $Y = X^{\frac{3}{4}} \eta/2 z$,

$$ P(Y)dy = \sqrt{\frac{Z}{\pi}} e^{-(z-\sqrt{Z})^2} dz. \quad (77) $$

This means that in $Q(X)$, $P_j(X)$ is narrowly centered around a specific point,

$$ Q(X)d^nX = \tilde{q}(z)dnz, \quad (78) $$

$$ \tilde{q}(z) = \sum_{j=1}^n |\psi_j|^2 g(z - z^{(j)}), \quad (79) $$

with

$$ g(z) = \left( \frac{Z}{\pi} \right)^{n/2} e^{-\frac{Zz^2}{\pi}}, $$

$$ z^{(j)}_k = \sqrt{Z}(2\delta_{jk} - 1). \quad (80) $$

For $p_j(X)$ in (74), only the values very close to $X = X^{\frac{3}{4}} \eta^{1/2} z^{(k)}$ are of interest, and either the numerator of (74) is negligible ($k \neq j$), or it coincides with the totally dominating term of the denominator ($k = j$). Using $z$ as argument (rather than $X$), we have in the limit of large $Z$,

$$ p_j(z^{(k)}) = \delta_{jk}. \quad (81) $$
In this limit, the peaks of (79) separate as well as narrow down. In terms of the semi-classical stochastic variables \( e \), according to (69) and (75),

\[
z_j = \frac{1}{2} X^{-\frac{3}{4}} \eta^{-\frac{1}{2}} \sum_{x=1}^{2X} e_{jx}.
\]  

(82)

The components of \( z \), each a sum of many small semi-classical variables, can be viewed as classical variables functioning as pointer variables. They are given by the unknown initial state of \( A \), but the distribution (79) of \( z \) in the final state, depends on the interaction between \( \mu \) and \( A \). One can say that the components of \( z \) are (non-local) variables hidden in the unknown initial state of \( A \).

The situation for \( n = 3 \), can easily be depicted in two dimensions (Fig. 1).

Thus we have seen how the distribution \( q(e) \), defined in (57) and appearing in the ensemble (56) of final states, corresponds to the pointer distribution \( \hat{q}(z) \) defined in (79) (with the relationship between \( z \) and \( e \) given by (82)). At the \( k \)th peak of \( \hat{q}(z) \), the corresponding state of \( \mu \) is the \( k \)th eigenstate of the observable \( R \) as indicated by (81).

The ensemble of final states (56) (with (57)) can be written

\[
\rho^{(f)}(j) = \sum_{j=1}^{n} |\psi_j|^2 |j\rangle_{\mu\mu} \otimes \rho_A^{(f,j)}
\]  

(83)

where the \( n \) states of \( A \),

\[
\rho_A^{(f,j)} = \left[ \prod_{l=1}^{n} \frac{(2X)!}{(X + X^{(j)}_l)!(X - X^{(j)}_l)!} \right]^{-1} \times
\]

\[
\times \sum_{X} P_j(X) \sum_{e \in \Omega_X} |j; 2X; e\rangle_A \langle j; 2X; e|,
\]

(84)

centered around \( X^{(j)}_k = X \eta(2\delta_{jk} - 1) \), are macroscopically distinguishable due to the negligible overlap of the different \( P_j(X) \).

The transition from a unimodal to a multimodal distribution can be followed in detail in equations (71), (72) and (73), by starting with a relatively small value for \( X \) and letting it increase to a large value \( (X \gg \eta^{-2}) \). This means that for a moment, we have given 2\( X \) the role of the variable \( x \) used in Sections 4 and 5.

Alternatively, one can follow this development in the continuous functions (79) and (80) (linked to (71)-(73) by (75), (77) and (78)) with increasing \( Z \). The transition of \( \hat{q}(z) \) from unimodal to \( n \)-modal takes place near \( Z = 1 \).

Instead of the model developed in this section, we could have studied the consequences of the slightly more general recursive relations (41) through successive mappings along increasing \( x \) in (29). These mappings are also successive steps in a random-walk or diffusion process. The mathematics would have been a bit more complicated, but the conclusion would have been of the same nature.
7 Six comments

7.1 The number of detector sets

So far, we have assumed one set of detectors for each of the $n$ eigenstates of $R$. In fact, only $n - 1$ detectors are needed. If there is no detector for the $n$th eigenstate, then (41) to (45) have to be slightly modified, but (37) to (40) still hold and the conclusion, as manifested by equations (58) and (59), is the same. These leads to changes in (69), (71) and (80) in Section 6 as follows:

$$X = (X_1, ..., X_{n-1}, 0); \quad X_j = \frac{1}{2} \sum_{x=1}^{2X} e_{jx}, \quad -X \leq X_j \leq X;$$

(85)

and

$$P_j(X) = P(X_j) \prod_{k \neq j,n} P(-X_k), \quad j = 1, ..., n - 1;$$

$$P_n(X) = \prod_{k=1}^{n-1} P(-X_k),$$

(86)

$$\sum_{X} P_j(X) = 1, \quad j = 1, ..., n.$$

In (80) $z$ is now restricted to $z_n = 0$, and we have more precisely,

$$g(z) = \left(\frac{Z}{\pi}\right)^{n/2} e^{-Zz^2},$$

(87)

$$z^{(j)}_k = \sqrt{Z(2\delta_{jk} - 1)}; \quad j = 1, ..., n - 1, n, \quad k = 1, ..., n - 1.$$

The splitting into a multimodal distribution in $z$ is thus of the same nature as before.

7.2 Entropy considerations

To find the entropy of the final state, only diagonal elements of the density matrix need to be considered. For given $j$ and $X$, there are

$$\prod_{l=1}^{n} \frac{(2X)!}{(X + X_l)!(X - X_l)!}$$

states, each with probability

$$|\psi_j|^2 2^{-2nX} (1 - \eta^2)^X \left(\frac{1 + \eta}{1 - \eta}\right)^{X_j}.$$
The average over $X$ of the logarithm of the inverse of (89) is to second order in $\eta$,
\begin{equation}
-\ln|\psi_j|^2 + 2nX\ln2 - X\ln(1-\eta^2) - X\eta\ln\frac{1+\eta}{1-\eta} =
\end{equation}
\begin{equation}
= -\ln|\psi_j|^2 + X(2n\ln2 - \eta^2).
\end{equation}
Averaging also over $j$, we find the entropy over the ensemble of final states,
\begin{equation}
-\sum_{k=1}^{n} |\psi_k|^2\ln|\psi_k|^2 + 2Xn\ln2 - X\eta^2.
\end{equation}
The first term here refers to the uncertainty inherent in the bifurcation following the interaction between $\mu$ and $A$, and it is not present after the bifurcation which is classical in nature (the pointer reading). The second term is an uncertainty brought in by $A$ but slightly reduced from its initial value (through correlation-building interaction, see immediately below) by the third term. Not included here is the entropy production from the final process of recording the result.

7.3 Build-up of correlations: a microperspective

It is easy to see, on the microlevel, the mechanism of correlation build-up for distributions of the type (52) when the number of variables is increased through increase of $x$. Let the probabilities for two different outcomes, $a$ and $b$, be $p$ and $1-p$, respectively. We consider only two variables ($\varepsilon = \pm 1$, $\varepsilon' = \pm 1$), where positive values strengthen $a$ and negative values strengthen $b$. Then the distribution of $\varepsilon$ and $\varepsilon'$ within the final state is
\begin{equation}
\frac{1}{4}(1 + \varepsilon\eta)(1 + \varepsilon'\eta)p + \frac{1}{4}(1 - \varepsilon\eta)(1 - \varepsilon'\eta)(1-p) =
\end{equation}
\begin{equation}
= \frac{1}{4}(1 + \varepsilon\varepsilon'\eta^2) + \frac{1}{4}(\varepsilon + \varepsilon')(2p-1)\eta.
\end{equation}
This gives the mean values,
\begin{equation}
\langle \varepsilon \rangle = \langle \varepsilon' \rangle = (2p-1)\eta. \quad \langle \varepsilon\varepsilon' \rangle - \langle \varepsilon \rangle \langle \varepsilon' \rangle = 4p(1-p)\eta^2.
\end{equation}
Thus, the mean values of $\varepsilon$ and $\varepsilon'$ are positive for $p > \frac{1}{2}$ and negative for $p < \frac{1}{2}$. Moreover, $\varepsilon$ and $\varepsilon'$ are positively correlated.

7.4 Orthogonality between scattering states and copied states

Let the copying part of $A$ consist of a lattice of $N$ similar charged spin particles. We assume the spins to be originally parallel, and we assume a fast charged particle belonging to the scattering state of $\mu$ to pass along the lattice. As a result of this all the spins are rotated by a small angle $\Delta \theta$. The squared modulus
of the scalar product of the new state (the state after copying) with the original state is typically
\[
(\cos(\frac{1}{2} \Delta \theta))^2N \approx (1 - \frac{1}{8} \Delta \theta^2)^2N \approx e^{-\frac{1}{4} N \Delta \theta^2},
\] (94)
which becomes extremely small for \( N \Delta \theta^2 \ll 1 \). This mechanism can thus justify the assumption
\[
|\langle 0, 0; \mathbf{e} | j, 0; \mathbf{e} \rangle|^2 = 0.
\] (95)
For different \( j \), spins belonging to lattices in different places are rotated, and the corresponding states are mutually orthogonal, as in (22) for \( x = x' = 0 \). This does not change for general \( x \) and \( x' \).

7.5 Sketch of a Gedankenexperiment

Let us discuss an outline of an experiment that would, in principle, simulate one step (out of very many stochastic steps) in a measurement process. The structure of such a set-up is shown in Figure 2.

A wave-packet of a spinless atom is entering at \( A \) and excited at \( B \) to a spin-1 state, then split at \( C \) by an inhomogenous magnetic field into \( S_z \)-components. The directions of motion of these component wave-packets are changed so that they are all running in parallel through \( D \), where a magnetic field in the \( z \)-direction splits their energy into different levels. \( D \) is assumed to be long enough so that the atom returns to the ground state within \( D \). The split wave-packet continues to \( E \) where detectors register arrival of the particle in either of the channels.

Due to the larger (smaller) phase space available for deexcitation of the component with higher (lower) energy, this component is enhanced (suppressed). Thus, if the amplitudes for \( S_z = -1, 0, +1 \) without a magnetic field in \( D \), are
\[
\psi_-, \psi_0, \psi_+ \quad (|\psi_-|^2 + |\psi_0|^2 + |\psi_+|^2 = 1),
\] (96)
then in the presence of a magnetic field \((0, 0, B)\) in \( D \), the probabilities for detection change from \(|\psi_-|^2, |\psi_0|^2, |\psi_+|^2\) into
\[
\frac{(|\psi_-|^2(1 - \alpha B), |\psi_0|^2, |\psi_-|^2(1 + \alpha B))}{1 + \alpha B(|\psi_+|^2 - |\psi_-|^2)},
\] (97)
where \( \alpha \) is a constant.

Here we have assumed the decay amplitude to be constant within the relevant range of momentum for the decay photon. Then the phase-space factor is proportional to
\[
\int \frac{d^3k}{|k|} \delta(|k| - \ldots)|M(k)|^2,
\] (98)
which is proportional to $1 - \alpha B$, 1 and $1 + \alpha B$, respectively, for the three cases. This is the basis of Equation (97).

It is important to note that the probability per unit time for deexcitation is proportional to

$$1 + \alpha B (|\psi_+|^2 - |\psi_-|^2),$$

which is the denominator of (97). Therefore, in a situation where $\varepsilon = B/|B|$ is $\pm 1$ with equal frequency, the average detection probabilities are unchanged,

$$|\psi_-|^2, \ |\psi_0|^2, \ |\psi_+|^2.$$  \hfill (100)

We can think of $\varepsilon$ as simulating one stochastic variable. Then the change from (100) into

$$\frac{(|\psi_-|^2(1 - \varepsilon \alpha |B|), |\psi_0|^2, |\psi_-|^2(1 + \varepsilon \alpha |B|))}{1 + \varepsilon \alpha |B|(|\psi_+|^2 - |\psi_-|^2)}$$  \hfill (101)

simulates one step in the random walk. The crucial feature here, in contrast to the stochastic variables $e_{jx}$ of (21), is that $\varepsilon$ is under the control of the experimenter.

It may seem paradoxical that the components formed at $B$ and split at $C$ with $a$ priori probabilities (100) given already at $B$, can be subject to "revision" by a later interaction taking place at $D$ changing the probabilities into (101).

The solution of the paradox is that in quantum mechanics, the entire process from $A$ to $E$ must be viewed as one whole. There is no observer intervening between $A$ and $E$.

Moreover, if the sign $\varepsilon$ of the magnetic field is unbiased, i.e., equally distributed between the values $\pm 1$, then the overall probabilities for detection at $E$, are those of (100), given already at $B$.

In this sense the Gedankenexperiment sketched here can be viewed as simulating one step in the measurement process as modelled in Section 6 above.

7.6 Quantum field theory and understanding measurement

One manifestation of the Bohr-Einstein debate on quantum measurement is the famous 1935 article by Einstein, Podolski and Rosen [1] and Bohr’s answer [2] to that. At the time of this debate, relativistic quantum field theory had not yet been developed. The ordering theorem, i.e., the expression of time-ordered products of quantum fields in terms of normal-ordered products [3], a basis of Feynman diagrams, was not yet available.

In the tradition that followed, the measurement problem was not stated in field-theoretical terms, but stayed within first-quantized theory. Similarly, quantum field theory was used to derive the measurable features of microworld processes, but not to describe the interface between microworld and macroworld. When the infrared divergences of quantum electrodynamics appeared, one had to let the final state include also soft bremsstrahlung [3]. Then the details of the
final state had to depend on the experimental resolution. In this work, one was quite close to a description taking into account also the measurement apparatus, but it did not really happen.

Acknowledgements

I thank Professor Kazimierz Rzazewski for discussions and for reading and criticizing an earlier version of this manuscripts. For discussions and technical support I also thank Anders Eriksson, Jens Fjelstad, Maria Grahn, Martin Nilsson-Jacobi, Kristian Lindgren and Carina Rehnstrom.

A Scattering process with sources and sinks

To describe a scattering process

\[ A + B \rightarrow C_1 + \ldots + C_m, \]  

we first consider two sources emitting the incoming particles at time \(-T\). We can think of them as one bilocal source creating the state \(|0\rangle_\mu\) of incoming particles \(A\) and \(B\). Let us similarly consider a set of multiple sinks ready to absorb and identify the states \(|1\rangle_\mu, \ldots, |n\rangle_\mu\) of outgoing particles \(C_1, \ldots, C_m\) at time \(T\).

We assume the dynamics to be described by renormalized quantum field theory, where an S-matrix element can be represented by a set of connected Feynman diagrams. Here we represent the whole sum over such diagrams by a shaded circle with ingoing and outgoing lines (Fig. 3).

We shall combine such diagrams with open half-circles, marked with the corresponding states, representing the source at time \(-T\) and the sinks at time \(T\) of the scattering states \(|j\rangle_\mu\) with the curved side as the active side, labelled by the emitted or absorbed state. However, rather than the transition amplitudes, we shall describe the density matrix of the outgoing state at time \(T\).

To get the density matrix, we connect the initial (time \(-T\)) state \(|0\rangle_\mu\) going into scattering, described by the scattering operator \(S\), whereas the adjoint state is taken into the scattering state by \(S^\dagger\). (This is a type of description used long ago by Kinoshita and Nakanishi. Since interaction Hamiltonians are hermitean and particle propagators are symmetric under time reversal, we get a whole series of diagrams involving emission, scattering and absorption and the inverse processes (Fig. 4)).

Taking together all diagrams, we find a geometrical series. The result can be viewed as the insertions of Fig. 5a and interpreted as renormalization of the emission process for the incoming state, and, similarly, renormalization of the
process of nothing at all happening (Fig. 5b). The final-state density matrix is

\[ F_j M_{j0} J_0 J_0^* M_{k0}^* F_k^* - F_j M_{j0} J_0 \left( \sum_{l=1}^{n} J_0^* M_{l0} F_l^* F_l M_{l0} J_0 \right) J_0^* M_{k0}^* F_k^* + \]

\[ + F_j M_{j0} J_0 \left( \sum_{l=1}^{n} J_0^* M_{l0}^* F_l^* F_l M_{l0} J_0 \right) J_0^* M_{k0}^* F_k^* + \ldots = \]

\[ = F_j M_{j0} J_0 \frac{1}{1 + \sum_{l=1}^{n} J_0^* M_{l0}^* F_l^* F_l M_{l0} J_0} J_0^* M_{k0}^* F_k^* = \]

\[ = \frac{|J_0|^2 F_j F_k^* M_{j0} M_{k0}^*}{1 + |J_0|^2 \sum_{l=1}^{n} |F_l|^2 |M_{l0}|^2} \]

where

\[ J_0, F_1, F_2, ..., F_n \]

represent the source and the assembly of sinks, respectively.

The term '1' in the denominator of the last expression in (103) represents the case of nothing happening. With a sufficiently strong source, it can be safely neglected. Then the source contributes identical factors in numerator and denominator, and the result reduces to

\[ \frac{F_j F_k^* M_{j0} M_{k0}^*}{\sum_{l=1}^{n} |F_l|^2 |M_{l0}|^2} \]

This depends on the absorption efficiencies of the sinks. Assuming for a moment these factors to be equal, we have again the same factors in numerator and denominator. The result is the normalized final state density matrix

\[ \frac{M_{j0} M_{k0}^*}{\sum_{l=1}^{n} |M_{l0}|^2} \]

which was our starting point in (11) or (17). Thus the non-linearity of (11) has been explained. The matrix elements \( M_{j0} \) are related to \( \psi_j \) through (13).

The interpretation of (106) is that it describes the case without a measurement apparatus. For a realistic measurement apparatus, the factors \( F_j \) of (105) are in general different and unknown, except for the restriction that the statistical distribution over them should not introduce any bias. Thus, we can identify them with \( B_j(\mathbf{e}) \) of Section 4.
B Factorization of final state interaction

We think of the interaction between the quantum system $\mu$ and the measurement apparatus $A$ as an electromagnetic interaction with very small energy and momentum transfer. Thus it can be described in terms of an exchange of soft photons. Emission and exchange of soft photons is an old and well-known example of factorizable processes in quantum electrodynamics. When it became understood, the picture of scattering became drastically changed, in the sense that no non-forward scattering takes place without soft-photon emission. Later, this was identified as coherent radiation from classical charged point sources moving into and out from a point-like scattering centre.

To show the factorization of soft photon emission and exchange, we consider an outgoing electron (charge $-e$, mass $m$) with final momentum $p$, described by a spinor $\bar{u}(p)$,

$$p^2 = m^2; \quad \bar{u}(p)(ip \cdot \gamma - m) = 0,$$

after emitting two soft photons with momenta $k_1, k_2$ and polarizations $\tau_1, \tau_2$,

$$k_1^2 = k_2^2 = 0; \quad k_1 \cdot \tau_1 = k_2 \cdot \tau_2 = 0; \quad |k_1|, |k_2| \ll m. \quad (108)$$

In the evaluation of the Feynman diagram of Fig. 5, the spinor $\bar{u}(p)$ for the outgoing electron is then replaced by an expression

$$e^2\bar{u}(p) \left[ \tau_1 \cdot \gamma \frac{i(p + k_1) \cdot \gamma + m}{(p + k_1)^2 + m^2}\tau_2 \cdot \gamma + (1 \leftrightarrow 2) \right] \frac{i(p + k_1 + k_2) \cdot \gamma + m}{(p + k_1 + k_2)^2 + m^2} =$$

$$= e^2 \frac{1}{2(p \cdot k_1 + p \cdot k_2)} \bar{u}(p) \times$$

$$\times \left[ \tau_1 \cdot \gamma \frac{(ip \cdot \gamma + m)\tau_2 \cdot \gamma (ip \cdot \gamma + m)}{2p \cdot k_1} + (1 \leftrightarrow 2) \right] =$$

$$= e^2 \frac{-p \cdot \tau_1 p \cdot \tau_2}{p \cdot k_1 + p \cdot k_2} \left( \frac{1}{p \cdot k_1} + \frac{1}{p \cdot k_2} \right) \bar{u}(p) = (s(k_1) \cdot \tau_1)(s(k_2) \cdot \tau_2)\bar{u}(p), \quad (109)$$

where

$$s_\mu(k) = -e \frac{ip_\mu}{p \cdot k} = -e \int_0^{\infty} dt \int d^3x e^{i(k \cdot x - |k|t)} \delta^3 \left( x - \frac{p}{p_0} t \right) \frac{p_\mu}{p_0} \quad (110)$$

is the Fourier transform of the current of a classical point charge $-e$ moving from $x = 0$ at time zero with the velocity $p/p_0$. The rest of the diagram is unchanged in the limit of small $k_1, k_2$. Equation (109) states that the emission of the two photons is described by one scalar emission factor for each photon. The corresponding holds for two photons being absorbed by an electron, as well as for one emitted photon and one absorbed.

24
For $r$ photons, use can be made of the identity

$$\sum_{(i_1 i_2 \ldots i_m)} \frac{1}{a_{i_1} (a_{i_1} + a_{i_2}) \ldots (a_{i_1} + a_{i_2} + \ldots + a_{i_m})} = \frac{1}{a_1 a_2 \ldots a_m}. \quad (111)$$

There is also a factor $(r!)^{-1}$. Summation over photon states and over $r$ gives rise to a coherent state generated by the classical current (110).

References

[1] B. Podolski A. Einstein and N. Rosen. *Phys. Rev.*, 47:777, 1935.

[2] N. Bohr. *Phys. Rev.*, 48:696, 1935.

[3] K.-E. Eriksson. Measurement as soft final-state interaction with a stochastic system. *arXiv:quant-ph/0206099 v2*, 2002.

[4] N. Gisin. *Phys. Rev. Letters*, 52:1657, 1984.

[5] N. Gisin and I. Percival. *J. Phys.*, A25:5677, 1992.

[6] J. M. Jauch and F. Rohrlich. *Helv. Phys. Acta*, 27:613, 1954.

[7] N. Nakanishi. *Progr. Theor. Physics (Japan)*, 19:159, 1958.

[8] G. C. Wick. *Phys. Rev.*, 80:268, 1950.
Figure 1: The distribution over pointer variables $\mathbf{z} = (z_1, z_2, z_3)$ for a three-state system is centered around $\mathbf{z}^{(1)}$, $\mathbf{z}^{(2)}$ and $\mathbf{z}^{(3)}$ in the plane $z_1 + z_2 + z_3 + \sqrt{Z} = 0$ with relative weights $|\psi_1|^2$, $|\psi_2|^2$ and $|\psi_3|^2$. For $b \ll a$, i.e., for large $Z$, the partial distributions are well separated. With increasing $Z$, the transition from a unimodal to a trimodal distribution takes place around $Z \approx 1$. 
Figure 2: Sketch of a Gedankenexperiment to simulate the first step in a measurement process, as described in subsection 7.5. An atomic beam is entering at A and excited by radiation at B into a spin-1 state, then split at C into the different $z$-components of the spin. In D there is a homogenous magnetic field $(0, 0, B)$, which makes the available phase space in the relaxation to the ground state of the three $S_z$-components different. After deexcitation in D, the atoms are detected at E.
Figure 3: Diagram elements for considering scattering, with production of initial particle state $|0\rangle_\mu$ and absorption of final particle state $|j\rangle_\mu$ included in the description.
Figure 4: Diagrams for the $jk$-component of the final-state density matrix, $\rho^{(f)}_{jk}$. 
Figure 5: (a) Insertions for the initial emission process into density-matrix diagrams. (b) Diagrams for the case of no process.
Figure 6: Feynman diagram for the emission of two soft photons.