Consistency, amplitudes and probabilities in quantum theory

2000 J. Phys. A: Math. Gen. 33 5615

(http://iopscience.iop.org/0305-4470/33/31/314)

J. Phys. A: Math. Gen. **33** (2000) 5615–5618. Printed in the UK PII: S0305-4470(00)12797-0

Consistency, amplitudes and probabilities in quantum theory

Yoel Tikochinsky† and Stephen F Gull
Astrophysics Group, Cavendish Laboratory, Madingley Road, Cambridge CB 0HE, UK

Received 21 March 2000, in final form 14 June 2000

Abstract. Caticha (1998 *Phys. Rev.* A **57** 1572) applies consistency arguments to derive the quantum mechanical rules for compounding probability amplitudes in much the same way as in earlier work by Tikochinsky (1988 *J. Math. Phys.* **29** 398, 1988 *Int. J. Theor. Phys.* **27** 543). These works are examined together to find the assumptions needed to obtain the most general results.

In a recent article Caticha [1] uses consistency arguments to derive the quantum mechanical rules for combining probability amplitudes. Caticha's work bears close resemblance, in both approach and execution, to earlier work by one of the present authors [2, 3]. With hindsight, it seems a good time to take stock and see what can be achieved by this approach and what are the assumptions needed to obtain the most general results.

As a preamble, we make the following observations. There exist well defined experimental procedures to prepare a system in certain initial quantum states, as well as ways to test whether a system so prepared arrives at a given final state. One way to prepare a system in a given quantum state is to pass it through a filter (e.g. a photon through a polarizing filter; a particle with spin through a Stern–Gerlach apparatus). Furthermore, there exist pairs of mutually orthogonal filters, such that a system passing through one filter will certainly be blocked by the second filter. For other (non-orthogonal) pairs of filters, it is a matter of chance whether a system prepared by passing through one filter will also pass the second. Moreover, the state of of a system prepared by passing consecutively several filters is determined solely by the last filter. This independence of the past history is referred to as the Markovian property of the quantum state. Thus, following Feynman [4], we treat the quantum state not as a vector in an abstract Hilbert state, but as an experimentally prepared entity.

With the above background in mind, our aim is to derive in a consistent way a substantial part of the formalism of quantum mechanics, namely, the rules for compounding probability amplitudes and for calculating the corresponding probabilities, starting with very simple assumptions. Note that the rules for assigning the elementary amplitudes are not part of this programme. Nevertheless, the part of the formalism that we shall prove has far-reaching consequences, as has been amply demonstrated, particularly by Feynman [4].

Our point of departure is the recognition that in quantum mechanics one cannot directly assign probabilities to processes. In contrast to the classical situation, not every proposition can be answered by yes or no (which slit did the particle go through?). Therefore Boolean algebra does not apply and the road opened by Cox [5] to introduce probabilities is not open to us. Probability must, therefore, be introduced indirectly as a function of the corresponding probability amplitude [3]. Let us review briefly how this is done.

† On leave from Racah Institute of Physics, The Hebrew University of Jerusalem, Israel.

0305-4470/00/315615+04$30.00

The basic entities of concern are the transition amplitudes $\langle B|A\rangle$ between experimentally determined initial and final states A and B. Both time-dependent transitions $\langle B(t_2)|(A(t_1))\rangle$ and transitions at a given time $t_2 = t_1$ are of interest. In fact, time serves as one of the parameters identifying the quantum state. To each transition one assigns a complex number—the probability amplitude for the process. This number is assumed to depend only on the given process $A \to B$ and to be independent of the past history (Markovian property).

Among the possible processes there are two kinds of special interest: *processes in series*, an example of which is $C \to B \to A$ with amplitude $\langle A|C\rangle_{\text{via } B}$, where C is made to pass through a filter B before A is verified, and *processes in parallel*, the simplest of which is

$$B \begin{matrix} \nearrow C_1 \searrow \\ \searrow C_2 \nearrow \end{matrix} A$$

where B can proceed to A only through two orthogonal filters C_1, C_2. The amplitude for this process will be denoted by $\langle A|B\rangle_{\text{via } C_1,C_2}$. Very special cases of these 'in series' or 'in parallel' processes are referred to as AND, OR setups in [1]. Our only assumption regarding these processes is that the amplitudes for the processes are given analytic functions of the partial complex amplitudes x and y, namely,

$$\langle A|C\rangle_{\text{via } B} = f(x, y) \tag{1}$$

where

$$x = \langle A|B\rangle \qquad \text{and} \qquad y = \langle B|C\rangle \tag{2}$$

and

$$\langle A|B\rangle_{\text{via } C_1 C_2} = g(x, y) \tag{3}$$

where

$$x = \langle A|B\rangle_{\text{via } C_1} \qquad \text{and} \qquad y = \langle A|B\rangle_{\text{via } C_2}. \tag{4}$$

Our task is to find the possible form of these functions, subject to consistency demands. Consider now the process $D \to C \to B \to A$ with amplitudes

$$z = \langle C|D\rangle \qquad y = \langle B|C\rangle \qquad \text{and} \qquad x = \langle A|B\rangle. \tag{5}$$

We could calculate the amplitude for the process in two different ways: (a) combine first $C \to B \to A$ to obtain $\langle A|C\rangle_{\text{via } B} = f(x, y)$ and then calculate $f(f(x, y), z)$ or (b) combine first $D \to C \to B$ to obtain $\langle B|D\rangle_{\text{via } C} = f(y, z)$ and then calculate $f(x, f(y, z))$. Consistency then demands that the two calculations give the same result, namely, the function $f(x, y)$ should obey the associative law

$$f(x, f(y, z)) = f(f(x, y), z). \tag{6}$$

Similarly, for processes in parallel, consistency entails

$$g(x, g(y, z)) = g(g(x, y), z). \tag{7}$$

Finally, consider the combined process

$$C \to B \begin{matrix} \nearrow C_1 \searrow \\ \searrow C_2 \nearrow \end{matrix} A.$$

One way to calculate the amplitude is to consider it as a process in series, with amplitude $f(g(x, y), z)$, where

$$x = \langle A|B\rangle_{\text{via } C_1} \qquad y = \langle A|B\rangle_{\text{via } C_2} \qquad \text{and} \qquad z = \langle B|C\rangle. \tag{8}$$

Another way to look at the same process is to consider it as a process in parallel, with an amplitude $g(f(x,z), f(y,z))$. Demanding that the two representations agree, we have the distributive law

$$f(g(x,y),z) = g(f(x,z), f(y,z)). \tag{9}$$

This is all that is needed to determine the functions $g(x,y)$ and $f(x,y)$. From here on the rest is mathematics. In particular, there is no need to assume commutativity for processes in parallel, as was done in [2]. The equality $g(x,y) = g(y,x)$ follows automatically from Cox's solution [6] of the associative law (7), as recounted, for example, in [1]. As shown in [2] and [1], given the functions $g(x,y)$ and $f(x,y)$ it is always possible to find a transformation $x' = H(x)$ which will bring g and f to the canonical form

$$[g(x,y)]' = x' + y' \qquad [f(x,y)]' = x'y'. \tag{10}$$

Conversely, starting with equation (10), one can make a transformation $x'' = K(x')$ such that, in terms of the new variables x'' and y'', the addition and multiplication laws (10) change their form, without changing their contents. There exist, beside equation (10), other solutions such as $g(x,y) = \phi(x)$, $g(x,y) = \psi(y)$ (with some restrictions on the functions $\phi(x)$ and $\psi(y)$) or even $g(x,y) =$ constant. These solutions are discarded as being non-generic.

From here on we shall restrict our discussion to transitions at a given time. Our aim is to amend the general proof of Born's law

$$\Pr(A|B) = |x|^2 \qquad x = \langle A|B\rangle \tag{11}$$

given in [3]. Here $\Pr(A|B)$ stands for the probability of transition, at a given time, from B to A. As shown in [3], assuming that the amplitude $\overline{x} = \langle B|A\rangle$ for the inverse transition $A \to B$ is a function of the amplitude $x = \langle A|B\rangle$, we obtain $\langle B|A\rangle = \langle A|B\rangle^*$. Futhermore, the probability for the process $B \to A$, assuming it to be a function of the amplitude $x = \langle A|B\rangle$, was shown there to be of the form

$$\Pr(A|B) = |x|^\alpha \qquad \alpha > 0. \tag{12}$$

Consider now all the orthogonal states A_i, which can be reached from B, with an amplitude $x_i = \langle A_i|B\rangle$. Since

$$\langle B|B\rangle = \sum_i \langle B|A_i\rangle\langle A_i|B\rangle = \sum_i x_i^* x_i \tag{13}$$

and since the probability of the certain event satisfies $\Pr(B|B) = 1$, we have by (12) and (13)

$$\Pr(B|B) = \left(\sum_i |x_i|^2\right)^\alpha = 1. \tag{14}$$

Hence, taking the logarithm of both sides, we obtain

$$\sum |x_i|^2 = 1. \tag{15}$$

However the totality of processes $B \to A_i$ forms an exhaustive and mutually exclusive set of alternatives, satisfying (see equation (12))

$$\sum \Pr(A_i|B) = \sum |x_i|^\alpha = 1. \tag{16}$$

Comparing (15) and (16) we find $\alpha = 2$ and

$$\Pr(A|B) = |x|^2. \tag{17}$$

(In [3] time dependence crept in leaving $\Pr(B(t_2)|B(t_1)) \neq 1$, thus disabling the preceding argument.)

In summary, the assumptions (i) that amplitudes for processes in series or in parallel are represented by analytic functions of the complex partial amplitudes and (ii) that the amplitude for the inverse process and the probability of the process are functions of the amplitude for the process are enough to derive the known quantum mechanical rules for combining amplitudes and for calculating the corresponding probabilities. This is achieved using general states and filters. That these assumptions are all that is needed was not fully realized in [2, 3]. The work of Caticha certainly helped to put things in sharper focus. In particular, as shown by Caticha, assumption (i) is enough to establish the linearity of the Schrödinger equation.

Acknowledgments

One of us (YT) would like to thank Geoff Stedman for drawing his attention to [1], and for reading and commenting on an earlier draft of this work. The hospitality of the University of Canterbury, New Zealand, is warmly acknowledged.

References

[1] Caticha A 1998 *Phys. Rev.* A **57** 1572
[2] Tikochinsky Y 1988 *J. Math. Phys.* **29** 398
[3] Tikochinsky Y 1988 *Int. J. Theor. Phys.* **27** 543
[4] Feynman R P, Leighton R B and Sands M 1965 *The Feynman Lectures on Physics* vol 3 (Reading, MA: Addison-Wesley)
Feynman R P 1985 *QED, Alix Mounter Memorial Lectures* (Princeton, NJ: Princeton University Press)
[5] Cox R T 1946 *Am. J. Phys.* **14** 1
[6] Cox R T 1961 *The Algebra of Probable Inference* (Baltimore, MD: John Hopkins University Press)