Relating Probability Amplitude Mechanics to Standard Statistical Models

Robert F. Bordley*

June 15, 2006

Abstract

The probabilities assessed for an event by a detailed experiment (e.g., one which measures both a particle's position on a detection screen and which of several slits it traversed to reach that screen) may differ from the probabilities assessed for that same event by a less detailed experiment (e.g., one which only measures the particle's position on the detection screen.) As this paper shows, probability amplitude mechanics writes the detailed experiment probabilities as a weighted average of the less detailed experiment probabilities and some factor measuring the variability overlooked by the less detailed experiment.

1 MOTIVATION

The n-slit interference experiments show that the probability assessed for a specific event, e.g., the particle reaching a specific point on a detector screen, will vary depending upon whether the experiment measuring that event was detailed (e.g., also assessed which slit the particle traversed en route to the screen) or less detailed (only assessed the particle's position on the detector screen). This paper will show that probability amplitude mechanics implies that the probability for an event, given a detailed experiment, is a weighted average of

- the probability of that same event, given a less detailed experiment, and
- a factor measuring the variability observed by the detailed experiment but overlooked by the less detailed experiment.

We then show that this alternate representation of probability amplitude mechanics might be deducible from fairly simple statistical assumptions. This

*Operating Sciences Department; General Motors Research Labs; Warren, Michigan 48090-9055

suggests the possibility of deriving probability amplitude mechanics from intuitive first principles, a result which would help resolve interpretational issues associated with quantum theory[1].

2 THE WEIGHTED AVERAGE FORMULA

2.1 More Detailed and Less Detailed Experiments

Let Ω be the universal set and $\emptyset$ be the null set. Following operational statistics(Foulis & Randall,1972a,1972b), we define the set of possible outcomes resulting from applying a specific operation C to a system by the partition

$$\Gamma_C = (A | \cup_A = \Omega; A_i \cap A_j = \emptyset, i \neq j)$$

Now consider a second operation F whose possible outcomes are represented by the partition

$$\Gamma_F = (E | \cup_E = \Omega; E_i \cap E_j = \emptyset, i \neq j)$$

If every possible outcome of F is a subset of every possible outcome of C, i.e., if $E \in \Gamma_F$ implies there exists an $A \in \Gamma_C$ with $E \subset A$, then operation C is less detailed (or more 'coarse') than operation F.

For operational statistics, the probability of observing event A, given we perform operation C, is $\Pr(A|\Gamma_C)$, the probability assessed for the state A given an operation whose possible outcomes are represented by the elements of Γ_C. Now consider the probability for this same state A if we planned to perform the more detailed operation F instead of C, i.e., $\Pr(A|\Gamma_F)$)[2].

2.2 Probability Amplitude Mechanics

To determine this probability, define the variance of event A as the sum of the variance of the real and imaginary wave function components of each event $E \in A$. Then the next section proves:

Theorem: Probability Amplitude Mechanics implies

$$\Pr(A|\Gamma_F) = (1 - \rho) \Pr(A|\Gamma_C) + \rho V_A$$

- where $\rho = \frac{r}{1+r}$ and r is the coefficient of variation of events $E \in A$, averaged over all $A \in \Gamma_C$. Thus r measures the variability overlooked by the less detailed experiment (and $0 \leq \rho \leq 1$.)

[1]When asked to explain probability amplitude mechanics, Feynman(1965) had replied, "We have no ideas about a more basic mechanism from which these results can be deduced."

[2]The fact that these two probabilities are different in the case of the n-slit interference experiment is, of course, one of the fundamental 'anomalies' of quantum mechanics.

- where V_A is the variance of event A divided by the summed variance of all events $A \in \Gamma_C$. The probability distribution, V_A, assigns higher probability to more variable states — somewhat like the maximum entropy principle.

When r is zero, there is no variability within the states, A, and everything the more detailed experiment measures is consistent with what the less detailed experiment had measured. In this case, $\Pr(A|\Gamma_F) = \Pr(A|\Gamma_C)$, i.e., the detailed and less detailed experiment assign the same state probabilities. When r is infinite, the more detailed experiment measures an infinitely large amount of variability not measured by the less detailed experiment. Hence the results of the less detailed experiment are essentially irrelevant and the probabilities assigned to various states are determined by V_A. When r is between zero and infinity, the detailed experiment's probability is intermediate between the less detailed experiment's probability and V_A. This, of course, is quite intuitive.

To write the formula in a somewhat more traditional matter, note that operational statistics mandates $\Pr(A|\Gamma_F) = \sum_{E \in A} \Pr(E|\Gamma_F)$, i.e., once we condition our probabilities on the same experiment, all the standard rules of probability apply. Making this substitution and rearranging the results of the Theorem gives

$$\Pr(A|\Gamma_C) = \Pr(A|\Gamma_F) + \frac{\rho}{1-\rho}(\Pr(A|\Gamma_F) - V_A) = \sum_{E \in A} \Pr(E|\Gamma_F) + r(\Pr(A|\Gamma_F) - V_A)$$

Hence the traditional interference term, $(\sum_{E,E^* \in A}[\Pr(E|\Gamma_F)\Pr(E^*|\Gamma_F)]^{.5}\cos(\theta_E - \theta_{E^*})$, has been replaced by $r(\Pr(A|\Gamma_F) - V_A)$. Note that as the amount of unmeasured variability, r goes to zero, probability amplitude mechanics gives additive probabilities. Hence we get a simple correspondence principle between classical and quantum mechanics.

2.3 Derivation from a Simple Statistical Model

The fact that the Theorem is so intuitive suggests that the same formula might be deducible from simpler statistical models. As a step toward constructing such a derivation, note that the probabilities assessed by the more detailed experiment, $\Pr(A|\Gamma_F)$, will be a function of those aspects of the system measured by the less detailed experiment (and reflected in $\Pr(A|\Gamma_C)$), and some aspects unmeasured by the less detailed experiment (reflected in some noise term.) The simplest statistical assumption about this noise term (given that it represents uncertainty about a probability) is that it follows a Dirichlet distribution. We similarly follow statistical mechanics in assuming that $\Pr(A|\Gamma_C)$ represents the mode (i.e., the most likely value) of this Dirichlet distribution.

Given these assumptions, the Appendix deduces $\Pr(A|\Gamma_F)$ as the mean of this distribution. We state this as a Lemma:
Lemma:$\Pr(A|\Gamma_F) = (1-\rho)\Pr(A|\Gamma_C) + \rho\frac{1}{|\Gamma_C|}$ where $0 \leq \rho \leq 1$.

which is identical to the result of the Theorem when each event has the same variance, i.e., $V_A = \frac{1}{|\Gamma_C|}$.

This special case does involve an interference term, $r(\Pr(A|\Gamma_F) - \frac{1}{|\Gamma_C|})$. But since $\Pr(A|\Gamma_C) > \sum_{E\in A}\Pr(E|\Gamma_F)$ if and only if $\Pr(A|\Gamma_F) > V_A$, the case of $V_A = \frac{1}{|\Gamma_C|}$ implies that whenever $\Pr(A|\Gamma_F)$ is larger then average, $\Pr(A|\Gamma_C) > \Pr(A|\Gamma_F)$ and whenever $\Pr(A|\Gamma_F)$ is smaller than average, $\Pr(A|\Gamma_C) < \Pr(A|\Gamma_C)$. Thus the Dirichlet special case implies that the probabilities from the less detailed experiment tend to be more extreme than the probabilities from the detailed experiment. Since this is not true in many quantum mechanical cases, future work will have to specify a distribution more general than the Dirichlet.

3 PROOF OF THE THEOREM

3.1 An Alternate Way of Writing the Wave Function

Probability amplitude mechanics writes the probability of A in terms of the wave functions, ϕ, as

$$\Pr(A|\Gamma_C) \propto \phi_A\phi_A^* = \frac{\phi_A\phi_A^*}{\sum_{A\in\Gamma_C}\phi_A\phi_A^*} \tag{1}$$

with

$$\phi_A = \sum_{E\in A}\phi_E \tag{2}$$

While the wave function, ϕ_E, is commonly written in the form $\phi_E = (\Pr(E))^{.5}\exp(i\theta_E)$, this paper writes it as $\phi_E = m_{E,R} + im_{E,I}$. Given this representation[3], (1) is equivalent to

$$\Pr(A|\Gamma_C) \propto [m_{A,R} + im_{A,I}][m_{A,R} - im_{A,I}] = \frac{m_{A,R}^2 + m_{A,I}^2}{\sum_{A\in\Gamma_C}(m_{A,R}^2 + m_{A,I}^2)} \tag{3}$$

For simplicity, we let Z be a dummy index which can either equal R or I.

3.2 The Variances associated with Wave Functions

For both values of Z, define the variances of the real and imaginary parts of the wave functions associated with events in A by

$$v_{A,Z} = \frac{1}{2|A|}\sum_{E,E^*\in A}(m_{E,R} - m_{E^*,R})^2 = \sum_{E\in A}m_{E,R}^2 - \frac{1}{|A|}(\sum_{E\in A}m_{E,R})^2 \tag{4}$$

[3]For Bohm(1952), $\theta_E = \frac{S_E}{h}$ which implies that $m_{E,R} = (\Pr(E))^{.5}\cos(S_E/h)$ and $m_{E,I} = (\Pr(E))^{.5}\sin(S_E/h)$ so that both m-functions can be negative.

These two variances, $v_{A,R}$ and $v_{A,I}$ measure the amount of variability 'internal' to the event A and which, presumably, would be overlooked by the less detailed experiment whose only outcomes are $A \in \Gamma_C$.

Equation (2) is equivalent to

$$m_{A,Z} = \sum_{E \in A} m_{E,Z} \quad for\ Z = R, I \tag{5}$$

Substituting (5) in (4) gives, with some rearrangement,

$$m^2_{A,Z} = |A|[\sum_{E \in A} m^2_{E,Z} - v_{A,Z}] \quad for\ Z = R, I \tag{6}$$

Defining $D_C = \sum_{A \in \Gamma_C, Z} m^2_{A,Z}$ and substituting (6) in (3) gives

$$\Pr(A|\Gamma_C) = \frac{|A|[\sum_{Z,E \in A} m^2_{E,Z} - \sum_Z v_{A,Z}]}{D_C} \tag{7}$$

Now consider an operation Γ_F which does measure the probability of each elemental event $E \in A$. Then the analogue of (3) gives

$$\Pr(E|\Gamma_F) = \frac{\sum_Z m^2_{E,Z}}{\sum_{Z,E \in \Gamma_F} m^2_{E,Z}} \tag{8}$$

Defining $D_F = \sum_{Z,E \in \Gamma_F} m^2_{E,Z}$, and substituting (8) in (7) gives

$$\Pr(A|\Gamma_C) = |A| \frac{\sum_{E \in A} \Pr(E|\Gamma_F) D_F - \sum_z v_{A,Z}}{D_C} \tag{9}$$

As a comparison, recall that writing the wave function as $\phi_E = (\Pr(E))^{.5} \exp(i\theta_E)$ leads to the formula

$$\Pr(A|\Gamma_C) = \sum_{E \in A} \Pr(E|\Gamma_F) + \sum_{E,E^* \in A} (\Pr(E|\Gamma_F) \Pr(E^*|\Gamma_F))^{.5} \cos(\theta_E - \theta_{E^*})$$

The key difference between the two formulas is that (9) uses variances to measure interference effects. Since interference effects are alien to standard statistics while variances are commonplace, this change is critical to translating probability amplitude mechanics into the vocabulary of standard statistics.

3.3 Comparing Detailed and Less Detailed Experiments

Operational statistics indicates that all probabilities derived from the same experiment obey the standard rules of probability. Hence $\Pr(A|\Gamma_F) = \sum_{E \in A} \Pr(E|\Gamma_F)$. Making this substitution in (9) gives

$$\Pr(A|\Gamma_C) = |A| \frac{\Pr(A|\Gamma_F) D_F - \sum_Z v_{A,Z}}{D_C} \tag{10}$$

Note that

$$\begin{aligned}
D_F &= \sum_{Z,E\in\Gamma_F} m_{E,Z}^2 = \sum_{Z,A\in\Gamma_C}\sum_{E\in A} m_{E,Z}^2 \\
&= \sum_{Z,A\in\Gamma_C}[\sum_{E\in A} m_{E,Z}^2 - \frac{1}{|A|}m_{A,Z}^2] + \sum_{Z,A\in\Gamma_C}\frac{m_{A,Z}^2}{|A|} \\
&= \sum_{Z,A\in\Gamma_C} v_{A,Z} + D_C\frac{\sum_{Z,A\in\Gamma_C}\frac{m_{A,Z}^2}{|A|}}{\sum_{Z,A\in\Gamma_C} m_{A,Z}^2} = \sum_{Z,A\in\Gamma_C} v_{A,Z} + D_C\sum_{A\in\Gamma_C}\frac{\Pr(A|\Gamma_C)}{|A|}
\end{aligned}$$

so that

$$\Pr(A|\Gamma_C) = \frac{|A|}{D_C}[\Pr(A|\Gamma_F)(\sum_{Z,A} v_{A,Z} + D_C\sum_{A}\frac{\Pr(A|\Gamma_C)}{|A|}) - \sum_{Z} v_{A,Z}]$$

We can rewrite this expression as

$$\Pr(A|\Gamma_F) = (1-\rho)Q(A|\Gamma_C) + \rho V_A \tag{11}$$

with

$$\begin{aligned}
Q(A|\Gamma_C) &= \frac{\Pr(A|\Gamma_C)/|A|}{\sum_A \Pr(A|\Gamma_C)/|A|} \\
V_A &= \frac{\sum_Z v_{A,Z}}{\sum_{A,Z} v_{A,Z}} \\
\rho &= \frac{\sum_{Z,A} v_{A,Z}}{\sum_A \Pr(A|\Gamma_C)/|A| + \sum_{Z,A} v_A^z} = \frac{\sum_{Z,A} v_{A,Z}}{\sum_{Z,E} m_{E,Z}^2}
\end{aligned}$$

Thus the probability of event A as assessed by the detailed operation, Γ_F, is a weighted average of

- the probability of event A as assessed by the less detailed operation, adjusted (via $Q(A|\Gamma_C)$), to increase that probability if A is more detailed (i.e., contains fewer elements of Γ_F) than other sets $A^* \in \Gamma_C$.
- the variance of elemental m-functions within set A divided by that variance summed across all $A \in \Gamma_C$.

To understand ρ, define $M_{A,Z}$ to be a random variable assuming values $m_{E,Z}$ for each $E \in A$ with probability $\frac{1}{|A|}$. The mean of this random variable is $\frac{m_{A,Z}}{|A|}$. The mean squared is $(\frac{m_{A,Z}}{|A|})^2$. The variance of this random variable is $\frac{v_{A,Z}}{|A|} = \frac{\sum_{E\in A} m_{E,Z}^2}{|A|} - (\frac{\sum_{E\in A} m_{E,Z}}{|A|})^2$. Hence the coefficient of variation, $r_{A,Z}$, of this random variable, is

$$r_{A,Z} = \frac{v_{A,Z}}{m_{A,Z}^2/|A|}$$

The average coefficient of variation across all $A \in \Gamma_C$ is

$$r = \frac{\sum_{A,Z} v_{A,Z}}{\sum_{A,Z} m_{A,Z}^2/|A|} = \sum_{A,Z} r_{A,Z} \frac{m_{A,Z}^2/|A|}{\sum_{A,Z} m_{A,Z}^2/|A|}$$

Then

$$\rho = \frac{\sum_{AZ} v_{A,Z}}{\sum_{EZ} m_{E,Z}^2} = \frac{\sum_{AZ} v_{A,Z}}{\sum_{AZ} v_{A,Z} + \sum_{AZ} m_{A,Z}^2/|A|} = \frac{r}{r+1}$$

For many physical applications, $|A|$ can be taken as constant across all elements of Γ_C. In this case, (11) simplifies to

$$\Pr(A|\Gamma_F) = (1-\rho)\Pr(A|\Gamma_C) + \rho V_A \qquad with\ \rho = \frac{r}{r+1}$$

which proves the Theorem[4].

APPENDIX:THE DIRICHLET DENSITY

If random variables M_A^2 and $M_{A^*}^2$ both have generalized Rayleigh (or non-central Chi Square) distributions(Park,1961) with $2\nu_A$ and $2\nu_{A^*}$ degrees of freedom and non-centrality factors, λ_A and λ_{A^*}, then the ratio $p_A = \frac{m_A^2}{m_A^2+m_{A^*}^2}$ has a non-central beta distribution. Because of the analytical complexity of the non-central beta distribution, we focus on the simpler beta distribution (which implies $\lambda_A = \lambda_{A^*} = 0$ with M_A^2 and $M_{A^*}^2$ having Chi Square Distributions.) The multivariate generalization of the beta distribution is the Dirichlet density(Johnson & Kotz,1970):

$$p_f(p_A, p_{A^*}, ..., p_{A^{**}}) \propto \prod_{A\in\Gamma_C} p_A^{\nu_A - 1}(1 - \sum_{A\in\Gamma_C} p_A)^{\omega_0 - 1}$$

with the expected value of p_A given by

$$E[p_A] = \frac{\nu_A}{\sum_{A\in\Gamma_C} \nu_A}$$

and the mode of p_A (for $\theta_A \geq 1$ for $A \in \Gamma_C$) given by

$$M[p_A] = \frac{\nu_A - 1}{\sum_{A\in\Gamma_C}(\nu_A - 1)} = \frac{\nu_A - 1}{\sum_{A\in\Gamma_C} \nu_A - |\Gamma_C|}$$

Then

$$E[p_A] = (1 - \frac{|\Gamma_C|}{\sum_{A^*\in\Gamma_C} \nu_{A^*}})M[p_A] + \frac{|\Gamma_C|}{\sum_{A\in\Gamma_C} \nu_A}\frac{1}{|\Gamma_C|}$$

Substituting $\rho = \frac{|\Gamma_C|}{\sum_{A\in\Gamma_C} \nu_A}$ gives the Lemma. Since $\nu_A \geq 1, A \in \Gamma_C$, we conclude that $0 \leq \rho \leq 1$.

[4]Note that $V_A = \sum_{(i,j)\in A} V_{ij} \frac{2}{A}$.

References

[1] Bohm, D. “A Suggested Interpretation of the Quantum Theory in terms of ‘Hidden Variables’.” *Physical Review.* 85,160-193(1952).

[2] Feynman, R., R. Leighton & M. Sands. *Feynman Lectures on Physics.* Addison-Wesley, New York, 1965.

[3] Foulis, D. & C. Randall. "Operational Statistics, I. Basic Concepts.” *Journal of Mathematical Physics.* 13, 1667-1675(1972).

[4] Foulis, D. & C. Randall. "Operational Statistics, II. Manuals of Operations & Their Logics.” *Journal of Mathematical Physics.* 13, 1667-1675(1972).

[5] Johnson, N. & S. Kotz. *Continuous Univariate Distributions-2.* Houghton Mifflin Company, Boston,1970.

[6] Park, J. “Moments of the Generalized Rayleigh Distribution.” *Quarterly of Applied Mathematics.* 19,45-49(1961).