qmc082.tex. Version of 30 September 2010.
Lecture Notes on Quantum Mechanics No. 8
R. B. Griffiths

Stochastic Processes

References:

CQT = R. B. Griffiths, *Consistent Quantum Theory* (Cambridge, 2002)
DeGroot and Schervish, *Probability and Statistics*, 3d ed (Addison-Wesley, 2002)
S. M. Ross, *Introduction to Probability Models* (Academic Press, numerous editions)

Contents

---

1 Introduction

⊙ A *stochastic process* or *random process* is a sequence of successive events in time, described in a probabilistic fashion.

• After the great success of Newton's mechanics in describing planetary motion, the belief among physicists was that time development in nature is fundamentally *deterministic*: if the state of the entire universe at some particular time is given with sufficient precision, then its future and past are determined, and can in principle be calculated.

∘ The most famous statement of this idea is by Laplace (1814):

– We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes.

• Various arguments show that such a calculation is in practice impossible. Quantum mechanics means it is not even possible in principle. We *have* to use probabilities.

⊙ References. Classical stochastic process are discussed in CQT Secs. 8.2, 9.1, and 9.2, and DeGroot and Schervish Sec. 2.4. For a very extensive treatment with a large number of examples see Ross. The quantum sample space is discussed in CQT Ch. 8, and the assignment of probabilities is the subject of the chapters that immediately follow.

2 Classical Stochastic Processes

2.1 Formalism

⊙ For simplicity consider a finite number of discrete times $t_0 < t_1 < \cdots < t_f$, and at each time the same sample space S of mutually exclusive possibilities, at most countably infinite. We will usually choose it to be finite and independent of the time.

∘ Example. A coin is tossed several times in succession: $S = \{H, T\}$, heads or tails.

∘ A die is rolled several times in succession: $S = \{1, 2, 3, 4, 5, 6\}$.

⊙ To describe the stochastic process itself we need a sample space $\breve{S}$ of sequences $\mathbf{s}$ of the form

$$\mathbf{s} = (s_0, s_1, s_2, \ldots s_f), \tag{1}$$

where each s_j is some element of S. We shall refer to a particular $\mathbf{s}$ as a *history*, and to $\breve{S}$ as the *history sample space*, or sample space of histories.

∘ Notice that the order of events in time matters: HHT, that is $s_0 = H, s_2 = H, s_2 = T$, is not the same thing as HTH for a coin tossed three times in a row. The different time orderings represent mutually exclusive possibilities.

⊙ Again for simplicity assume that the event algebra $\breve{\mathcal{E}}$ consists of all subsets (including $\emptyset$ and $\breve{S}$ itself) of $\breve{S}$: any collection of histories from the sample space belongs to the event algebra.

⊙ Finally, probabilities must be assigned to all items in the event algebra. We do this by assigning a nonnegative number $\Pr(\mathbf{s})$ to each history in $\breve{S}$ in such a way that they sum 1. An element of $\mathcal{E}$ is assigned a probability equal to the sum of the probabilities of the individual histories that it contains.

∘ There is no general rule for assigning such probabilities; anything is allowed, provided the numbers are nonnegative and sum to 1. Typically one is trying to produce a *stochastic model* of some series of events, so one makes an assignment for which the mathematics is not too difficult, perhaps containing some parameters which could be chosen to fit experimental results.

⊙ If the s_j that make up a history $\mathbf{s}$ are themselves integers or real numbers, as is often the case, they are *random variables*, since, e.g., s_2 assigns to $\mathbf{s}$ the third element in $(s_0, s_1, s_2 \ldots)$. Consequently there is a one-to-one correspondence between elements of the sample space and the collections of values taken by these $f+1$ random variables, and one can regard the random variables taken together as constituting the sample space.

• Given the joint distribution $\Pr(\mathbf{s})$ one can calculate the marginal or *single time* distribution $\Pr(s_j)$ for any particular j:

$$\Pr(s_j) = \sum_{s_0}\sum_{s_1}\cdots\sum_{s_{j-1}}\sum_{s_{j+1}}\cdots\sum_{s_f}\Pr(s_0, s_1, \ldots s_{j-1}, s_j, s_{j+1}, \ldots s_f). \tag{2}$$

∘ However, from a knowledge of $\Pr(s_j)$ for every j from 0 to f it is, in general, impossible to construct the joint distribution $\Pr(\mathbf{s})$. The latter contains more information than found in the collection of all the single time probabilities. It allows one to compute the *correlations* between the values of the different s_j representing states of the system at different times.

• As the reader will have noticed, there is nothing really special about stochastic processes from the perspective of general probability theory. Sample space, event algebra, probabilities: all follow the usual rules. The only special thing about a stochastic process is the way in which one views it and the types of questions one asks.

2.2 Independent processes

• Consider the specific case of a coin tossed N times in a row, and let $s_j = H$ or T be the outcome on toss j. The sample space is of size 2^N. A simple way of assigning probabilities is to suppose that each time the coin is tossed the probability is p that the result is H and $q = 1-p$ that it is T, with p and q independent of the toss, and that the tosses are *statistically independent*, so that the probability associated with some sequence or history is just the product of the probabilities of the individual outcomes. Thus $\Pr(HHT)$ is p^2q, because H appears twice and T appeared once.

⊙ More generally an *independent* stochastic process has a joint probability distribution given by

$$\Pr(s_0, s_1, \dots s_f) = \Pr(s_0)\Pr(s_1)\dots\Pr(s_f), \tag{3}$$

so it is an exception to the rule that the joint distribution contains more information than the collection of one-time distributions.

• It may be that the one-time distributions are all identical functions of their arguments:

$$\Pr(s_j) = g(s_j), \tag{4}$$

where the function $g(s)$ is the same, independent of j. In this case one says that the random variables s_0, s_1, etc. are *identically distributed*. But one can also have a situation where $\Pr(s_j) = g_j(s_j)$, i.e., there are different functions giving probabilities of the different random variables.

• The case of "independent identically-distributed random variables" comes up so often in books on probability that it is often abbreviated to i.i.d. Note that the coin tossing example we started with is of this form if one replaces H with $+1$ and T with -1 so as to turn the outcome into a random variable. A die rolled several times in succession (assuming it is not carefully rolled in exactly the same way each time) is described by an i.i.d.

◦ Despite the simplicity of i.i.d.'s, figuring out their statistical properties can require some effort.

⋆ Exercise. Suppose an honest die is rolled N times in a row and a record is made of the number of spots. What is the probability that the sum will be smaller than $2N$, assuming that N is large? If that is too easy, consider the case where all the p_s are equal except $p_6 = 1/5$.

• In a sense what is distinctive about i.i.d.'s, and to some extent more general independent processes, is that the *time ordering* plays no role. The fact that the die has come up with $s = 1$ on three successive rolls does *not* increase the probability that it will come up $s = 6$ the next time.

⋆ Exercise. In fact if the game has just started and you have observed $s = 1$ on the first three rolls you might want to conclude that the probability of $s = 6$ on the fourth roll is *less* than 1/6. Discuss.

2.3 Markov processes

• A Markov process is a simple type of stochastic process in which the time order in a sequence of events plays a significant role: what has just happened can influence the probability of what happens next.

⊙ An example is a random walk in one dimension. Let the time $t = j$ be a nonnegative integer, and let the location of the walker at time j be an integer s_j, positive or negative. At each time the walker can hop to the left, $s_{j+1} = s_j - 1$, with probability p, or to the right, $s_{j+1} = s_j + 1$, with probability r, or stay put, $s_{j+1} = s_j$, with probability q; $p + q + r = 1$. Obviously, knowing something about the walker's location at time $t = j$ will provide useful information about its location at $t = j + 1$; the s_j are not statistically independent.

• To get a finite sample space one can suppose that

$$-N \le s_j \le N, \tag{5}$$

where N is some positive integer, and use periodic boundary conditions: the random walker hops from N to $-N$ instead of hopping from N to $N+1$, and from $-N$ to N instead of to $-N-1$. The sample space $\breve{S}$ is then the set of all sequences (1) with the s_j lying in the range (5) for every j, a collection of $(2N+1)^{f+1}$ histories.

∘ Some sequences will be assigned zero probability because we only allow hops such that $|s_{j+1} - s_j| \leq 1$, and these could be eliminated from the sample space without changing anything, but by the same token they can simply be left in the same space and assigned zero probability.

∘ Assigning probabilities to random walks is a particular case of the general scheme used for Markov processes described next.

⊙ The joint probability distribution for a Markov process, of which the random walk is a particular case, is constructed in the following way. Write

$$\Pr(s_0, s_1) = \Pr(s_1 \,|\, s_0) \Pr(s_0), \quad \Pr(s_0, s_1, s_2) = \Pr(s_0, s_1) \Pr(s_2 \,|\, s_0, s_1), \tag{6}$$

and so forth. Thus if we know the marginal probability distribution $\Pr(s_0)$ and the conditional probabilities $\Pr(s_1 \,|\, s_0)$, $\Pr(s_2 \,|\, s_0, s_1)$, $\Pr(s_3 \,|\, s_0, s_1, s_2)$, and so forth, we can multiply them together to calculate $\Pr(s_0, s_1, \ldots s_f)$. Thus far no assumption has been made; this could be done for any stochastic process.

⊙ The *Markov* assumption is that for every j it is the case that

$$\Pr(s_{j+1} \,|\, s_0, s_1, \ldots s_j) = \Pr(s_{j+1} \,|\, s_j) = M(s_{j+1}, s_j), \tag{7}$$

i.e., the conditional probability on the left side depends on s_{j+1} and s_j, as one would expect, but is then *independent* of $s_0, s_1, \ldots$ up to s_{j-1}.

• One refers to $M(s_{j+1}, s_j)$ as the (Markov) *transition matrix*. Since a conditional probability is a "genuine" probability distribution for its left argument (the one preceding $|$), it follows that $M(s, s') \geq 0$, and for every s'

$$\sum_s M(s, s') = 1. \tag{8}$$

Think of $M(s, s')$ as a matrix $M_{ss'}$. Each column consists of nonnegative numbers that sum to 1.

∘ A matrix $M_{ss'}$ or $M(s, s')$ with nonnegative elements that satisfies (8) is called a *stochastic matrix*.

• We have assumed that $M(s_{j+1}, s_j)$ depends only on the values s_{j+1} and s_j of the two arguments, which is to say we are dealing with a *stationary* Markov process. One could also make the transition matrix depend explicitly on the time step: $M(j; s_{j+1}, s_j)$. The result would be a *nonstationary* Markov process.

∘ The random walker is described by a stationary random process, but the walker itself is not stationary: it hops from integer to integer, although sometimes it stays put. Thus "stationary" refers to the fact that the hopping probabilities which enter the transition matrix do not depend upon the time.

• One can think of the Markov property (7) as a "lack of memory" of what happened previously. E.g., the walker's next position depends on its current position, but apart from that its earlier positions have no influence on what happens next.

⊙ Putting together (6) and (7) one arrives at the formula

$$\Pr(s_0, s_1, \ldots s_f) = M(s_f, s_{f-1}) M(s_{f-1}, s_{f-2}) \cdots M(s_1, s_0) \Pr(s_0) \tag{9}$$

for the joint probability distribution.

• From this one can calculate the marginal distribution for any s_j:

$$\Pr(s_j) = \sum_{s_0,s_1,\dots s_{j-1}} M(s_j, s_{j-1}) M(s_{j-1}, s_{j-2}) \cdots \Pr(s_0) = M^j \Pr(s_0) \tag{10}$$

where the right side uses an obvious matrix notation, with $\Pr(s_0)$, and thus $\Pr(s_j)$, a column vector.

⋆ Exercise. How should (9) and (10) be modified for a nonstationary Markov process?

⋆ Exercise. Show that if M is a stochastic matrix and $\Pr(s_0)$ an arbitrary probability distribution, the joint distribution defined by the right hand side of (9) has the property that for every j it is the case that $\Pr(s_{j+1} \mid s_0, s_1, \dots s_j) = \Pr(s_{j+1} \mid s_j)$; i.e., the first equality in (7) is satisfied.

• Note. Probabilists often use the transpose of our M as the transition matrix, and modify the right side of (10) accordingly, with $\Pr(s_j)$ a row rather than a column vector, and the rows rather than the columns of the transition matrix summing to 1. The row sum rather than the column sum defines what one means by a stochastic matrix.

• A matrix $M(s, s')$ for which each row as well as each column sums to 1 is called *doubly stochastic*.

⊙ Back to the random walk. For this case the transition matrix is:

$$M(s, s') = \begin{cases} p & \text{if } s = s' - 1 \\ q & \text{if } s = s' \\ r & \text{if } s = s' + 1. \end{cases} \tag{11}$$

with, if we use the restriction (5) to keep the sample space finite, an obvious modification to take care of the periodic boundary condition.

⋆ Exercise. Make the obvious modification assuming $N = 1$, and write down M as a 3×3 matrix with rows and columns in the order -1, 0, 1 from top to bottom and from left to right. Check that the columns add to 1.

⋆ Exercise. The matrix you have just constructed also has the property that each row sums to 1, so it is doubly stochastic. However, you can make a small modification in it by adding ϵ to one matrix element and subtracting it from another so that the matrix remains stochastic (columns sum to 1) but is no longer doubly stochastic. Do this, and then explain the significance of the modified entries in terms of hopping probabilities.

• Rather than using periodic boundary conditions for the random walk it is possible to employ "reflecting" boundary conditions: when $s_j = N$, at the next step the walker either stays put or hops to $N - 1$; when $s_j = -N$ it either stays put or hops to $-(N - 1)$, with suitable probabilities (for which there is not a unique choice).

⋆ Exercise. For the $N = 1$ case modify the stochastic matrix you obtained previously using periodic boundary conditions, so that it becomes appropriate for reflecting boundary conditions. Explain briefly the significance of the changes. Is the transition matrix still doubly stochastic?

2.4 Statistical inference

• Statistical inference in the case of a stochastic process can be carried out using conditional probabilities in the usual way.

⊙ Let us consider a specific example involving a random walk, where N is large so we won't have to worry about periodic boundary conditions. Suppose that at $t = 0$ the walker is at $s = 0$ and at $t = 3$ the walker is at $s = 1$. What is its location at $t = 2$?

• We use the first datum, $s = 0$ at $t = 0$, to set up the probability distribution $\Pr(\mathbf{s})$, see (9). There will be $3^3 = 27$ histories of the form $(s_0 = 0, s_1, s_2, s_3)$ with nonzero probability, assuming that p, q, r in (11) are all positive, because at each step the walker has 3 choices. The probability of a given history will then be $p^i q^j r^k$ if the history involves i hops to the left, j cases of staying put, and k hops to the right, with $i + j + k = 3$.

∘ Use the second datum, $s_3 = 1$, to select from the 27 histories just mentioned the ones we are interested in, where $s_3 = +1$. In the following table the probability is listed just below each history.

$$\begin{matrix} (0,-1,0,1) & (0,0,0,1) & (0,0,1,1) & (0,1,0,1) & (0,1,1,1) & (0,1,2,1) \\ pr^2 & q^2 r & q^2 r & pr^2 & q^2 r & pr^2 \end{matrix} \tag{12}$$

• These six probabilities are *weights* in the sense that they are positive numbers giving the *relative* probabilities of the different possibilities, even though they do not add up to 1. To convert them into the (conditional) probabilities we are interested it, divide each by the sum, $3r(q^2 + pr)$. (It is possible to rewrite this in terms of only two hopping probabilities, say p and r, using the fact that $p + q + r = 1$, but that does not necessarily lead to a simpler expression.)

• So where was the walker at $t = 2$? There is no definite answer to this question, but we can assign probabilities to the different values of s_2 using the weights in (12)—note that s_2 is the next-to-the-last entry in (s_0, s_1, s_2, s_3). The three possibilities $s_2 = 0$ or 1 or 2 are then assigned the following conditional probabilities:

$$\begin{aligned} &\Pr(s_2 = 0 \,|\, s_3 = 1) = \frac{q^2 + 2pr}{3(q^2 + pr)}, \quad \Pr(s_2 = 1 \,|\, s_3 = 1) = \frac{2q^2}{3(q^2 + pr)}, \\ &\Pr(s_2 = 2 \,|\, s_3 = 1) = \frac{pr}{3(q^2 + pr)}. \end{aligned} \tag{13}$$

⋆ Exercise. One check on (13): set $q = 0$ and work out the probabilities in this case, where there are a fewer walks (with nonzero probabilities) to consider.

∘ One could very well include $s_0 = 0$ along with $s_3 = 1$ as another condition to the right of the bar | in (13), since the two data actually enter the problem in a symmetrical way. Conditions are often omitted from conditional probabilities when they are evident from the context, so one could also omit $s_3 = 1$.

⋆ Exercise. What is $\Pr(s_2)$ given *only* the condition that $s_0 = 0$, no constraint at time $t = 3$?

⋆ Exercise. Suppose you were only given that $s_3 = 1$ and no information about where the walker was at $t = 0$. Could you still find $\Pr(s_2)$? Or at least make a plausible guess?

3 Quantum Stochastic Processes

3.1 Introduction

• A quantum stochastic process consists of sequences of quantum properties (subspaces of the Hilbert space) at successive times to which probabilities can be assigned.

∘ Classical stochastic processes provide a good starting point for thinking about the quantum case, but, as one would expect, the latter involves some additional ideas.

• There are two problems. The first is to set up a quantum sample space of histories and the corresponding event algebra. This topic is addressed below. The second is to assign probabilities in a manner related to Schrödinger's equation, which is more complicated, and is taken up in a later set of notes.

3.2 Quantum histories

⨀ Consider a finite sequence of times $t_0 < t_1 < t_2 < \cdots t_f$. A *quantum history* assigns to the quantum system at each of these times some quantum property: a subspace of the Hilbert space or, equivalently, its projector. We can write such a history schematically in the form

$$F_0 \odot F_1 \odot F_2 \odot \cdots F_f, \tag{14}$$

where each F_m is the projector (property) at time t_m.

∘ Read (14) as "F_0 is true at t_0, F_1 is true at t_1, $\dots F_f$ is true at t_f." The significance of $\odot$ will be explained shortly.

∘ Example for a spin-half particle when $f = 2$, a history with events at $f + 1 = 3$ different times.

$$[z^+] \odot [z^-] \odot [x^+]. \tag{15}$$

∘ Note that $F_m = I$ is an acceptable property of the quantum system, but since it is always true, inserting it at time t_m tells us nothing at all: this time might just as well be omitted from the history. Similarly, one could insert an additional time, say time $t_{2.5}$ someplace between t_2 and t_3, and let $F_{2.5} = I$. This would make no difference.

• WARNING! We are *not* assuming that the F_m projectors are related by Schrödinger's equation or the time development operators; do *not* assume that $F_{m+1} = T(t_{m+1}, t_m)F_m T(t_m, t_{m+1})$. For the present just assume that each F_m can be chosen independently of all the others.

• The next idea, which goes back to C. Isham, can be motivated in the following way. The classical sample space corresponding to flipping a coin three times in succession is formally the same as that obtained by flipping each of three (distinct) coins just once. Thus if the coins are labeled 1, 2, and 3 (penny, nickel, dime) there are eight mutually exclusive events in the sample space, things such as H_1, T_2, T_3, just as if a single coin, say a quarter, were flipped three times in a row. We already know how to describe three distinct quantum objects at a single time: use the Hilbert space corresponding to the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$. Why not employ the same idea for histories?

⨀ Indeed it works, and we shall think of the history (13) as a projector on the *history* Hilbert space

$$\breve{\mathcal{H}} = \mathcal{H}_0 \odot \mathcal{H}_1 \odot \cdots \mathcal{H}_f, \tag{16}$$

a tensor product of $f + 1$ copies of the Hilbert space $\mathcal{H}$ used to represent properties of the system at a single time.

• The symbol $\odot$ means the same thing as the usual tensor product symbol $\otimes$; however, it is convenient to have a distinctive notation when dealing with events at successive times.

∘ The tensor product of projectors, as in (14), is a projector. For a system with parts labeled 0, 1, 2, etc., the meaning would be: system 0 has property F_0, system 1 has property F_1, and so forth. This corresponds to our interpretation of (14): the system of interest has property F_0 at time t_0, F_1 at time t_1, and so forth.

• On any tensor product there are, of course, entangled states and projectors onto entangled states. However, our discussion will be limited to the simplest situation in which histories of the quantum system are either represented by products of projectors, as in (14), or sums of projectors of this form.

3.3 Sample space and event algebra

⊙ A quantum sample space is always a decomposition of the identity for some suitable Hilbert space, a collection of projectors that sum to the identity. When interested in the properties of a system with Hilbert space $\mathcal{H}$ at a single time we use a sample space consisting of projectors that form a decomposition of the identity I operator on $\mathcal{H}$. Thus a sample space of histories should correspond to a decomposition of the identity $\breve{I}$ of the history Hilbert space $\breve{\mathcal{H}}$ introduced above in (16).

• Be careful to distinguish the history identity

$$\breve{I} = I_1 \odot I_2 \odot \cdots I_f, \tag{17}$$

an operator on the Hilbert space $\breve{\mathcal{H}}$ of histories, from the identity I on the Hilbert space $\mathcal{H}$ of the system at a single time. In (17) we have used subscripts for I corresponding to those on the right side of (16).

⊙ There are many ways to construct a history sample space. Here is one which is relatively straightforward. At time t_m choose a decomposition $\{P_m^{\alpha_m}\}$ of the single-time identity I_m on $\mathcal{H}_m$:

$$I_m = \sum_{\alpha_m} P_m^{\alpha_m}. \tag{18}$$

• Notation. The subscript m is used to label the time t_m, the identity I_m on the Hilbert space $\mathcal{H}_m$ at this time (which will usually be the same as $\mathcal{H}$, the ordinary Hilbert space used to describe the system we are interested in), and the projectors which form a decomposition of the identity at this time. The different projectors that constitute this decomposition are then distinguished from each other by a *superscript* label α_m; note that this is a label and not a power. Because the square of a projector is equal to the projector there is no reason to raise a projector to some power, so it is (relatively) safe to use the superscript position for a label. You can think of α_m as taking on integer values $1, 2, \ldots$; or $0, 1, \ldots$ if you prefer.

∘ The decomposition might be the same for every m, but more flexibility is needed if we want to include something like (15) in our discussion.

⋆ Exercise. Explain why.

⊙ Given the decompositions in (18) we can set up a *history sample space* consisting of projectors on $\breve{\mathcal{H}}$ of the form

$$Y^\alpha = P_0^{\alpha_0} \odot P_1^{\alpha_1} \odot \cdots P_f^{\alpha_f}, \tag{19}$$

where

$$\alpha = (\alpha_0, \alpha_1, \ldots \alpha_f) \tag{20}$$

is an $(f+1)$-component label.

⋆ Exercise. Show that

$$Y^\alpha Y^\beta = \delta_{\alpha\beta} Y^\alpha \tag{21}$$

where $\alpha = \beta$ means $\alpha_m = \beta_m$ for every m, and that

$$\sum_\alpha Y^\alpha = \breve{I}. \tag{22}$$

⊙ We shall say that the sample space constructed in this way is a *product sample space*, because we started off with separate samples spaces (18), one for each time, and formed the history sample space by taking the product of these spaces: the former is made up of all possible tensor products (in the sense of $\odot$) of projectors from the single-time sample spaces.

• Although product sample spaces as just defined are simple and often suffice, they do have their limitations. Thus there are cases in which the history sample space, even though it contains only projectors which are tensor products (in the sense of $\odot$) of projectors at different times, does *not* make use of a single decomposition of the identity at each time. A simple example for a spin-half particle and $f = 1$ (two times) is the set of projectors

$$[z^+] \odot [z^+], \quad [z^+] \odot [z^-], \quad [z^-] \odot [x^+], \quad [z^-] \odot [x^-]. \tag{23}$$

⋆ Exercise. Check that the projectors in (23) form a decomposition of the history identity $\breve{I} = I_0 \odot I_1$ by showing that they are mutually orthogonal and that they sum to $\breve{I}$.

• Another example that does not quite fit the pattern in (19) is that in which one supposes that the quantum system starts off in a particular state $|\psi_0\rangle$ at t_0 and at later times t_1, t_2, etc. has events of the sort drawn from the corresponding decompositions in (18). We could, in fact, use a decomposition

$$I_0 = [\psi_0] + (I_0 - [\psi 0]) = P_0^1 + P_0^2 \tag{24}$$

at t_0, and construct the sample space as in (19). But since we are not really interested in histories where the system does not start in the state $|\psi_0\rangle$ or $[\psi_0]$, it is convenient to sum all the Y^α with $\alpha_0 = 2$, corresponding to histories that start off with $(I - [\psi 0])$, into a single projector

$$\bar{Y} = P_0^2 \odot I_1 \odot I_2 \cdots I_f, \tag{25}$$

whose meaning is that the system was *not* in the state $[\psi_0]$ at t_0, and nothing else is being said about what happened to it later. Then we can forget about $\bar{Y}$, since it will be assigned zero probability.

◦ The sample space then consists of $\bar{Y}$ along with the Y^α for which $\alpha_1 = 1$.

⋆ Exercise. Use the scheme just discussed to make up a sample space which contains the history (15), assuming that $[\psi_0] = [z^+]$. List all of the projectors in the sample space.

⊙ There is no reason in principle why history sample spaces should not contain projectors onto states that are "entangled" at different times, analogous to the entangled states at a single time which quantum theory allows for the tensor product Hilbert space (in the sense of $\otimes$) of a system containing several parts. Whether these more general possibilities are useful for describing the time development of quantum systems in certain circumstances has not yet been studied. In these notes we restrict ourselves to cases where the history sample space projectors are products (in the sense of $\odot$) of projectors at different times.

⊙ Once a sample space of histories has been established, the *event algebra* consists of sums of projectors from this sample space, just as in the case of properties of a system at one time. The connection between sample space and event algebra, at least formally, is just the same as for quantum properties at a single time.

⊙ Two history sample spaces, and the corresponding event algebras are *compatible* if the projectors in one commute with projectors in the other. This means one can construct a common refinement whose event algebra includes both of the original event algebras. Otherwise they are incompatible and cannot be combined (single framework rule).

• This compatibility rule is exactly the same as for alternative sample spaces (decompositions of the identity) of a quantum system at one time.

$\odot$ An additional rule comes into play when assigning probabilities to histories using extensions of the Born rule to three or more times; in this case one has to apply a stronger compatibility condition (known as *consistency*); more about that later,