WHY QUANTUM THEORY?

LUCIEN HARDY
Centre for Quantum Computation,
The Clarendon Laboratory,
Parks road, Oxford OX1 3PU, UK
hardy@qubit.org

Abstract. The usual formulation of quantum theory is rather abstract. In recent work I have shown that we can, nevertheless, obtain quantum theory from five reasonable axioms. Four of these axioms are obviously consistent with both classical probability theory and quantum theory. The remaining axiom requires that there exists a continuous reversible transformation between any two pure states. The requirement of continuity rules out classical probability theory. In this paper I will summarize the main points of this new approach. I will leave out the details of the proof that these axioms are equivalent to the usual formulation of quantum theory (for these see reference [1]).

1. Introduction

The usual formulation of quantum theory is very obscure employing complex Hilbert spaces, Hermitean operators and so on. While many of us, as professional quantum theorists, have become very familiar with the theory, we should not mistake this familiarity for a sense that the formulation is physically reasonable. Quantum theory, when stripped of all its incidental structure, is simply a new type of probability theory. Its predecessor, classical probability theory, is very intuitive. It can be developed almost by pure thought alone employing only some very basic intuitions about the nature of the physical world. This prompts the question of whether quantum theory could have been developed in a similar way. Put another way, could a nineteenth century physicist have developed quantum theory without any particular reference to experimental data? In a recent paper I have shown that the basic structure of quantum theory for finite and countably infinite dimensional Hilbert spaces follows from a set of five reasonable axioms [1]. Four of these axioms are obviously consistent with both classical probability theory and with quantum theory. The remaining axiom states that there exists a continuous reversible transformation between any two pure states. This axiom rules out classical probability theory and gives us quantum theory. The key word in this axiom is the word "continuous". If it is dropped then we get classical probability theory instead. The proof that quantum theory follows from these axioms, although involving simple mathematics, is rather lengthy. In this paper I will simply discuss the main ideas referring interested readers to the main paper [1].

T. Placek and J. Butterfield (eds.), Non-locality and Modality, 61–73.

Various authors have set up axiomatic formulations of quantum theory, for example see references [2, 3, 4, 5, 6, 7, 8, 9, 10, 11] (see also [12, 13, 14]). Much of this work is in the quantum logic tradition. The advantage of the present work is that there are a small number of simple axioms which can be easily motivated without any particular appeal to experiment, and, furthermore, the mathematical methods required to obtain quantum theory from these axioms are very straightforward (essentially just linear algebra).

2. Basic notions

We will consider situations in which a preparation apparatus prepares systems which may be transformed by a transformation apparatus and measured by a measurement apparatus. Associated with any given preparation will be a *state*. *The state is defined to be (that thing described by) any mathematical object that can be used to determine the probability associated with each outcome of any measurement that may be performed on a system prepared by the associated preparation.* The point is that, if one knows the state, one can predict probabilities for any measurement that may be performed. It is not entirely clear that one will be able to ascribe states to preparations. The first axiom, to be introduced later, will make this possible by assuming that the same probability is obtained under the same circumstances. If we can ascribe a state it is clear from the definition above that one way of describing the state is by that mathematical object which simply lists all the probabilities for every outcome of every conceivable measurement that could possibly be made on the system. This would be a very long list. Since most physical theories have some structure, it is likely that this would be too much information. We can imagine that a set of K appropriately chosen probability measurements will be just sufficient and necessary to determine the state (so K is the smallest number of probabilities required to specify the state). We will call these the *fiducial measurements*. We can list just the probabilities corresponding to these fiducial measurements in the form of a column vector. Thus, the state can be written

$$\mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_K \end{pmatrix}. \tag{1}$$

We will call the integer K the *number of degrees of freedom*. This number plays an important role in this work.

The allowed states $\mathbf{p}$ will belong to some set S. We expect that there will exist sets of states which can be distinguished from each other in this set by a single shot measurement. Consider one such set. If Alice picks a state from this set and sends it to Bob then Bob can set up a measurement apparatus such that each state gives rise to a disjoint set of outcomes. By knowing which outcomes are associated with which state, Bob can tell Alice which state she sent. Let the maximum number of states in any such set be called N. We will call N the *dimension* (because in quantum theory it corresponds to the dimension of the Hilbert space).

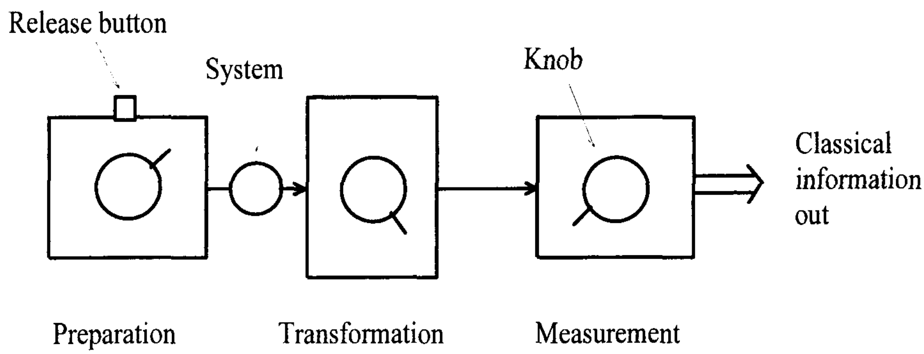

Figure 1. The situation considered consists of a preparation device with a knob for varying the state of the system produced and a release button for releasing the system, a transformation device for transforming the state (and a knob to vary this transformation), and a measuring apparatus for measuring the state (with a knob to vary what is measured) which outputs classical information.

Associated with any particular type of system will be the two integers K and N. It turns out that in classical probability theory we have $K = N$ and in quantum theory we have $K = N^2$. We will explain why this is the case later.

First, let us describe the type of scenario we wish to consider. This is shown in Figure 1. We have three types of apparatus. The preparation apparatus prepares systems in some state. It has a knob on it for varying the type of state prepared. It also has a release button, whose role will be described shortly. The system then passes through a transformation apparatus. This has a knob on it which varies the transformation effected. Unless otherwise stated, we will assume that the transformation device is set to leave the state unchanged (i.e. effect the identity transformation). Finally the system impinges onto a measurement apparatus. This has a knob on it to vary the measurement being performed. It also has some classical information coming out. Either we obtain a non-null outcome, labeled $l = 1$ to L, or we obtain a *null outcome*. We require that if the release button is pressed on the preparation apparatus (and assuming that the transformation is set to the identity) then we will certainly obtain a non-null outcome. On the other hand, if the release button is not pressed then we will certainly obtain a null outcome. To illustrate this we could think of an array of detectors labeled $l = 1$ to L. If none of the detectors click then we can say this is a null result. Since we allow null outcomes we need not assume that states are normalized.

All quantities are reducible to measurements of probability. For example, any measurement of an expectation value is really a probability weighted sum. Therefore, we need only consider measurements of probability. Henceforth, when we refer to a "measurement" or a "probability measurement" we mean specifically a measurement of the probability that the outcome belongs to some non-null subset of outcomes with a given knob setting on the measurement apparatus.

If we never press the release button then all the fiducial probability measurements will be equal to zero (so the state will be represented by a column vector with K zero's). We will call this state the null state.

It is normal in probability theory to talk about pure states and mixed states. A mixed state is any state which can be simulated by a mixture of two distinct states. Thus, we prepare randomly either state A or state B with probabilities λ and $1-\lambda$ where $0<\lambda<1$. Pure states are defined to be those states (except the null state) which are not mixed states. Pure states will turn out to be extremal states in the set of allowed states (this set being convex).

We will now describe classical probability theory and then quantum theory. We will find that it is possible to give the two theories a very similar mathematical structure. This will help us to appreciate the similarities and differences between the two theories.

3. Classical probability theory

Consider a ball that can be in one of N boxes (or be missing). The state is fully determined by specifying the probabilities, p_n, for finding the ball in each box. This information can, as in the previous section, be written

$$\mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_N \end{pmatrix}. \tag{2}$$

Since the ball may be missing, the sum of the probabilities in this vector must be less than or equal to one. There are N entries in $\mathbf{p}$. Hence, $K=N$. There are some interesting special cases. The states

$$\mathbf{p}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \mathbf{p}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \mathbf{p}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \text{etc.} \tag{3}$$

represent the case where the ball is definitely in one of the boxes. These states cannot be simulated by mixtures of other states and hence are pure states for this system. The state

$$\mathbf{p}_{\text{null}} = \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{4}$$

represents the case where the ball is missing. These $N+1$ states are extremal in the space of allowed states. Since we are casting classical probability theory and quantum theory in similar mathematical forms, let us consider how we can represent measurements in the classical case. One measurement we could make is to look and see if the ball is in box 1. The probability of finding the ball in box 1 is p_1. We can write this as

$$p_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \cdot \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_N \end{pmatrix} = \mathbf{r}_1 \cdot \mathbf{p}. \tag{5}$$

Hence, we can identify the vector $\mathbf{r}_1$, defined as

$$\mathbf{r}_l = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{6}$$

with the measurement where we look to see if the ball is in box 1. We can write down similar vectors for the other boxes. However, we could perform more complicated measurements. For example, we could toss a λ biased coin and look in box 1 if it came up heads and in box 2 if it came up tails. In this case the measurement being performed would be represented by the vector

$$\mathbf{r} = \lambda \mathbf{r}_1 + (1-\lambda)\mathbf{r}_2 \tag{7}$$

since then $\mathbf{r} \cdot \mathbf{p} = \lambda p_1 + (1-\lambda) p_2$. In general it can be shown that the probability associated with any measurement is given by

$$\text{prob}_{\text{meas}} = \mathbf{r} \cdot \mathbf{p} \tag{8}$$

where $\mathbf{r}$ is associated with the measurement and $\mathbf{p}$ is associated with the state.

Consider a classical bit. This is a system with $N = 2$. In this case the extremal states are

$$\mathbf{p}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \mathbf{p}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \mathbf{p}_{\text{null}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{9}$$

The set of allowed states $S_{\text{classical}}$ are given by the convex hull of these extremal states as shown in Figure 2a. Note that the normalized states (for which $p_1 + p_2 = 1$) lie on the hypotenuse. Note also that the pure states form a discrete set. There is no continuous path from one pure state to another which goes through the pure states.

We see that classical probability theory is characterized by $K = N$, by the set $S_{\text{classical}}$ of allowed states $\mathbf{p}$ and the set $R_{\text{classical}}$ of allowed measurements $\mathbf{r}$, and by the formula $\text{prob}_{\text{meas}} = \mathbf{r} \cdot \mathbf{p}$.

http://www.springer.com/978-1-4020-0662-3

Non-locality and Modality
Placek, T.; Butterfield, J. (Eds.)
2002, X, 352 p., Softcover
ISBN: 978-1-4020-0662-3