Bidirectional classical stochastic processes with measurements and feedback

G. E. Hahne *†
P. O. Box 2748, Sunnyvale, CA 94087-0748 USA

NAS Technical Report NAS-06-002

January 2006

PACS Numbers: 02.50.Ey, 02.50.Ga, 03.65.Ta

Abstract

A measurement on a quantum system is said to cause the "collapse" of the quantum state vector or density matrix. An analogous collapse occurs with measurements on a classical stochastic process. This paper addresses the question of describing the response of a classical stochastic process when there is feedback from the output of a measurement to the input, and is intended to give a simplified model for quantum-mechanical processes that occur along a space-like reaction coordinate. The classical system can be thought of in physical terms as two counterflowing probability streams, which stochastically exchange probability currents in a way that the net probability current, and hence the overall probability, suitably interpreted, is conserved. The proposed formalism extends the mathematics of those stochastic processes describable with linear, single-step, unidirectional transition probabilities, known as Markov chains and stochastic matrices. It is shown that a certain rearrangement and combination of the input and output of two stochastic matrices of the same order yields another matrix of the same type. Each measurement causes the partial collapse of the probability current distribution in the midst of such a process, giving rise to calculable, but non-Markov, values for the ensuing modification of the system's output probability distribution. The paper concludes with an analysis of a simple classical probabilistic version of a so-called grandfather paradox.

1 Introduction

In a previous paper [1], the author proposed a formalism for describing the evolution of a Schrödinger wave function for a single particle along a spacelike

*Ames Associate, retired from NASA/Ames Research Center
†email: ghahne@mail.arc.nasa.gov

reaction coordinate, where the time was taken as one of the transverse coordinates. The principal objective of that study was to establish a version of one-body nonrelativistic quantum mechanics in which the time plays a natural role as an operator/observable. The present work is a preliminary attempt to address another aspect of the formalism in [1] (mentioned in Sec. 4 therein), that is, is there a self-consistent theory of measurement in a quantum mechanics in which there is feedback from the outcome of a measurement to the input to the measurement? This problem does not occur in conventional quantum mechanics when the time is the coordinate of evolution (but see [2] for an extension of quantum mechanics that can describe physical processes with feedback in the time dimension). The present study is preliminary in that we shall propose, and analyze the consistency of, only a classical analog to such a quantum system and of certain measurements upon it.

We can represent a physical system by a point moving in a transverse space, such that the point's transverse position is a function (which is, in general, multi-valued) that is parametrized by a spacelike coordinate of evolution, and such that both forward and backward motion along the evolution coordinate are possible physically and are distinguished mathematically. We divide the transverse space, which can include the time direction, by a fine-grained mesh into a large number of boxes, n^F of which are associated with forward motion, and n^B with backward motion. In turn, the elements of the former are grouped into subsets called coarse-grained boxes, and similarly for the latter. The system point can jump from any one box to another, including to/from forward from/to backward motion, in one of a sequence of lumped zones of interaction. We shall not attempt to give concrete realizations to the physics of these zones, but merely presume a given set of transition probabilities associated with each zone. A second type of zone will represent measurements: We assume that measurement zones do not give rise to forward↔backward transitions, but merely distinguish an incoming signal according to which of the coarse-grained boxes detected the trajectory's transit in each instance that the system point passes through such a zone.

Classical mechanics, probability theory and the associated stochastic processes have simplifying advantages over quantum systems: (i) there are no interference phenomena in combining sub-processes, (ii) closed channels—domains inaccessible to classical mechanical systems—do not carry probability currents, and (iii) a measurement of the first kind (in Pauli's sense, [3], p. 75) affects a probability distribution for a single system, but, insofar as we sum over all possible measurement outcomes, does not affect the distribution for an ensemble of systems. There are many textbooks on the subject of probability and stochastic processes, e.g. [4], [5], [6], [7], [8], and specifically on Markov processes, e.g. [9] Ch. XV, [10], [11], [12] Ch. 7, [13], [14] Ch. 6.

We shall keep to the analysis of state spaces that are discrete, and adhere to the nomenclature recommended in [6], p. 188, Table 6.1: the entities to be studied will be called Markov *chains*; Markov *processes* deal with continuous state spaces, as in Brownian motion. Markov chains can be associated with either a discrete or a continuous evolution parameter, which is usually—but herein not necessarily—understood to be the time. This paper will deal only with chains with a discrete evolution coordinate, and with a finite state space.

An $n \times n$ matrix A, with exclusively nonnegative real entries A_{jk}, is called *stochastic*, if its *column* sums are all equal to 1, that is

$$\sum_{j=1}^{n} A_{jk} \;=\; 1, \quad \text{for } k = 1, 2, \ldots, n. \tag{1}$$

This notation is the transpose of that often used in mathematical texts, (see, for example, [14], Ch. 6.1), and is adopted in order to facilitate the eventual comparision to formulas in quantum mechanics. In fact, let U be an n-component column ("state") vector with nonnegative components U_k, and let the components of the n-vector V be

$$V_j \;=\; \sum_{k=1}^{n} A_{jk} U_k; \tag{2}$$

then if also

$$\sum_{k=1}^{n} U_k \;=\; 1, \tag{3}$$

the property (1) entails

$$\sum_{j=1}^{n} V_j \;=\; 1. \tag{4}$$

A stochastic matrix as A is often called a matrix of transition probabilities.

The product of two stochastic matrices of the same order is also stochastic ([14], Ch. VI, Th. 1.1(d)). The present work is based on a result derived below: there is another way to combine two stochastic matrices, which depends on a kind of rearrangement of what constitutes input and what output, and that also yields a stochastic matrix, provided that the combining process converges.

The remainder of this paper is organized as follows. In section 2, we shall define, and exhibit some properties of, a nonconventional synthesis of stochastic matrices that yields another stochastic matrix. These results will be shown to describe what will be called "bidirectional Markov chains" in that both input and output to the process will occur at both ends of an interval in a coordinate describing the evolution of a stochastic process. In section 3, we shall study the effect of a sequence of "coarse-grained" intermediate measurements of the first kind on the predicted output of such a system; the transition matrix for such a measurement depends on the input, and hence the measurement process is a nonlinear, i.e., non-Markovian, mapping of input into output. Section 4 completes the paper with a comparision of conditional probabilities with the collapse of a probability distribution, and analyzes a classical-probabilistic version of a grandfather paradox.

2 Bidirectional Markov chains

In this section we shall deconstruct stochastic matrices of the same order in a parallel manner and fashion a synthesis of such matrices in a way that yields another stochastic matrix that is not the matrix product of its ingredient matrices.

Let A be as in (1), and let n^F, n^B be an integer partition of n:

$$n^F + n^B = n. \tag{5}$$

The integer n^F will be the number of forward- (F)-propagating components, or channels, of the probability state vector, while n^B will be the number of backward- (B)-propagating channels. Accordingly, we can partition A as

$$A = \begin{bmatrix} (A^{FF})_{\alpha\beta} & (A^{FB})_{\alpha b} \\ (A^{BF})_{a\beta} & (A^{BB})_{ab} \end{bmatrix}, \tag{6}$$

where the subscripts stand for a complete array of sub-matrix elements, with $\alpha, \beta = 1, 2, \ldots, n^F$ and $a, b = 1, 2, \ldots, n^B$. Consistent with (2), the causal order in the superscripts on the sub-blocks of A are to be read from right to left; similarly, the input to output channels are ordered from right to left in the subscripts.

Let us introduce two more $n \times n$ stochastic matrices C and M, partitioned as A, with C as well as A nonspecial; the M stands for "measurement". For the purposes of this section, the matrix M is presumed to be a fixed stochastic matrix

$$M = \begin{bmatrix} (M^{FF})_{\alpha\beta} & 0 \\ 0 & (M^{BB})_{ab} \end{bmatrix}, \tag{7}$$

where M^{FF} and M^{BB} are nonspecial $n^F \times n^F$ and $n^B \times n^B$ stochastic matrices, respectively. The zeroes for the off-diagonal blocks of M correspond to the physical assumption that a measurement process gives rise to no $F \leftrightarrow B$ transitions.

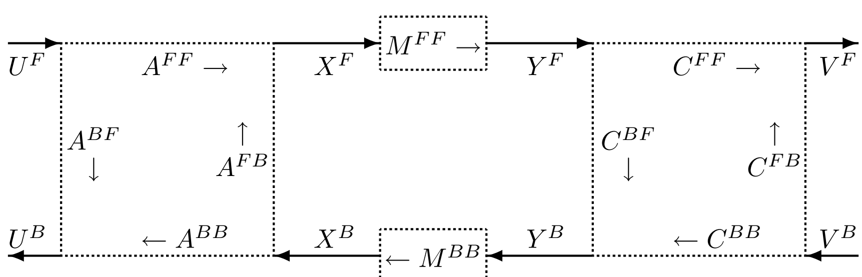

Figure 1. Probability flows and transitions.

Figure 1 shows the hookups of the channels carrying probability: U^F, X^F, Y^F and V^F are the vectors of forward-flowing probability, with n^F channels each, as in $(U^F)_\alpha,\ \alpha = 1, 2, \ldots, n^F$; V^B, Y^B, X^B and U^B are the vectors of backward-flowing probability, with n^B components each, as in $(V^B)_a,\ a = 1, 2, \ldots, n^B$. The vectors U^F and V^B are the prescribed input, while V^F and U^B comprise the overall output. The arrows within the boxes A, M, and C represent the action of the sub-matrices in diverting and mixing the flow of probability current vectors.

There is a feedback loop given by the causal sequence X^F, M^{FF}, Y^F, C^{BF}, Y^B, M^{BB}, X^B, and A^{FB}. In what follows, we shall consider the time to be one of the transverse directions: Although conventional physical systems can be in only one spatial position at a given time, such systems can be at more than one time in a given spatial position; therefore, we must consider the circumstance that the system passes through the feedback loop either zero, or one, or two, etc., times before exiting.

We define I^{FF} and I^{BB} to be the $n^F \times n^F$ and $n^B \times n^B$ unit matrices, respectively. Also, let $\bar{R}^F$ and $\bar{R}^B$ be single-rowed matrices with n^F and n^B columns, respectively, all entries being 1 in both matrices, e.g.,

$$(\bar{R}^F)_\alpha \;=\; 1, \quad \text{for } \alpha = 1, 2, \ldots, n^F. \tag{8}$$

Then, since A, M, and C are stochastic matrices, we have

$$\bar{R}^F A^{FF} + \bar{R}^B A^{BF} \;=\; \bar{R}^F, \tag{9a}$$
$$\bar{R}^F A^{FB} + \bar{R}^B A^{BB} \;=\; \bar{R}^B, \tag{9b}$$
$$\bar{R}^F M^{FF} \;=\; \bar{R}^F, \tag{9c}$$
$$\bar{R}^B M^{BB} \;=\; \bar{R}^B, \tag{9d}$$
$$\bar{R}^F C^{FF} + \bar{R}^B C^{BF} \;=\; \bar{R}^F, \tag{9e}$$
$$\bar{R}^F C^{FB} + \bar{R}^B C^{BB} \;=\; \bar{R}^B. \tag{9f}$$

The "equations of motion" of the system in Fig. 1 are as follows:

$$X^F \;=\; A^{FF} U^F + A^{FB} X^B, \tag{10a}$$
$$U^B \;=\; A^{BF} U^F + A^{BB} X^B, \tag{10b}$$
$$Y^F \;=\; M^{FF} X^F, \tag{10c}$$
$$X^B \;=\; M^{BB} Y^B, \tag{10d}$$
$$V^F \;=\; C^{FF} Y^F + C^{FB} V^B, \tag{10e}$$
$$Y^B \;=\; C^{BF} Y^F + C^{BB} V^B. \tag{10f}$$

Let us check first that overall probability current is conserved in the transition from input to output. We apply (9a) and (9b) to the sums-over-channels of corresponding sides of (10a) and (10b), and apply similarly (9c) to (10c), (9d) to (10d), and (9e) and (9f) to the sum of (10e) and (10f), thereby obtaining

$$\begin{aligned}
\bar{R}^F X^F + \bar{R}^B U^B &= \bar{R}^F U^F + \bar{R}^B X^B, \\
\bar{R}^F Y^F &= \bar{R}^F X^F, \\
\bar{R}^B X^B &= \bar{R}^B Y^B, \\
\bar{R}^F V^F + \bar{R}^B Y^B &= \bar{R}^F Y^F + \bar{R}^B V^B.
\end{aligned}$$

Therefore, the apparatuses A, M, and C conserve probability current, so that we have the combined result

$$\bar{R}^F V^F + \bar{R}^B U^B = \bar{R}^F U^F + \bar{R}^B V^B. \tag{12}$$

We also need to prove that, whenever the components of the input vectors U^F and V^B are nonnegative, the components of the output vectors V^F and U^B are likewise nonnegative. It is convenient to obtain this result from an explicit form for the matrix mapping input into output. We can use (10c) and (10d) to eliminate Y^F and X^B from (10a), (10b), (10e), (10f), then use the first and fourth of the latter set to obtain X^F and Y^B in terms of U^F and V^B, and then use the second and third to obtain V^F and U^B in terms of U^F and V^B. We first define the auxiliary matrices L^{FF} and L^{BB}, where L stands for “loop”:

$$L^{FF} = \left[I^{FF} - A^{FB} M^{BB} C^{BF} M^{FF}\right]^{-1}, \tag{13a}$$

$$L^{BB} = \left[I^{BB} - C^{BF} M^{FF} A^{FB} M^{BB}\right]^{-1}. \tag{13b}$$

The product matrices $A^{FB} M^{BB} C^{BF} M^{FF}$ and $C^{BF} M^{FF} A^{FB} M^{BB}$ are assumed to be sufficiently close to the $n^F \times n^F$ and $n^B \times n^B$ zero matrices, respectively, so that the inverses in (13) exist. Then we have

$$X^F = L^{FF}\left[A^{FF} U^F + A^{FB} M^{BB} C^{BB} V^B\right], \tag{14a}$$

$$Y^B = L^{BB}\left[C^{BF} M^{FF} A^{FF} U^F + C^{BB} V^B\right]. \tag{14b}$$

We now represent the complete mapping of input into output by an $n \times n$ matrix S, such that

$$\begin{bmatrix} V^F \\ U^B \end{bmatrix} = \begin{bmatrix} S^{FF} & S^{FB} \\ S^{BF} & S^{BB} \end{bmatrix} \begin{bmatrix} U^F \\ V^B \end{bmatrix}. \tag{15}$$

Then (10) and (14) entail

$$S^{FF} = C^{FF} M^{FF} L^{FF} A^{FF}, \tag{16a}$$

$$S^{FB} = C^{FF} M^{FF} L^{FF} A^{FB} M^{BB} C^{BB} + C^{FB}, \tag{16b}$$

$$S^{BF} = A^{BF} + A^{BB} M^{BB} L^{BB} C^{BF} M^{FF} A^{FF}, \tag{16c}$$

$$S^{BB} = A^{BB} M^{BB} L^{BB} C^{BB}. \tag{16d}$$

Given convergence, (13a) and (13b) can be represented in infinite series expansions

$$L^{FF} = I^{FF} + \sum_{m=1}^{\infty} \left(A^{FB}M^{BB}C^{BF}M^{FF}\right)^m, \tag{17a}$$

$$L^{BB} = I^{BB} + \sum_{m=1}^{\infty} \left(C^{BF}M^{FF}A^{FB}M^{BB}\right)^m. \tag{17b}$$

Each summand in these series consists of products of matrices with nonnegative elements, so that all the elements of both L^{FF} and L^{BB} are nonnegative. In turn, we infer from (16) that all the elements of the matrix S are nonnegative, and (12) implies that the column sums of S are all $+1$; hence, S is a stochastic matrix mapping overall input into overall output. We have thereby achieved the principal goal of this section.

We remark further on the convergence of the sums for L^{FF} and L^{BB}. These sums represent the complete feedback loop taken at the entry points at the beginning of X^F, and at the beginning of Y^B, respectively, in Fig. 1. Both A and C have $n(n-1)$ free parameters, subject to the inequalities that each element is nonnegative and less than 1. If one or both off-diagonal blocks A^{FB} and C^{BF} have all elements sufficiently small the sum in (17) will converge: In fact, suppose that

$$\left(C^{BF}M^{FF}A^{FB}M^{BB}\right)_{ab} = \epsilon D_{ab}, \tag{18}$$

where $0 \leq \epsilon < 1$, and where

$$\sum_{a=1}^{n^B} D_{ab} \leq 1, \quad \text{for } b = 1, 2, \ldots, n^B. \tag{19}$$

Then we have

$$\sum_{a=1}^{n^B} (D^2)_{ab} = \sum_{a,c=1}^{n^B} D_{ac}D_{cb} \leq \sum_{c=1}^{n^B} D_{cb} \leq 1. \tag{20}$$

Similarly, the column sums of $(D^r)_{ab}$, and therefore the individual elements of D^r, are majorized by 1. Hence each element in the matrix sum in (17b) converges. Since we have

$$\begin{aligned}&\left(A^{FB}M^{BB}C^{BF}M^{FF}\right)^m \\ &\quad = A^{FB}M^{BB}\left(C^{BF}M^{FF}A^{FB}M^{BB}\right)^{m-1}C^{BF}M^{FF}, \ m = 2, 3, \ldots\end{aligned} \tag{21}$$

the sum in (17a) also converges. It can be arranged that divergence occurs as the upper bound ϵ approaches 1. Since no element of D can exceed 1, the set of matrices A and C that give rise to a divergent sum will be of lower dimensionality, a kind of boundary set within the set of all possible stochastic matrices A and C. There is no full $2n(n-1)$-dimensional domain of stochastic matrices A and C that yields divergent feedback processes.

3 Intermediate measurements

In this section, we shall partition both the n^F F-type and the n^B B-type channels into coarse-grained subsets, and presume that at each crossing of the system point through the device M, a measurement is made of which coarse grain the system traversed. We shall also presume, analogous to the quantum-mechanical postulate of Lüders ([15] Eq. (7)—see also [16], Eq. (43), et seq., and [17], Eq. (2)), that within a coarse grain the measurement process does not change the relative strengths of the probability currents in the constituent fine-grained channels. The latter assumption entails a nonlinearity in the stochastic process, at least for individual trials—see the analysis below. As in Sec. 4, the above model for the effect of a coarse-grained measurement is straightforwardly related to the usual constructs for conditional probabilities. Note that we are assuming a measurement to be of the first kind ([3], p. 75), that is, a second measurement of the coarse-grained structure of the current vector following the first would yield the same result as the first—the measurement has no dynamical effect on the system but only determines the coarse-grained state of the system.

Let $\{n_1^F, n_2^F, \ldots, n_{\nu^F}^F\}$ and $\{n_1^B, n_2^B, \ldots, n_{\nu^B}^B\}$ be ordered sequences of positive integers partitioning n^F and n^B, respectively, with partial sums

$$p_\nu^F = \begin{cases} 0, & \text{if } \nu = 0, \\ \sum_{\nu'=1}^{\nu} n_{\nu'}^F, & \text{if } \nu = 1, 2, \ldots, \nu^F, \end{cases} \tag{22a}$$

$$p_\nu^B = \begin{cases} 0, & \text{if } \nu = 0, \\ \sum_{\nu'=1}^{\nu} n_{\nu'}^B, & \text{for } \nu = 1, 2, \ldots, \nu^B. \end{cases} \tag{22b}$$

We also define the projection matrices

$$\left(I_\nu^{FF}\right)_{\alpha\beta} = \begin{cases} \delta_{\alpha\beta}, & \alpha, \beta = p_{\nu-1}^F + 1, \ldots, p_\nu^F, \text{ for each } \nu = 1, \ldots, \nu^F, \\ 0, & \text{otherwise}, \end{cases} \tag{23a}$$

$$\left(I_\nu^{BB}\right)_{ab} = \begin{cases} \delta_{ab}, & a, b = p_{\nu-1}^B + 1, \ldots, p_\nu^B, \text{ for each } \nu = 1, \ldots, \nu^B, \\ 0, & \text{otherwise}. \end{cases} \tag{23b}$$

It is also convenient to define certain probabilities that are associated with the circumstance that an input current to an apparatus is presumed, or known after the fact, to be diverted into a particular subset of the set of output channels. In fact, let Ω be the matrix associated with an apparatus such as A, M, or C, and let $\Omega^{\Xi\Xi'}$ be the subblocks with labels $\Xi, \Xi' = F, B$, and let input and output vectors be $W_{\text{in}}^{\Xi'}$ and W_{out}^{Ξ}, respectively. These are related by

$$W_{\text{out}}^{\Xi} = \Omega^{\Xi\Xi'} W_{\text{in}}^{\Xi'} / N\left(\Omega^{\Xi\Xi'}, W_{\text{in}}^{\Xi'}\right), \tag{24}$$

where N is a normalization factor. Note that all the input probability current is presumed to be allocated to the subset of Ξ-type channels, and in the same fine-grained proportion that is governed by the input and by the matrix Ω. In order to conserve probability current we must have

$$N\big(\Omega^{\Xi\Xi'}, W_{\rm in}^{\Xi'}\big) \;=\; \bar{R}^{\Xi}\Omega^{\Xi\Xi'}W_{\rm in}^{\Xi'}/\bar{R}^{\Xi'}W_{\rm in}^{\Xi'}. \tag{25}$$

Then a detection in the $\nu^{\rm th}$ F-type coarse-grained channel entails the following mapping of X^F into Y^F:

$$Y^F \;=\; M_\nu^{FF}\big(X^F\big) \;=^{\rm def}\; I_\nu^{FF}X^F/N\big(I_\nu^{FF}, X^F\big). \tag{26}$$

Similarly, a detection in the $\nu^{\rm th}$ B-type channel entails the following mapping:

$$X^B \;=\; M_\nu^{BB}\big(Y^B\big) \;=^{\rm def}\; I_\nu^{BB}Y^B/N\big(I_\nu^{BB}, Y^B\big). \tag{27}$$

We now specify the set of possible scenarios (i.e., the sample space) afforded by the above-described system (cf. Fig. 1), and assign a transition matrix to each scenario. The physical system's trajectory is taken to be continuous, so that in progressing from an input to an output channel, a sequence of measurements of F and B type can be assigned; we need not consider, for example, two successive F-type measurements in the sample space. In particular, we have for U^F to V^F transitions the possible sequences

$$\begin{aligned}
&U^F \to M_{\nu_0}^{FF} \to V^F, \\
&U^F \to M_{\nu_0}^{FF} \to M_{\nu_{1'}}^{BB} \to M_{\nu_1}^{FF} \to V^F, \\
&\text{etc.,}
\end{aligned} \tag{28}$$

where the "etc." stands for two, three, and so on, times around the feedback loop. For U^F to U^B transitions we have the sequences

$$\begin{aligned}
&U^F \to U^B, \\
&U^F \to M_{\nu_0}^{FF} \to M_{\nu_{0'}}^{BB} \to U^B, \\
&U^F \to M_{\nu_0}^{FF} \to M_{\nu_0'}^{BB} \to M_{\nu_1}^{FF} \to M_{\nu_{1'}}^{BB} \to U^B, \\
&\text{etc.}
\end{aligned} \tag{29}$$

For V^B to V^F transitions we have

$$\begin{aligned}
&V^B \to V^F, \\
&V^B \to M_{\nu_{0'}}^{BB} \to M_{\nu_0}^{FF} \to V^F, \\
&V^B \to M_{\nu_{0'}}^{BB} \to M_{\nu_0}^{FF} \to M_{\nu_{1'}}^{BB} \to M_{\nu_1}^{FF} \to V^F, \\
&\text{etc.,}
\end{aligned} \tag{30}$$

and for V^B to U^B the sequences

$$\begin{aligned}
&V^B \to M_{\nu_{0'}}^{BB} \to U^B, \\
&V^B \to M_{\nu_{0'}}^{BB} \to M_{\nu_1}^{FF} \to M_{\nu_1'}^{BB} \to U^B, \\
&\text{etc.}
\end{aligned} \tag{31}$$

We assume that both the ν^F and, separately, the ν^B coarse-grained channels cover a large number of adjacent, but nonoverlapping, time windows, and that each passage through the apparatuses A and C takes a substantial positive amount of time, such that in practice it is possible to take an originally unordered set of measurements and infer the physical sequence of coarse grains at which the system trajectory crossed the measuring apparatus M. It is, therefore, possible to assign to each set of measured data uniquely to one of the processes listed in (28)–(31). (We shall consider below the case that this sequencing of the measured data is not feasible.) This construction entails the result that most of the possible sequences of measurements in (28)–(31) will have zero probability of occurring: in the third line of (29), for example, if the time interval associated with $M^{FF}_{\nu_1}$ is the same, or earlier than, the time interval associated with $M^{BB}_{\nu'_0}$, the probability of this sequence occurring is zero. These zero probabilities are built into the dynamics of the system by the matrices A and C, and need not be invoked as separate hypotheses.

We also assume that the matrices A^{FB} and/or C^{BF} have sufficiently small elements such that the probability that the system undergoes more than a moderate number of feedback loops is small, that is, that the sums in (17a) and (17b) converge rapidly.

Let us describe in detail a nontrivial example, from which one can infer a rule of calculation for any of the scenarios listed or implied in (28)–(31). Let us calculate the conditional probability that a trajectory enters, occasions a chosen sequence of exactly four measurements, and exits, as in the third line of (29); the trajectory is presumed to cross $M^{FF}, M^{BB}, M^{FF}, M^{BB}$ in the sequence of coarse-grained windows labeled $\nu_0, \nu_{0'}, \nu_1, \nu_{1'}$, respectively. We then have ten stages connected by nine processes, defined recursively as follows:

symbol/stage	definition	
U^F	*input*	(32a)
X^F_0	$A^{FF}U^F/N\big(A^{FF}, U^F\big)$	(32b)
Y^F_0	$I^{FF}_{\nu_0}X^F_0/N\big(I^{FF}_{\nu_0}, X^F_0\big)$	(32c)
$Y^B_{0'}$	$C^{BF}Y^F_0/N(C^{BF}, Y^F_0\big)$	(32d)
$X^B_{0'}$	$I^{BB}_{\nu_{0'}}Y^B_{0'}/N(I^{BB}_{\nu_{0'}}, Y^B_{0'}\big)$	(32e)
X^F_1	$A^{FB}X^B_{0'}/N\big(A^{FB}, X^B_{0'}\big)$	(32f)
Y^F_1	$I^{FF}_{\nu_1}X^F_1/N(I^{FF}_{\nu_1}, X^F_1\big)$	(32g)
$Y^B_{1'}$	$C^{BF}Y^F_1/N\big(C^{BF}, Y^F_1\big)$	(32h)
$X^B_{1'}$	$I^{BB}_{\nu_{1'}}Y^B_{1'}/N(I^{BB}_{\nu_{1'}}, Y^B_{1'}\big)$	(32i)
$U^B_1\big(\nu_{1'}, \nu_1, \nu_{0'}, \nu_0; U^F\big)$	$A^{BB}X^B_{1'}/N\big(A^{BB}, X^B_{1'}\big)$.	(32j)

(The first four arguments of U^B_1 signify the prescribed measurement sequence in order from right to left; the subscript 1 means that the trajectory traverses exactly one complete feedback loop in getting from U^F to U^B.) Should any one of the above-computed rhs's—apart from the normalization factors—be zero, we stop the computation and say that the given scenario cannot happen, that is, it is impossible for the system trajectory to enter at U^F, give rise to the specified sequence of measured values, and exit at U^B with nonzero current. For each specified sequence of coarse-grained measurements, therefore, we have a well-defined output: if the sequence is not a possible process we put $U^B = 0$, and if the sequence is possible, we obtain an output with the same normalization as the input,

$$\bar{R}^B U_1^B\big(\nu_{1'}, \nu_1, \nu_{0'}, \nu_0; U^F\big) \;=\; \bar{R}^F U^F. \tag{33}$$

The output U^B in (32j) is a conditional probability, that is, given the input vector U^F and given the sequence of four measurements, it describes the distribution of fine-grained probabilities of outcomes U^B. We shall now propose how to determine the probabilities for the fine-grained transitions $U^F \to U^B$ of (32), such that the sequence of measurements is not preassigned but is in a sense part of the output: that is, what is the probability, with input U^F as the only "given", that the outcome comprises exactly the sequence of measurements of (32) followed by the trajectory exiting in one of the fine-grained channels of U^B? Our claim is that the result is the product of the U^B of (32j) with the nine normalization factors of the rhs's of (32b)–(32j), that is,

$$\begin{aligned} U_1^B[\nu_{1'}, \nu_1, \nu_{0'}, \nu_0]\big(U^F\big) \;&=\; N\big(A^{FF}, U^F\big)\ldots N\big(A^{BB}, X_{1'}\big) \\ &\qquad \times U_1^B\big(\nu_{1'}, \nu_1, \nu_{0'}, \nu_0; U^F\big) &\text{(34a)} \\ &=\; A^{BB} I^{BB}_{\nu_{1'}} C^{BF} I^{FF}_{\nu_1} A^{FB} I^{BB}_{\nu_{0'}} C^{BF} I^{FF}_{\nu_0} A^{FF} U^F. &\text{(34b)} \end{aligned}$$

We emphasize that, in (34), the notation implies that the output on the lhs is not conditioned on the sequence of measurements, but treats that sequence as part of the output information that is subject to chance. We argue in favor of (34a) as follows: The given sequence of measurements, together with the input $\big(U^F\big)$ and output $\big(U^B\big)$ modes, allow us to infer the unique sequence of encounters of the system trajectory with the apparatuses A, M, and C, as in (32). We know, therefore, that in the first passage through A (in (32b)) the transitions entailed by A^{FF}, and not those of A^{BF}, occurred. The minimal assumption is that the components of the new vector $\tilde{X}_0^F$ have the same ratios to one another as they would have in the absence of this knowledge, but that the vector's components are each enhanced by a common factor such that the net output current is equal to that of the input; the partial output current without the above, or any, information as to how the input current was diverted by A, is simply $\tilde{X}_0^F\big(U^F\big) = A^{FF}U^F$ with no multiplying factor. A similar argument applies to the steps (32d), (32f), (32h), and (32j). With respect to trajectory's first passage across a measuring device in (32c), the probability that the device M will register the coarse grain ν_0 is, by the usual rule for conditional probabilities, just the ratio given by $N\big(I^{FF}_{\nu_0}, \tilde{X}_0^F\big)$; the distribution of output over the complete set of fine-grained F-type channels is therefore $Y^F[\nu_0]\big(U^F\big) = I^{FF}_{\nu_0}\tilde{X}_0^F$, with no multiplying factor. A similar argument obtains for (32e), (32g), and (32i). Therefore, (34) is established.

We note that

$$I^{FF} = \sum_{\nu=1}^{\nu^F} I_\nu^{FF}, \tag{35a}$$

$$I^{BB} = \sum_{\nu=1}^{\nu^B} I_\nu^{BB}. \tag{35b}$$

Therefore, we have

$$U_1^B(U^F) \stackrel{\text{def}}{=} \sum_{\nu_0=1}^{\nu^F} \sum_{\nu_{0'}=1}^{\nu^B} \sum_{\nu_1=1}^{\nu^F} \sum_{\nu_{1'}=1}^{\nu^B} U_1^B[\nu_{1'}, \nu_1, \nu_{0'}, \nu_0](U^F) \tag{36a}$$

$$= A^{BB} C^{BF} A^{FB} C^{BF} A^{FF} U^F, \tag{36b}$$

where the last rhs is just the probability distribution in U^B starting with U^F that is predicted with exactly one feedback loop, and with no detailed coarse-grained measurements having been made, i.e., $M^{FF} = I^{FF}$ and $M^{BB} = I^{BB}$. (More precisely, following the interpretation of Lüders [15], we make a measurement of just the unit operator I^{FF} or I^{BB} on passage of the system trajectory through the device M.) This result corresponds to a term in the series expansion of the second summand on the rhs of (16c). If we sum over all possible continuous paths and all possible measurement outcomes in the transition from U^F to U^B, we recover the whole rhs of (16c), and similarly for the other blocks of (16). That is, the overall outcome when no measurements are made, as in (16), can be constructed from an ensemble of results of detailed measurements on the system.

We now consider that, in passing from a U^F channel to a U^B channel, the trajectory crosses M twice in the F direction, with registered values ν_0, ν_1, and twice in the B direction with registered values $\nu_{0'}, \nu_{1'}$, but such that it is not possible to infer from these data in which order the respective pairs of F-type and B-type crossings occurred along the trajectory. In particular, the system can pass through the same coarse grain twice ($\nu_1 = \nu_0$ or $\nu_{1'} = \nu_{0'}$). The total number of possible processes is

$$\text{number} = \left(\nu^F\right)^2 \left(\nu^B\right)^2. \tag{37}$$

We characterize sets of data in the following four ways: (i) $\nu_1 = \nu_0$ and $\nu_{1'} = \nu_{0'}$, (ii) $\nu_1 = \nu_0$ and $\nu_{1'} \neq \nu_{0'}$, (iii) $\nu_1 \neq \nu_0$ and $\nu_{1'} = \nu_{0'}$, and (iv) $\nu_1 \neq \nu_0$ and $\nu_{1'} \neq \nu_{0'}$. The unordered sequences have (i) one, (ii) two, (iii) two, and (iv) four ordered ways to be realized; we must sum over probabilities for the respective distinct ordered processes in order to infer the net probability for an unorderable sequence of measured coarse-grain values to be detected. Given only that the process $U^B \leftarrow U^F$ has occurred with two crossings each of M^{FF} and M^{BB}, the probability that this outcome occurs but no detailed measurements are made on the state of the system at a crossing can still be obtained by summing probabilities over the disjoint subsets of unordered measurements.

4 Discussion

The principle of evaluating conditional probabilities describes the effect of a collapse of a probability distribution. In fact, let $\mathcal{S}$ be an index set with elements $\zeta \in \mathcal{S}$. Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be nonempty subsets of $\mathcal{S}$. Let $p(\zeta)$ be a probability distribution such that $0 \leq p(\zeta) \leq 1$, and

$$\sum_{\zeta \in \mathcal{S}} p(\zeta) \;=\; 1, \tag{38a}$$

$$\sum_{\zeta \in \mathcal{S}_1} p(\zeta) \;=\; P_1, \tag{38b}$$

$$\sum_{\zeta \in \mathcal{S}_2} p(\zeta) \;=\; P_2 \neq 0. \tag{38c}$$

Now suppose that it be given that $\mathcal{S}_2$ is "true", in that we now take as input information the circumstance that $p(\zeta) = 0$ for ζ in the complement of $\mathcal{S}_2$ in $\mathcal{S}$. Then we infer a collapsed probability distribution

$$p(\zeta|\mathcal{S}_2) \;=^{\text{def}}\; \begin{cases} p(\zeta)/P_2, & \text{if } \zeta \in \mathcal{S}_2, \\ 0, & \text{otherwise.} \end{cases} \tag{39}$$

We obtain the usual conditional probability law

$$P(\mathcal{S}_1|\mathcal{S}_2) \;=\; \sum_{\zeta \in \mathcal{S}_1 \cap \mathcal{S}_2} p(\zeta|\mathcal{S}_2) \;=\; P(\mathcal{S}_1 \cap \mathcal{S}_2)/P_2, \tag{40}$$

in an obvious notation. The probabilities $p(\zeta)$ and $p(\zeta|\mathcal{S}_2)$ can also be described as the distributions before and after the limiting case of a non-interventional measurement on the system, respectively.

Let us now study the case that there are no intermediate measurements performed in the system described in Fig. 1, i.e., $M^{FF} = I^{FF}$ and $M^{BB} = I^{BB}$. Let us also take $n^F = 2$ and $n^B = 1$, and label the F-type channels with subscripts 1 and 2. We shall analyze a probabilistic version of a grandfather paradox ([18], *passim*), and show that in this context of classical probability flows, without the destructive interference that can be provided by quantum mechanics, a feedback loop is incapable of *decreasing* the F to F survival probability of a state or channel to a lower value than it would have if the feedback were absent. Let us take

$$A \;=\; \begin{bmatrix} 1 & 0 & \alpha \\ 0 & 1 & \beta \\ 0 & 0 & 1-\alpha-\beta \end{bmatrix}, \tag{41a}$$

$$C \;=\; \begin{bmatrix} 1-\gamma & 0 & 0 \\ 0 & 1 & 0 \\ \gamma & 0 & 1 \end{bmatrix}, \tag{41b}$$

$$\tag{41c}$$

where $0 \leq \alpha, \beta, (1-\alpha-\beta), \gamma \leq 1$ and $\alpha\gamma < 1$. According to (13) we have

$$L^{FF} = \begin{bmatrix} 1/(1-\alpha\gamma) & 0 \\ \beta\gamma/(1-\alpha\gamma) & 1 \end{bmatrix}, \tag{42a}$$

$$L^{BB} = 1/(1-\alpha\gamma), \tag{42b}$$

so that

$$S^{FF} = \begin{bmatrix} (1-\gamma)/(1-\alpha\gamma) & 0 \\ \beta\gamma/(1-\alpha\gamma) & 1 \end{bmatrix}, \tag{43a}$$

$$S^{BF} = \begin{bmatrix} \gamma(1-\alpha-\beta)/(1-\alpha\gamma) & 0 \end{bmatrix}. \tag{43b}$$

Given that the inputs are $(U^F)_1 = 1$, $(U^F)_2 = 0$, and $V^B = 0$, the outputs V^F and U^B are given by the first columns of S^{FF} and S^{BF}. We consider γ to stand for the strength of a signal sent backwards from the apparatus C, and the $(U^F)_1 \to (V^F)_1$ transmission to be adjusted by the feedback parameter α. When $\alpha = 0$ the feedback loop is open. As α increases from zero to one, the output signal

$$(V^F)_1 = (1-\gamma)/(1-\alpha\gamma) \tag{44}$$

always increases.

It is plausible, therefore, to infer that a nonzero feedback loop in classical probability can only increase the survival probability of the classical state of interest, here the first component of the F-type state vector, above its value when the feedback loop is zeroed. We infer that only quantum-mechanical feedback will have the capability of diminishing the survival probability of a physical state within a feedback sysyem of the type of Fig. 1. Two recent online preprints ([19], [20], and references given therein) study time travel from the viewpoint of conventional quantum mechanics. I believe that the proper framework for analysis of quantum-mechanical time-travel phenomena is a quantum analog of the classical system described in the present paper, for which the basic dynamics is described in [2]; the task of analysis of quantum-mechanical feedback and measurement within the formalism of [2] remains to be accomplished.

References

[1] G. E. Hahne. *J. Phys.* **A36**, 7149 (2003). Online at arXiv:quant-ph/0404012.

[2] G. E. Hahne. *J. Phys.* **A35**, 7101 (2002). Online at arXiv:quant-ph/0404103.

[3] W. Pauli. *General Principles of quantum mechanics.* Springer, Berlin, 1980.

[4] S. Karlin and H. M. Taylor. *A first course in stochastic processes.* Academic Press, New York, 1975, 2nd edition.

[5] A. Papoulis. *Probability, random variables, and stochastic processes.* McGraw-Hill, New York, 1965.

[6] E. Parzen. *Stochastic processes.* Holden-Day, San Francisco, 1962.

[7] A. Leon-Garcia. *Probability and random processes for electrical engineering.* Addison Wesley, Reading, MA, 1994.

[8] R. v. Mises. *Mathematical theory of probability and statistics.* Edited and complemented by H. Geiringer. Academic Press, New York, 1964.

[9] W. Feller. *An introduction to probability theory and its applications*, Volume 1. Wiley, New York, 1968, 3rd edition.

[10] D. T. Gillespie. *Markov processes.* Academic Press, San Diego, CA, 1992.

[11] J. R. Norris. *Markov chains.* Cambridge U. Press, Cambridge, UK, 1997.

[12] P. E. Pfeiffer. *Concepts of probability theory.* Dover, New York, 1978.

[13] A. T. Bharucha-Reid. *Elements of the theory of Markov processes and their applications.* McGraw-Hill, New York, 1960.

[14] H. Minc. *Nonnegative matrices.* Wiley, New York, 1988.

[15] G. Lüders. *Ann. Physik (Leipzig)* **8**, 322 (1951).

[16] W. H. Furry. in *Lectures in Theoretical Physics*, Vol. VIIIA, edited by W. E. Brittin, Gordon and Breach, New York, 1966.

[17] N. Gisin. *Phys. Rev. Lett.* **52**, 1657 (1984).

[18] P. J. Nahin, *Time Machines.* Springer, New York, 1999, 2nd edition.

[19] D. T. Pegg. arXiv:quant-ph/0509141 v1, 17 June 2005

[20] D. M. Greenberger and K. Svozil. arXiv:quant-ph/0506027 v2, 21 June 2005