Notes on Zurek's derivation of the quantum probability rule

C. M. Caves

2004 January 29; modified 2005 July 29

This is my take on WHZ's derivation of the quantum probability rule. I worked this out myself in December of 2003 after reading carefully WHZ's PRL [1] but then found that much of it had been anticipated by Schlosshauer and Fine [2] and by Barnum [3] in their discussions of WHZ's derivation. My assessment is perhaps harsher than either of theirs, however, and pithier than Howard's, which has the typical Howardian underbrush of verbiage and caveats. The document was modified at the time of the Being Bayesian in a Quantum World meeting in Konstanz, where I gave brief presentation based on it.

It is hard to tell from WHZ's discussion whether he sees his derivation as justifying the Born rule as the way for an observer to assign subjective probabilities or as the rule for objective probabilities that adhere within a relative state. For example, in his PRL discussion of a system-environment state in Schmidt form, WHZ comments, "Given the state of the combined $\mathcal{SE}$ in Schmidt form—with complex α_k and with $\{|\sigma_k\rangle\}$ and $\{\varepsilon_k\rangle\}$ orthonormal—what sort of invariant *quantum facts* can be known about $\mathcal{S}$?" The point of his argument appears to be that the probabilities are invariant quantum facts, as opposed to the phases of the coefficients α_k, which are said to be unknown. On the other hand, if he is after ontological probabilities, why would he begin the paper with a paraphrase of Schrödinger, "One can know precisely the state of a composite object (consisting, for example, of the system $\mathcal{S}$ and the environment $\mathcal{E}$) and yet be ignorant of the state of $\mathcal{S}$ alone," and why, in the part of his argument that gets equal probabilities for equal amplitudes, would he state, "When all of the coefficients of swapped states are equal, the observer with access to $\mathcal{S}$ alone cannot detect the effect of the swap." Mohrhoff [4] does a good job of drawing attention to the philosophical inconsistencies in WHZ's presentation. I concentrate here on the details of the derivation, and I assume that WHZ is thinking in terms of objective probabilities associated with relative states, as seems to be clear from the abstract of his follow-on PRA [5], where he writes, "The probabilities derived in this manner are an objective reflection of the underlying state of the system—they represent experimentally verifiable symmetries, and not just a subjective 'state of knowledge' of the observer."

What WHZ would like to do is the following. He wants to find the probability

[1] W. H. Zurek, "Environment-assisted invariance, entanglement, and probabilities in quantum physics," *Phys. Rev. Lett.* **90**, 120404 (2003), `quant-ph/0211037`.

[2] M. Schlosshauer and A. Fine, "On Zurek's derivation of the Born Rule," *Found. Phys.* **35**, 197–213 (2005), `quant-ph/0312058`.

[3] H. Barnum, "No-signalling-based version of Zurek's derivation of quantum probabilities: A note on 'Environment-assisted invariance, entanglement, and probabilities in quantum physics,' " `quant-ph/0312150`.

[4] U. Mohrhoff, "Probabilities from envariance?" *Int. J. Quant. Inf.* **2**, 221 (2004), `quant-ph/0401180`.

[5] W. H. Zurek, "Probabilities from entanglement, Born's rule $p_k = |\psi_k|^2$ from envariance," *Phys. Rev. A* **71**, 052105 (2005), `quant-ph/0405161`.

$p_A\big(j;\{|a_k\rangle\},|\psi_A\rangle\big)$ for outcome j associated with an orthonormal basis $\{|a_k\rangle\}$ when system A has state vector $|\psi_A\rangle$. (We use this very pedantic notation to be sure that we don't lose track of any of any dependencies that might be important on the way to the standard rule.) Thus he clearly assumes the standard Hilbert-space structure of pure states and quantum questions as rays in Hilbert space. The state vector is expanded in the outcome basis as

$$|\psi_A\rangle = \sum_k \alpha_k |a_k\rangle \ .$$

He begins with the further assumption that the very notion of "outcomes"—in the language of measurement theory, the notion of a measurement with these outcomes—means that the system interacts via a unitary U with an "environment" B, which begins in a standard state $|e_0\rangle$, in such a way that after the system and environment interact, the overall state takes the Schmidt form

$$U|\psi_A\rangle \otimes |e_0\rangle = |\Psi_{AB}\rangle = \sum_k \alpha_k |a_k\rangle \otimes |b_k\rangle \ ,$$

with the environment states $|b_k\rangle$ being orthonormal. (For any interaction, one can write a relative-state decomposition based on the systen states $|a_k\rangle$, but the relative environment states are not generally orthogonal.) Thus he further assumes the tensor-product structure for composite systems and that unitaries describe quantum dynamics. Notice that what I am saying is that in WHZ's approach, it is the Schmidt relative state that *defines* the notion of outcomes for system A; without the entanglement with system B, one cannot even talk about outcomes for the basis $\{|a_k\rangle\}$.

It is built into this discussion—built into the formulation at the most fundamental level—that WHZ also assumes that the probabilities don't depend on the environmental basis $|b_k\rangle$ that becomes "correlated" with the outcome basis, because it has already been assumed that the probabilities that he is seeking, $p_A\big(j;\{|a_k\rangle\},|\psi_A\rangle\big)$, have no dependence on the environmental states $|b_k\rangle$. This is a kind of foundational noncontextuality assumption that underlies the whole approach. I will call it *environmental noncontextuality* (EN) for lack of a better name. Let's give him all the assumptions made up till now.

Further unitary evolution of the environment after the interaction does not change the probabilities. This is not really an additional assumption because all such a unitary can do is to change the orthonormal environment states $|b_k\rangle$, and it's already been assumed that this basis doesn't affect the probabilities. It does, however, allow us to rewrite the desired probability as

$$p_A\big(j;\{|a_k\rangle\},|\psi_A\rangle\big) = F\big(|a_j\rangle;|\Psi_{AB}\rangle\big) \ .$$

Here F is a function of (i) joint AB pure states whose Schmidt decomposition picks out the outcome basis on A and (ii) a vector in the outcome basis. The function has the property

$$F\big(|a_j\rangle;|\Psi_{AB}\rangle\big) = F\big(|a_j\rangle;1_A \otimes U_B|\Psi_{AB}\rangle\big) \ ;$$

i.e., F is constant on the equivalence classes of joint pure states defined by the action of unitaries U_B on system B. This is assumption (4) of Schlosshauer and Fine, and Howard calls this the *no-signalling assumption*, because changes in the probabilities at A as a consequence of actions taken at B could be used to send signals. This assumption is the expression of what we called EN above, but it is a neat way to write EN and tease out its consequences.

Now we're ready to introduce WHZ's concept of *envariance*: $|\Psi_{AB}\rangle$ is *envariant* under a unitary U_A on A if the effect of U_A on $|\Psi_{AB}\rangle$ can be compensated by a unitary U_B on B, i.e., $U_A \otimes U_B|\Psi_{AB}\rangle = |\Psi_{AB}\rangle$. We see immediately that for an envariant transformation U_A,

$$F\big(|a_j\rangle;|\Psi_{AB}\rangle\big) = F\big(|a_j\rangle;U_A\otimes U_B|\Psi_{AB}\rangle\big) = F\big(|a_j\rangle;U_A\otimes 1_B|\Psi_{AB}\rangle\big)\ ,$$

or translated to probabilities,

$$p_A\big(j;\{|a_k\rangle\},|\psi_A\rangle\big) = p_A\big(j;\{U_A|a_k\rangle\},U_A|\psi_A\rangle\big)\ .$$

WHZ wants to view envariance as the key to his derivation, but it is just a way to write the consequences of EN, when they provide any useful constraints, in terms of system unitaries, instead of environment unitaries. It turns out not to be necessary to translate EN to system unitaries for any of the steps in the derivation.

WHZ uses envariance, however, to draw his first big conclusion. A unitary U_A that is diagonal in the outcome basis, i.e., that introduces phases in the outcome basis, is envariant because it can be compensated by a U_B that introduces the opposite phases in the $|b_j\rangle$ basis. This implies, of course, that the probabilities $p(j;\{|a_k\rangle\},|\psi_A\rangle)$ are independent of the phases of the amplitudes α_k. Though this is an important conclusion, it is, within the assumptions, a recognition of the ambiguity in where phases are placed in the Schmidt decomposition (i.e., in the amplitudes or in the states) and thus is a direct consequence of EN, without the need to invoke envariance.

Now we get to the more important use of envariance, the case where several of the amplitudes α_k have the same magnitude. The objective is to use envariance to show that the corresponding probabilities are equal. This result in hand, one is off and running with well developed techniques to get the standard quantum rule, although we will want to examine carefully the justification of these well developed techniques in the present context. The method is illustrated without loss of generality by assuming that there are only two outcome states, $|0\rangle$ and $|1\rangle$, and that the system state is $|\psi_A\rangle = (|0\rangle+|1\rangle)/\sqrt{2}$, with corresponding Schmidt state $|\Psi_{AB}\rangle = (|00\rangle+|11\rangle)/\sqrt{2}$. The unitary U_A that swaps $|0\rangle$ and $|1\rangle$ is envariant because it can be compensated by a counter-swap on system B. So we have from envariance that

$$\begin{aligned}
p_A\big(0;\{|0\rangle,|1\rangle\},|\psi_A\rangle\big) &= F\big(|0\rangle;|(|00\rangle+|11\rangle)/\sqrt{2}\big) = F\big(|0\rangle;|(|10\rangle+|01\rangle)/\sqrt{2}\big)\ ,\\
p_A\big(1;\{|0\rangle,|1\rangle\},|\psi_A\rangle\big) &= F\big(|1\rangle;|(|00\rangle+|11\rangle)/\sqrt{2}\big) = F\big(|1\rangle;|(|10\rangle+|01\rangle)/\sqrt{2}\big)\ ,
\end{aligned}$$

We want to get $p_A\big(0;\{|0\rangle,|1\rangle\},|\psi_A\rangle\big) = p_A\big(1;\{|0\rangle,|1\rangle\},|\psi_A\rangle\big)$, and one can see immediately that the envariance-inspired relations by themselves aren't going to help at all in reaching this conclusion because they can never change the alternative under consideration from 0 to 1. All they say is that there are two different forms for each of these two probabilities, without providing any connection between them. Moreover, notice that these relations follow directly from EN: they're an immediate consequence of the formulating the notion of outcomes and probabilities in terms of a Schmidt state.

We need a further assumption to get anywhere. We can exchange the roles of A and B, imagining that we are trying to determine probabilities for "outcomes" j associated with the basis $\{|b_k\rangle\}$ for system B in a state

$$|\psi_B\rangle = \sum_k \alpha_k |b_k\rangle \ .$$

Given the same formulation as for A, we can write the probabilities for the outcomes as

$$p_B\big(j; \{|b_k\rangle\}, |\psi_B\rangle\big) = G\big(|b_j\rangle; |\Psi_{AB}\rangle\big) \ ,$$

where G plays the role of F. The assumption we need is that

$$F\big(|a_j\rangle; |\Psi_{AB}\rangle\big) = G\big(|b_j\rangle; |\Psi_{AB}\rangle\big) \ ,$$

when the amplitudes α_k have equal magnitudes. This is apparently WHZ's "pedantic" assumption in his PRL, which he gets by fiddle-faddling around because he hasn't given a precise mathematical formulation of what he's doing (you won't find in his paper any equations that spell out explicitly how the probabilities depend on the other mathematical objects). This is assumption (3) of Schlosshauer and Fine, and Howard calls it the Perfect Correlation Principle (PCP), which is what I will call it. It is another kind of noncontextuality assumption, since F and G arise in entirely different contexts; PCP seems natural because of the symmetry between A and B in the joint state $|\Psi_{AB}\rangle$. The point is that WHZ's derivation depends on an unstated assumption that one can interchange the roles of systems A and B in the case of Schmidt states with amplitudes of equal magnitude. In his PRA, WHZ has moved to asserting that he doesn't need this assumption, but again the reason is that he hasn't given a precise mathematical formulation of what he is doing.

If we had the idea that $F\big(|a_j\rangle; |\Psi_{AB}\rangle\big)$ and $G\big(|b_j\rangle; |\Psi_{AB}\rangle\big)$ were marginal probabilities for outcomes on the joint state, then we could justify PCP using fairly minimal assumptions. This way of thinking isn't legal, because we're trying to derive the probability rule for A (or B) by mucking around with the joint system AB, so introducing any probability concepts on AB is not allowed. Still, if we imagine that there is a joint probability $P_{AB}\big(|a_j\rangle, |b_k\rangle; |\Psi_{AB}\rangle\big)$, whose marginals are F and G, then it is perhaps a very minimal assumption to write

$$P_{AB}\big(|a_j\rangle, |b_k\rangle; |\Psi_{AB}\rangle\big) = 0 \quad \text{for} \quad j \neq k \ .$$

Assuming that F and G are the marginals of P_{AB}, we get the desired equality immediately:

$$\begin{aligned} F\big(|a_j\rangle; |\Psi_{AB}\rangle\big) &= \sum_k P_{AB}\big(|a_j\rangle, |b_k\rangle; |\Psi_{AB}\rangle\big) \\ &= P_{AB}\big(|a_j\rangle, |b_j\rangle; |\Psi_{AB}\rangle\big) \\ &= \sum_k P_{AB}\big(|a_k\rangle, |b_j\rangle; |\Psi_{AB}\rangle\big) = G\big(|b_j\rangle; |\Psi_{AB}\rangle\big) \ . \end{aligned}$$

Again, however, this line of argument is not legitimate, so any good feeling it attaches to PCP should be discounted.

Now we can go back to our discussion of the two alternatives, and go straight to the result, as do Schlosshauer/Fine and Barnum:

$$\begin{aligned}
p_A\big(0;\{|0\rangle,|1\rangle\},|\psi_A\rangle\big) &= F\big(|0\rangle;(|00\rangle+|11\rangle)/\sqrt{2}\big) && \text{(starting point)}\\
&= F\big(|0\rangle;(|10\rangle+|01\rangle)/\sqrt{2}\big) && \text{(envariance of swaps on } A)\\
&= G\big(|1\rangle;(|10\rangle+|01\rangle)/\sqrt{2}\big) && \text{(PCP)}\\
&= G\big(|1\rangle;(|00\rangle+|11\rangle)/\sqrt{2}\big) && \text{(envariance of swaps on } B)\\
p_A\big(1;\{|0\rangle,|1\rangle\},|\psi_A\rangle\big) &= F\big(|1\rangle;(|00\rangle+|11\rangle)/\sqrt{2}\big) && \text{(PCP)}\;.
\end{aligned}$$

When you examine this string of equalities carefully, you find that the portrayal of envariance as the key assumption is just an artful dodge. The two envariance steps come directly from EN, without detouring through envariance, and EN is an underlying noncontextuality assumption. The two PCP steps are just as critical, and they are also noncontextuality assumptions. Both Schlosshauer/Fine and and Howard are onto this, in somewhat different language.

WHZ goes on to a standard argument that gets one from equal probabilities in the equal amplitude case to the general quantum probability rule for unequal amplitudes. Schlosshauer and Fine let him off the hook completely on this further argument, and Howard only comments about the need for a continuity assumption to get away from rational probabilities. My initial inclination was also to let him off the hook, but I now think there is a real question about this further argument, having nothing to do with continuity.

We'll consider only the two-outcome version of the argument, which goes as follows. The state of system A is $|\psi_A\rangle = \alpha|0_A\rangle + \beta|1_A\rangle$, where we assume that $|\alpha|^2 = m/M$ is rational or can be sufficiently well approximated by a rational approximation. We imagine that in the correlated state that defines the notion of outcomes,

$$|\Psi_{AB}\rangle = \alpha|0_A\rangle\otimes|0_B\rangle + \beta|1_A\rangle\otimes|1_B\rangle\;,$$

the two states of system B can be written as

$$|0_B\rangle = \frac{1}{\sqrt{m}}\sum_{l=1}^{m}|b_l\rangle \qquad\text{and}\qquad |1_B\rangle = \frac{1}{\sqrt{M-m}}\sum_{l=m+1}^{M}|b_l\rangle\;.$$

This requires, of course, that B have at least M dimensions, in which case one can always find a basis in which $|0_B\rangle$ and $|1_B\rangle$ have the desired form. The joint state of systems A and B assumes the form

$$|\Psi_{AB}\rangle = \sum_{l=1}^{m}\frac{\alpha}{\sqrt{m}}|0_A\rangle\otimes|b_l\rangle + \sum_{l=m+1}^{M}\frac{\beta}{\sqrt{M-m}}|1_A\rangle\otimes|b_l\rangle\;,$$

which leaves all expansion coefficients with the same magnitude, $1/\sqrt{M}$. One can argue about this approach in any context where it is used, but here we're only interested in its use within WHZ's derivation.

To fit the state into his derivation, what he does is to say that now we have to imagine that there is yet another system, C, that interacts with the first two systems in such a way that the final state of all three systems is of the desired Schmidt form:

$$|\Psi_{ABC}\rangle = \sum_{l=1}^{m} \frac{\alpha}{\sqrt{m}} |0_A\rangle \otimes |b_l\rangle \otimes |c_l\rangle + \sum_{l=m+1}^{M} \frac{\beta}{\sqrt{M-m}} |1_A\rangle \otimes |b_l\rangle \otimes |c_l\rangle \ .$$

This step is necessary to define the notion of the outcomes $0l$ and $1l$ for systems A and B. WHZ uses his previous work, applied to AB vs. C, to conclude that the outcomes jl have equal probabilities, given by $1/M$, and thus that the probability for outcome 0 is

$$p_A\big(0; \{|0_A\rangle, |1_A\rangle\}, |\psi_A\rangle\big) = \sum_{l=1}^{m} p_{AB}(0l; \{|j_A\rangle \otimes |b_l\rangle\}, |\Psi_{AB}\rangle) = \sum_{l=1}^{m} \frac{1}{M} = \frac{m}{M} = |\alpha|^2 \ ,$$

which, modulo the continuity assumption, gets him the standard quantum rule.

But let's think about this for a minute. We were originally told that the very notion of outcomes for system A required us to think about a joint pure state with the appropriate Schmidt decomposition. Now we're told that the notion of outcomes requires us to think about a much more complicated three-system joint state, where the two additional systems must have a dimension big enough to accommodate the rational approximation to the desired probabilities. Does this mean the notion of outcomes depends on the value of the amplitudes? This is a very unattractive alternative, so what we really must think is that for all amplitudes, the notion of outcomes requires us to think in terms of a big three-system joint state, where B and C have arbitrarily large dimensions. We're now supposed to believe that the notion of outcomes for system A requires us to think in terms of two other systems correlated in a particular way, which has no apparent relation to the number of outcomes of system A. Even a relative-state believer would find this hard to swallow, and it makes the PCP assumption far less natural, because this construction wrecks the nice-looking symmetry between A and the systems to which it is coupled and even between AB and C. It is a heck of a lot less attractive than the original picture we were presented and really should have been stated at the outset.

In the end one is left wondering what makes the envariance argument any more compelling than just asserting that a swap symmetry means that a state with equal amplitudes has equal probabilities and then moving on to the argument that extends to rational amplitudes.