Stochastic Quantum Dynamics I. Born Rule

Robert B. Griffiths
Version of 25 January 2010

Contents

References:
CQT = *Consistent Quantum Theory* by Griffiths (Cambridge, 2002), Chs. 8, 9, 10, 11.

1 Introduction

★ Quantum dynamics, i.e., the time development of quantum systems, is fundamentally *stochastic* or probabilistic.

• It is unfortunate that this basic fact is not clearly stated in textbooks, including QCQI, which present to the student a strange mishmash of unitary dynamics followed by probabilities introduced by means of *measurements*.

• While the measurement talk is fundamentally sound when it is properly interpreted it can be quite confusing.

◦ Measurements: topic for separate set of notes

• What we will do is to introduce probabilities by means of some simple examples which show what one must be careful about when dealing with quantum systems, and the fundamental principles which allow the calculation of certain key probabilities related to quantum dynamics.

2 Born Rule

2.1 Statement of the Born Rule

★ The Born rule can be stated in the following way. Let t_0 and t_1 be two times, and $T(t_1, t_0)$ the unitary time development operator for the quantum system of interest.

∘ This system must be isolated during the time period under discussion, or at least it must not decohere, as otherwise the unitary $T(t_1, t_0)$ cannot be defined.

★ Assume a normalized state $|\psi_0\rangle$ at t_0 and an orthonormal basis $\{|\phi_1^k\rangle\}, k = 0, 1, \ldots$ at t_1. Then

$$\Pr(\phi_1^k) = \Pr(\phi_1^k \,|\, \psi_0) = |\langle\phi_1^k|T(t_1, t_0)|\psi_0\rangle|^2 = \langle\psi_0|T(t_0, t_1)|\phi_1^k\rangle\langle\phi_1^k|T(t_1, t_0)|\psi_0\rangle \tag{1}$$

is the probability of $|\phi_1^k\rangle$ at t_1 given the initial state $|\psi_0\rangle$ at t_0. In the argument of Pr() and sometimes in other places we omit the ket symbol; $\Pr(\phi_1^k)$ means the probability that the system is in the state $|\phi_1^k\rangle$, or in the ray (one-dimensional subspace) generated by $|\phi_1^k\rangle$, or has the property $|\phi_1^k\rangle$ corresponding to the projector $|\phi_1^k\rangle\langle\phi_1^k| = [\phi_1^k]$. The condition ψ_0 to the right of the vertical bar $|$ is often omitted when it is evident from the context. (For more on conditional probabilities, see separate notes.)

∘ The final equality in (1) is a consequence of the general rule $\langle\chi|A|\omega\rangle^* = \langle\omega|A^\dagger|\chi\rangle$, together with the fact that $T(t_1, t_0)^\dagger = T(t_0, t_1)$.

∘ Mnemonic. In (1) subscripts indicate the time. In the matrix elements the time t_j as an argument of T always lies next to the ket/bra corresponding to this time.

★ It should be emphasized that the Born rule, (1), is on exactly the same level as the time-dependent Schrödinger equation in terms of fundamental quantum principles. That is, it is a postulate or an axiom that does not emerge from anything else.

∘ But there are various ways of checking that it is reasonable, consistent, etc. One of these is:

☐ Exercise. Show that $\sum_k \Pr(\phi_1^k) = 1$. [Hint: $\sum_k |\phi_1^k\rangle\langle\phi_1^k| = ?$]

∘ As a matter of history, Born published his rule about six months after Schrödinger published his equation. His original statement was not in the form (1), but the basic idea was there, and he deserves credit for it.

★ Examples

• Example 1. One qubit, trivial dynamics $T(t_1, t_0) = I$, $|\psi_0\rangle = |0\rangle$, $|\phi_1^k\rangle = |k\rangle$.

$$\Pr(0) = |\langle 0|I|0\rangle|^2 = 1, \quad \Pr(1) = |\langle 1|I|1\rangle|^2 = 0. \tag{2}$$

• Example 2. Same $|\psi_0\rangle$ and $T(t_1, t_0) = I$, but at t_1 use a basis

$$|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}, \quad |-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}. \tag{3}$$

Then

$$\Pr(+) = |\langle +|0\rangle|^2 = 1/2 = \Pr(-). \tag{4}$$

☐ Exercise. What happens if instead of $|\psi_0\rangle = |0\rangle$ we assume $|\psi_0\rangle = |-\rangle$? Assume $T(t_1, t_0) = I$, and work out the probabilities for $|\phi_1^k\rangle$ in both of the preceding examples.

• Example 3. Same $|\psi_0\rangle$, but instead of the trivial I assume that

$$T(t_1, t_0) = H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{5}$$

i.e., the one-qubit Hadamard gate.

∘ One finds in place of (2) the probabilities $\Pr(0) = \Pr(1) = 1/2$; for the basis in (3) the results are $\Pr(+) = 1$, $\Pr(-) = 0$.

□ Exercise. Check these.

□ Exercise. One can think of H as representing a particular rotation of the Bloch sphere. Does this help in understanding the resulting probabilities?

2.2 Incompatible sample spaces

★ Conceptual pitfall. "Because both (2) and (4) are correct results, we can conclude that at t_1 the qubit certainly has the property $|0\rangle$ and also it either has the property $|+\rangle$ or the property $|-\rangle$, with equal probabilities for these last two possibilities." But this is *not* so, because it is quantum nonsense: both $|0\rangle$ and $|+\rangle$ correspond to rays in the quantum Hilbert space, but there is no ray that corresponds to $|0\rangle$ AND $|+\rangle$, or to $|0\rangle$ AND $|-\rangle$.

★ To avoid this trap go back to the fundamentals of (ordinary or "classical") probability theory (discussed in a different set of notes): $(\mathcal{S}, \mathcal{E}, \mathcal{P})$. Sample space $\mathcal{S}$, event algebra $\mathcal{E}$ must be introduced before we can define a probability (measure) $\mathcal{P}$.

• The sample space is a collection of *mutually exclusive* possibilities, one and only one of which actually occurs or is true in a given situation. Examples: H or T for a coin toss; $s = 1$, 2, 3, 4, 5, 6 for a die.

★ A quantum sample space *always* consists of a decomposition of the Hilbert space identity I, a collection $\{P_k\}$ of nonzero projectors such that

$$I = \sum_k P_k, \quad P_k = P_k^\dagger, \quad P_j P_k = \delta_{jk} P_k \tag{6}$$

• $P_k P_j = 0$ for $j \neq k$ tells us that the events or properties represented by the P_k are mutually exclusive: if one occurs a different one does not occur.

• The fact that $\sum_k P_k = I$ tells us that one of these events or properties must be true, since I is the quantum property that is always true.

∘ Sometimes (6) is called a *projective decomposition* to distinguish it from a more general way of writing I as a sum of positive operators that need not be projectors. Our use will be: "decomposition" means a projective decomposition, and the more general sum will be called a "POVM" (positive operator valued measure).

★ We will assume the event algebra consists of every projector on the Hilbert space that is the sum of one or more of the projectors in the collection $\{P_k\}$, along with the 0 projector, corresponding to the empty set in classical probability.

□ Exercise. Show that if there are n projectors in the decomposition (6) there are 2^n projectors in the event algebra.

★ For the Born rule, (1), we choose

$$P_k = [\phi_1^k] = |\phi_1^k\rangle\langle\phi_1^k|, \tag{7}$$

which satisfies (6) because the $\{|\phi_1^k\rangle$ form an orthonormal basis.

• So in example 1 we have

$$P_1 = [0] = |0\rangle\langle 0|, \quad P_2 = [1] = |1\rangle\langle 1|, \tag{8}$$

whereas in example 2

$$Q_1 = [+] = |+\rangle\langle +|, \quad Q_2 = [-] = |-\rangle\langle -|, \tag{9}$$

where we have used Q rather than P to label the operators in the second decomposition.

★ It is easily checked that for every choice of j and k, $P_jQ_k \neq Q_kP_j$ for the operators defined in (8) and (9). This is an example of *quantum incompatibility*: noncommuting operators are a sure sign that we are in the quantum domain where classical physics will no longer work.

★ Given two decompositions $\{P_j\}$ and $\{Q_k\}$ of the identity I we shall say they are *compatible* if $P_jQ_k = Q_kP_j$ for every choice of j and k, and *incompatible* if one of the P projectors fails to commute with one of the Q projectors.

□ Exercise. Show that if $P_jQ_k = Q_kP_j$ then the product is itself a projector (possibly the 0 projector), whereas if $P_jQ_k \neq Q_kP_j$ neither product is a projector.

□ Exercise. Show that if $P_jQ_k \neq Q_kP_j$ for some choice of j and k, there is another distinct pair j', k' such that $P_{j'}Q_{k'} \neq Q_{k'}P_{j'}$. [Hint. If P_j is a projector so is $I - P_j$.]

• Example of compatible decompositions. Consider a qutrit; the standard orthonormal basis consists of the three kets $|0\rangle$, $|1\rangle$, $|2\rangle$. Let $[j] = |j\rangle\langle j|$ be the corresponding projector. Then define

$$P_1 = [0], \quad P_2 = [1] + [2], \quad Q_1 = [0] + [1], \quad Q_2 = [2]. \tag{10}$$

□ Exercise. Show that $\{P_1, P_2\}$ and $\{Q_1, Q_2\}$ are both decompositions of the qutrit identity I, and that they are compatible.

★ If $\{P_j\}$ and $\{Q_k\}$ are compatible decompositions, the collection $\{P_jQ_k\}$ after throwing away the terms that are zero is itself a decomposition of the identity called the *common refinement*.

□ Exercise. Check that they add up to I.

□ Exercise. What is the common refinement in the case of the example in (10) Why do you suppose it is called a *refinement*?

★ Extremely important rule of quantum probabilities and quantum reasoning. *Incompatible sample spaces cannot be combined!*

○ This is an instance of what is known as the *single framework rule*.

• Return to the conceptual pitfall discussed earlier. It takes correct results, namely (2) and (4), associated with incompatible sample spaces, namely (8) and (9), and tries to combine them, with a nonsensical result.

★ By carrying out enough violations of the single framework rule it is possible to construct contradictions, propositions that are (apparently) both true and false. Most quantum paradoxes are constructed in this manner.

○ Further discussion: see CQT Sec. 4.6, and for the paradoxes see Ch. 22.

2.3 Born rule using pre-probabilities

★ One often finds (1) written in the form

$$\Pr(\phi^k) = |\langle \phi_1^k | \psi_1 \rangle|^2, \tag{11}$$

where

$$|\psi_1\rangle = T(t_1, t_0)|\psi_0\rangle \tag{12}$$

is the result of integrating Schrödinger's equation from t_0 to t_1 starting with $|\psi_0\rangle$ as the initial state.

• Not only is (11) very compact, it is also an efficient way of calculating probabilities for a number of different k values, given that one is always interested in a particular state $|\psi_0\rangle$ at t_0. One only has to integrate Schrödinger's equation once—this is usually the most expensive part of the calculation—and then evaluate several inner products, which is relatively simple.

★ While there is no problem in regarding $|\psi_1\rangle$ as a convenient *calculational tool*, a grave problem arises if one attempts to think of it as representing a *physical property* at time t_1, at least when using the sample space (7) of properties associated with the orthonormal basis elements $\{|\phi_1^k\rangle\}$. The reason is that $|\psi_1\rangle\langle\psi_1|$ will in general not commute with the projectors $|\phi_1^k\rangle\langle\phi_1^k|$.

□ Exercise. Show that unless all the probabilities in (1) except one are zero, there will be some $|\phi_1^k\rangle\langle\phi_1^k|$ that does not commute with $|\psi_1\rangle\langle\psi_1|$.

• For this reason it is best to think of $|\psi_1\rangle$ *not* as representing a physical property but instead as a calculational tool for computing probabilities by means of the Born rule. In CQT it is referred to as a *pre-probability.*

★ In classical physics probabilities are not physical properties, instead they are part of the conceptual or theoretical machinery we use to think about physical properties or events in the presence of some uncertainty. In this sense probabilities are less "real" than the events to which they refer. In quantum mechanics, pre-probabilities, since they are tools for calculating probabilities, are if anything somewhat less real than the probabilities, and thus definitely less real than physical properties. Keeping this in mind is helpful in avoiding the confusion that can accompany the concept of a quantum "state" or "wave function." Sometimes this denotes a real physical property, but often it is being used as pre-probability, and the failure of textbooks to make this distinction is one of the sources of difficulty that students have learning the subject.

◦ For more on pre-probabilities see CQT Sec. 9.4

• Even in classical physics the term "state" can have more than one meaning. Sometimes it refers to a particular mechanical configuration corresponding to a point in the classical phase space. But it is also used for a probability distribution on the phase space. Similarly in quantum theory "state" can have various meanings. Sometimes it is a "pre-probability," the counterpart of a classical probability distribution, and sometimes a quantum "property," which corresponds to either a point or a region in the classical phase space.

2.4 Generalizations of the Born rule

• The *rank* of a projector P is the number of nonzero eigenvalues, which is also its trace and the dimension of the subspace onto which it projects. If this subspace is one-dimensional, a ray, then P can be written as a dyad, $P = |\psi\rangle\langle\psi|$, and is said to be a *pure-state* projector. But in general a projector may project onto a space of higher dimension.

★ The Born rule in (1) is written as

$$\Pr(P_k|\psi_0) = \langle\psi_0|T(t_0,t_1)P_kT(t_1,t_0)|\psi_0\rangle = \langle\psi_1|P_k|\psi_1\rangle \tag{13}$$

when at t_1 one is considering a general decomposition $\{P_k\}$ of the identity I. Since the dyads $P_k = |\psi_1^k\rangle\langle\psi_1^k|$ for an orthonormal basis form such a decomposition, (1) is a special case of (13), and we shall henceforth refer to (13) as "the Born rule," regarding (1) as a particular example or application.

★ The next generalization seeks to treat t_1 and t_0 in a fully symmetrical fashion We proceed in two steps. First let us suppose that at t_0 we are interested in an orthonormal basis $\{|\psi_0^j\rangle\}, j = 1, 2, \ldots$ comparable to, but in general different from, the orthonormal basis $\{|\phi_1^k\rangle\}$ at t_1. Define the matrix (which could also be written M_{kj})

$$M(k,j) = |\langle\phi_1^k|T(t_1,t_0)|\psi_0^j\rangle|^2 = |\langle\psi_0^j|T(t_0,t_1)|\phi_1^k\rangle|^2. \tag{14}$$

• Because the basis elements are normalized, one has

$$\sum_j M(k,j) = 1 \text{ for all } k, \quad \sum_k M(k,j) = 1 \text{ for all } j. \tag{15}$$

∘ Such a matrix is said to be *doubly stochastic*. A matrix in which only the second condition holds is a *stochastic* matrix in the theory of Markov processes (see separate notes on probabilities) if one uses the physicists' convention for the order of j and k. (Unfortunately, the opposite convention is employed by probabilists.)

• Note that there is nothing in (14) that requires that t_0 is a time earlier than t_1; the formula works just as well for either order. Of course $T(t_1, t_0)$ may very well be a different operator if its two arguments are interchanged, but that is a separate issue.

• Just as in the theory of Markov processes, the $M(k, j)$ matrix is not itself a collection of probabilities, but it can be used along with other information (real or hypothetical) to generate a probability distribution. If, for example, one has reason to assign for every j a probability $p_0(j)$ to ket $|\psi_0^j\rangle$ (or the corresponding ray) of the quantum system at time t_0, then the joint probability for $|\psi_0^j\rangle$ at $t = 0$ and $|\phi_1^k\rangle$ at t_1 is given by

$$\Pr(\psi_0^j, \phi_1^k) = \Pr(j,k) = M(k,j)p_0(j). \tag{16}$$

∘ The Born rule as stated in (1) is a special case of this procedure when one assumes that $p_0(j = J) = 1$ for some particular value J of j, so zero for all other values, and $|\psi_0\rangle$ in (1) is $|\psi_0^J\rangle$.

∘ Note that one could just as well begin by assuming probabilities $p_1(k)$ for the $|\phi_1^k\rangle$ at time t_1, and from these construct a joint probability distribution $\Pr(j, k) = p_1(k)M(k, j)$.

★ The next generalization is to extend (14) to the case of general decompositions of the identity

$$I = \sum_j P_j \text{ at } t_0, \quad I = \sum_k Q_k \text{ at } t_1. \tag{17}$$

where it becomes:

$$M(k,j) = \text{Tr}\left[Q_k T(t_1,t_0) P_j T(t_0,t_1)\right]. \tag{18}$$

• But now (15) is replaced by

$$\sum_j M(k,j) = \text{Tr}(Q_k), \quad \sum_k M(k,j) = \text{Tr}(P_j), \tag{19}$$

so these sums are no longer, in general, equal to 1, but to the dimension of the space onto which Q_k or P_j projects. Of course if either of these is 1, we are back to the previous case.

• One can again use the $M(k,j)$ to generate a joint probability distribution by using other information or making additional assumptions. Once again assume one has reason to assign to each property P_j at t_0 a probability $p_0(j)$. Then in place of (16) write

$$\Pr(P_j, Q_k) = \Pr(j,k) = M(k,j)p_0(j)/\text{Tr}(P_j), \tag{20}$$

where the factor of $1/\text{Tr}(P_j)$ is obviously necessary, given (19), in order to have $\sum_k \Pr(j,k) = p_0(j)$.

★ But what is the quantum *sample space* that $\Pr(P_j, Q_k)$ refers to? Especially if, as is often the case, some or perhaps all of the projectors P_j do not commute with some (or all) of the projectors Q_k?

• The answer can be understood through the following analogy. The sample space for an ordinary coin tossed twice in a row is exactly the same as the sample space for two coins tossed just once. Similarly we might expect that the sample space for one qubit at two successive times is at least formally the same as that of two qubits at the same time. The Hilbert space for two qubits (see separate notes on composite systems) is the *tensor product* of Hilbert spaces for the individual qubits. Thus:

★ The Hilbert space used to represent the properties of a quantum system at two different times is the tensor product $\mathcal{H} \odot \mathcal{H}$, with $\odot$ nothing but an alternative form of the symbol $\otimes$, where $\mathcal{H}$ is the Hilbert space appropriate to that system at a single time.

◦ The generalization to three or more times should be obvious.

• Therefore the sample space of the two-time quantum system must be some decomposition of the identity $I \odot I$ (or $I \otimes I$). The decomposition employed in the preceding discussion is the *product* of sample spaces,

$$I \odot I = \sum_{jk} P_j \odot Q_k, \tag{21}$$

at the individual times. That is, for every pair of subscripts j, k, $P_j \odot Q_k$ is a projector representing a quantum property, in fact the property: "P_j at time t_0 AND Q_k at time t_1." It is for this sample space, and thus to properties of this sort, that probabilities $\Pr(P_j, Q_k)$ are assigned by, for example, (20).

2.5 Limitations of the Born rule

★ The Born rule has a serious limitation in that it can only be used for *two* times, an initial time and a final time. Naive attempts to extend it to three or more times result in paradoxes and contradictions. Textbook quantum mechanics deals with this situation through a set of ad hoc rules referring to measurements and wave function collapse, a witches' brew that even the experts sometimes admit they don't understand very well. A more systematic approach requires the use of quantum histories and consistency conditions (Chs. 8, 10 and 11 of CQT).

3 Two qubits

3.1 Pure states at t_1

• Studying two qubits will provide a number of examples for applying the Born rule.

★ Consider the controlled-not gate in Fig. 1.

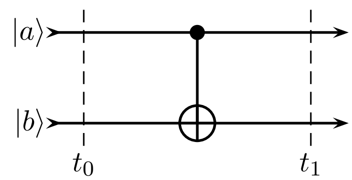

Figure 1: Controlled-not gate

• Suppose the initial state is $|ab\rangle = |10\rangle$ and at the final time we use the standard or computational basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. Then since for the controlled-not gate $T(t_1, t_0)|10\rangle = |11\rangle$, we see that at t_1, $\Pr(11) = 1$ and all the other probabilities are 0. By contrast, if the initial state is $|{+0}\rangle$ then

$$|\psi_1\rangle = T(t_1, t_0)|{+0}\rangle = (|00\rangle + |11\rangle)/\sqrt{2}. \tag{22}$$

One can then "read off" the probabilities as the absolute squares of the coefficients in this expression for $|\psi_1\rangle$:

$$\Pr(00) = 1/2 = \Pr(11), \quad \Pr(01) = 0 = \Pr(10). \tag{23}$$

• Note how this works in general. To find the probabilities corresponding to the basis $\{|\phi_1^k\rangle\}$, expand $|\psi_1\rangle = T(t_1, t_0)|\psi_0\rangle$ in this basis:

$$|\psi_1\rangle = \sum_k c_k |\phi_1^k\rangle; \quad \Pr(\phi_1^k) = |c_k|^2 \tag{24}$$

★ Beware of treating $|\psi_1\rangle$ (or the corresponding ray) as a physical property at t_1; see the discussion in Sec. 2.3. In general its projector $|\psi_1\rangle\langle\psi_1|$ does not commute with the projectors for the different states $|\phi_1^k\rangle$ of the sample space.

∘ There is, however, an exception. If $|\psi_1\rangle$ is a member of the orthonormal basis $\{|\phi_1^k\rangle\}$, then the Born rule implies that it will occur with probability 1, and all other possibilities have probability zero. In this case $|\psi_1\rangle$ can represent the real physical state, just as a probability of 1 is an exception to the general rule that probabilities provide only partial knowledge.

★ Another way of calculating probabilities is to use the decomposition of the identity

$$I = \sum_k [\phi_1^k] = \sum_k |\phi_1^k\rangle\langle\phi_1^k|, \tag{25}$$

and write

$$|c_k|^2 = |\langle\phi_1^k|\psi_1\rangle|^2 = \langle\psi_1|\phi_1^k\rangle\langle\phi_1^k|\psi_1\rangle$$
$$= \langle\psi_1|[\phi_1^k]|\psi_1\rangle = \Pr(\phi_1^k), \tag{26}$$

which is using Born's rule in the form (15).

3.2 Properties of one qubit at t_2

★ Suppose we have two qubits a, b and $T(t_1, t_0)|\psi_0\rangle = |\psi_1\rangle$. What is the probability that qubit a is in the state $|1\rangle$, aside from whatever qubit b may be doing?

• One way to address this is to expand $|\psi_1\rangle$ in the standard basis as

$$|\psi_1\rangle = \sum_{jk} c_{jk}|jk\rangle, \tag{27}$$

so that

$$\Pr(j,k) = \Pr(a{=}j, b{=}k) = |c_{jk}|^2 \tag{28}$$

and then calculate the *marginal* probability distribution using the usual formula from ordinary probability theory:

$$\Pr(a{=}j) = \sum_k \Pr(j,k), \tag{29}$$

from which it follows that $\Pr(a{=}1) = |c_{10}|^2 + |c_{11}|^2$.

• But why use the standard basis? It is true that the question asks for the probability that a is in the state $|1\rangle$, which means we have to use the standard basis for qubit a, but we could use some other basis, such as $\{|+\rangle, |-\rangle\}$, for qubit b. Would we get the same answer? Yes, but it is more obvious if we use an alternative approach based on the decomposition of the identity

$$I = [0]\otimes I + [1]\otimes I = [0]_a + [1]_a, \tag{30}$$

where $[0]_a$ means the operator $|0\rangle\langle 0|$ on qubit a. Then use (13):

$$\Pr(a{=}1) = \langle\psi_1|[1]_a|\psi_1\rangle. \tag{31}$$

The right side makes no reference to any basis for b.

• Another way to evaluate this probability is to take $|\psi_1\rangle$ and "expand it" in the orthonormal basis of a which interests us,

$$|\psi_1\rangle = |0\rangle\otimes|\beta^0\rangle + |1\rangle\otimes|\beta^1\rangle, \tag{32}$$

where the kets $|\beta^j\rangle$ belonging to the Hilbert space of qubit b are simply expansion coefficients; in general they are neither normalized nor orthogonal to each other. It is then easy to see that (31) implies that

$$\Pr(a{=}1) = \langle\beta^1|\beta^1\rangle, \tag{33}$$

and, once again, there is no reference to any orthonormal basis for b, since the $|\beta^j\rangle$ in (32) are uniquely determined once one has chosen a basis for a.

◦ To see that the expansion (32) is possible, it is helpful to introduce an orthonormal basis for b—any one will do, so we might as well use the standard basis—and expand $|\psi_1\rangle$, as in (27). Then by collecting terms one sees that

$$|\beta^j\rangle = \sum_k c_{jk}|k\rangle. \tag{34}$$

3.3 Correlations and conditional probabilities

★ The probabilities for a and b given in (23) are *correlated* with one another in the sense that they are not statistically independent. A useful way to view such correlations is through *conditional probabilities*.

• See separate notes on probabilities.

• In the case of interest to us, the joint probability distribution (23) implies that

$$\Pr(a{=}0) = 1/2 = \Pr(a{=}1), \quad \Pr(b{=}0) = 1/2 = \Pr(b{=}1), \tag{35}$$

and therefore

$$\Pr(b{=}0 \,|\, a{=}0) = \Pr(00)/\Pr(b{=}0) = 1 \neq \Pr(b{=}0) = 1/2. \tag{36}$$

That is, given that $a = 0$, we can be sure (conditional probability equal to 1) that $b = 0$, and the converse is also true: if $b = 0$, then surely $a = 0$ as well. (Here, as in the arguments of Pr, "$a = 0$" is shorthand for "a is in the state $|0\rangle$" or "a has the property $[0]$.")

★ These formulas apply even if qubits a and b are widely separated from each other, say in different laboratories.

• There is nothing "magical" about correlations of this sort, and it is easy to construct similar examples in classical stochastic processes. Unfortunately the situation has become somewhat muddied in quantum textbooks because of the careless introduction probabilities by means of measurements and an associated wave function "collapse" that is supposedly produced by measurements. Once one realizes that this collapse is nothing but an algorithm for computing conditional probabilities the magic and its associated confusion disappears.

∘ Measurements discussed in a separate set of notes.