J. L. MARTIN

Basic quantum mechanics

J. L. MARTIN
KINGS COLLEGE, LONDON

Basic quantum mechanics

Clarendon Press · Oxford · 1981

Oxford University Press, Walton Street, Oxford OX2 6DP

OXFORD LONDON GLASGOW
NEW YORK TORONTO MELBOURNE WELLINGTON
KUALA LUMPUR SINGAPORE JAKARTA HONG KONG TOKYO
DELHI BOMBAY CALCUTTA MADRAS KARACHI
NAIROBI DAR ES SALAAM CAPE TOWN

Published in the United States by
Oxford University Press, New York.

British Library Cataloguing in Publication Data

Martin, John Legat
Basic quantum mechanics.—(Oxford physics series).
1. Quantum theory
I. Title
530.1'2 QC174.12 80-41845

ISBN 0-19-851815-3
ISBN 0-19-851816-1 Pbk

Typeset by Eta Services, Beccles, Suffolk.
Printed in Great Britain by Billing & Sons Ltd.
Guildford London Oxford Worcester

Preface

I cannot believe that God plays dice with the cosmos.

Albert Einstein.

We throw the dice. It is up to the Lord to fix how they are to fall.

Proverbs 16:33.

THE basic rules of quantum mechanics are not at all complicated: the main problem is the hurdle of unfamiliarity. We have no right to expect that 'everyday' experience will be of any help in those regions of physics which cannot be handled in an everyday way. This is easy to accept. Yet it has taken centuries to learn this lesson, and it has only been relatively recently that physicists have been forced to accept that 'everyday' mechanics must be modified to deal with phenomena on the cosmological or on the atomic scale. In the main, quantum mechanics is concerned with the atomic scale. Moreover, quantum mechanics is *statistical* in nature: it deals with questions of probability. This is in itself a psychological barrier for many people: even Albert Einstein was inflexibly opposed to a theoretical scheme which leaves so much to chance.

The basic rules are laid down in the first three chapters and repeatedly illustrated by reference to one and only one example: the states of polarization of a single photon. At this stage the mathematics is kept as simple as it can be, since it is important in a subject where conceptual difficulties may be severe not to make matters worse!

Understanding the first three chapters is not the same as being able to *use* quantum mechanics. The rules may be simple, but the range of applications is very rich, and the solutions of some problems lie well beyond the reach of present mathematical expertise. In order to be a competent quantum 'mechanic', one must become familiar with techniques appropriate to more general systems, and understand why classical ideas like energy and momentum reappear in quantum mechanics. Such topics, along with their applications in particular to electron spin, the hydrogen atom, and the harmonic oscillator, take us up to chapter 11.

Very few realistic quantum problems are exactly solvable, and the next six chapters deal with the more frequently used approximation methods. Throughout I have tried to offer realistic applications of each method, to show that they all have a genuine value. The book ends with a brief sketch of a few of the many difficulties and problems which have plagued the quantum philosophers from the very beginning.

Though the text originates from two undergraduate courses given at King's College, it is not a mere expansion of the lecture notes; most topics

are presented in much greater depth. The emphasis throughout is on a careful and precise statement of quantum principles, with the mathematics no more complicated than is required for this aim. I have not followed the historical development of the subject, fascinating though it is, as it covers 25 years of groping and false starts. I feel that the best time to meet, for example, Bohr's 1913 model of the hydrogen atom is *after* meeting Schrodinger's 1925 model. In this way, one may appreciate much more fully both the brilliance and the shortcomings of Bohr's approach.

There are problems scattered through the text to encourage the reader to test his grasp of the subject reasonably frequently. Sometimes the result of a problem is used at a later point in the text; this should not cause any difficulty. When there is a need for techniques which can be found in any of the dozens of books on mathematical methods, they are not reproduced; instead they are summarized as lists of mathematical prerequisites.

It is obvious that the author of a book such as this must owe a great deal to others. In particular I wish to acknowledge both the helpful discussions with several colleagues, especially Professor E. J. Burge, and the very many texts on quantum mechanics which already exist: these are so numerous that it would be impracticable to list them all, and invidious to list only some. I am grateful to the Editors of *Physics Today* for permission to include the quotations on pp. 103 and 122. Several of the diagrams, mainly in chapters 16 to 18, were generated by FORTRAN programs invoking the very useful DIMFILM routines available at the University of London Computer Centre.

King's College, London.
April 1981 *J.L.M.*

Contents

1. Basics

Why do we need quantum mechanics?

IT is now about three hundred years since Newton laid the foundations of his mechanics, and rather more than a hundred since Maxwell formulated the equations of electromagnetic theory. Up till about the end of the nineteenth century it was generally believed that Newtonian mechanics and Maxwell's equations were the complete and final expression of the basic laws of physics, and that future progress lay merely in minor refinements and more exact solutions. This belief was not unreasonable, since the success of Newtonian mechanics had been so spectacular till then. It had even been said by some that there was nothing left to be discovered by the next generation! However, there were misgivings; in connection with an apparently insurmountable difficulty in black-body radiation, Lord Kelvin referred in 1904 to 'a cloud which has obscured the brilliance of the molecular theory of heat and light during the last quarter of the nineteenth century'; and indeed it was becoming clear that certain awkward experimental results would compel a major theoretical change. It was not till 1925 that the final shape of the 'new' dynamics was found; since then the basic ideas have not changed.

From our present point of view, perhaps the most important feature of Newtonian mechanics is that it is *deterministic*. If we know every detail of the state of a physical system at a particular time, then we are able in principle to predict the precise result of an observation made at a later time. However, to know whether the physical world is indeed deterministic we must make an appeal to experiment, and experiment at the atomic level apparently suggests that the physical world is *not* deterministic (at least, not in the sense just described). Thus Newtonian mechanics needs to be modified.

Why then has Newtonian mechanics been so successful? An important feature of quantum mechanics is that certain fundamental physical quantities (previously believed to be continuous) are observed to take only a discrete set of values. For example, angular momentum will take a value which is an integer multiple of a basic unit, $\frac{1}{2}h = 0{\cdot}527 \times 10^{-34}$ joule-seconds (J s), and no other value. This 'natural' unit is almost unimaginably small compared with 'everyday' angular momenta, and it is not surprising that it was overlooked. Generally speaking, it is correct to use Newtonian mechanics for systems big enough or massive enough for quantum effects to be negligible, e.g. a railway train, or even a particle large enough to be visible in an optical microscope.

Will quantum mechanics survive, or will it be modified as Newtonian mechanics has been? No-one knows, but the history of science suggests that we would do well not to accept any theory as necessarily final, however beautiful or however reasonable it may seem.

The photoelectric effect

A good example of a phenomenon inexplicable by classical theories is the *photoelectric effect*. In 1888, Hallwachs found that a negatively charged metal plate could be discharged by ultraviolet light, and it was soon understood that this results from the ejection of electrons from the surface of the metal. Ten years later, Lenard investigated the phenomenon, obtaining results that were incomprehensible to the theory of the time. He showed that for monochromatic light of frequency ν the kinetic energy of an ejected electron could not exceed a maximum which depended linearly on ν:

$$\text{K.E.} \leqslant h\nu - E_0;$$

h was apparently a universal constant† ($\sim 6{\cdot}6 \times 10^{-34}$ J s), while E_0 depended on the metal being irradiated. Classically, one would have expected that the maximum kinetic energy would increase with the incident intensity, but this did not happen; perhaps more surprisingly, for radiation whose frequency was too low ($h\nu < E_0$) no electrons at all were ejected, *however intense* the radiation.

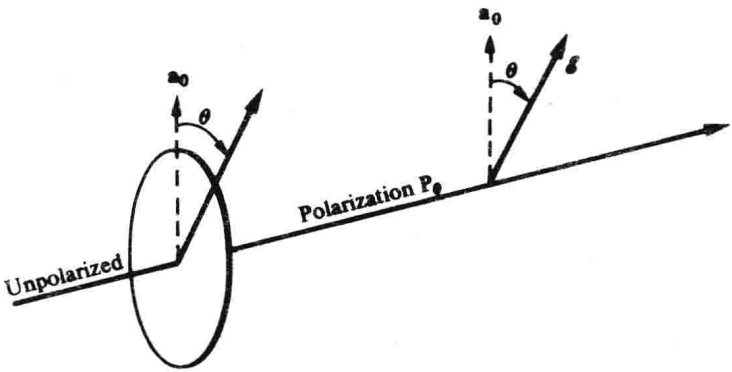
In 1905 Einstein gave a quantum explanation of this phenomenon which still stands today. The monochromatic beam is in this situation to be regarded as a beam of *photons* each carrying the same characteristic amount of energy $h\nu$. When a photon in the beam interacts with an electron in the metal, the electron may be ejected with kinetic energy not greater than $h\nu - E_0$, where E_0 is the minimum energy required to overcome the potential 'cliff' at the surface. The photon description fits the observed effects exactly. Any classical attempt at a description fails since it cannot allow for the discrete nature of the energy in a monochromatic beam.

Plane-polarized light

A useful illustration which we shall use repeatedly is provided by the properties of polarized light. Imagine a uniform monochromatic collimated beam of light. Such a beam has definite frequency, direction, and intensity, but is not quite completely specified; there remains the possibility of polarization. We shall consider plane-polarization to begin with.

A beam is *plane-polarized* if the electric vector $\mathscr{E}$ in the beam is everywhere parallel to some vector $\boldsymbol{a}$, which must itself be perpendicular to the direction of propagation. The direction of $\boldsymbol{a}$ is conveniently specified by the angle θ between $\boldsymbol{a}$ and a fixed vector $\boldsymbol{a}_0$; we shall call the corresponding polarization P_θ. There is one such polarization P_θ for each θ in the range $0 \leqslant \theta < \pi$; note that P_θ and $\mathrm{P}_{\theta+\pi}$ are the same polarization, since the *sense* of the vector $\boldsymbol{a}$ is not relevant.

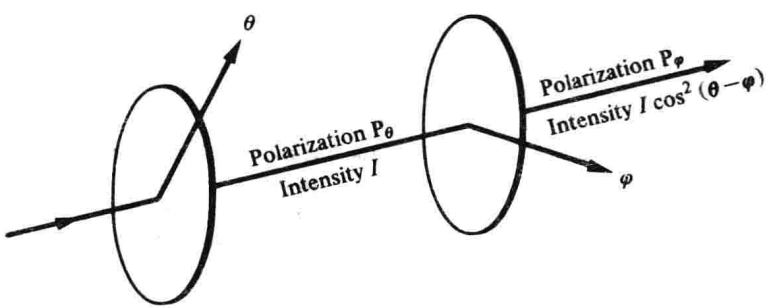
† **Planck's constant.** Introduced by Planck (1900) for reasons from *statistical* mechanics, this constant appears throughout the entire range of microscopic physics.



One way of obtaining a beam polarized at the angle θ is to pass it through an appropriately oriented plane-polarizer (such as a Nicol prism or a Polaroid filter). By an obvious convention we say that a plane-polarizer which produces the polarization P_θ is oriented at the angle θ.

Plane-polarized photons

Phenomena like the photoelectric effect lead us to regard a beam of P_θ-polarized light as a beam of photons, and we then say that each photon in the beam is 'in the **state** P_θ' (see below, p. 6).† Quantum mechanics is the study of the relations between such states.



Consider the following known experimental facts: a monochromatic beam with polarization P_θ encounters a polarizer set at angle ϕ; then

(1) the beam exits with polarization P_ϕ (as we already know);

† In the text bold type will always indicate the point at which a *definition* of an essential concept is given. This point need not be where the concept is first mentioned

(2) the beam is still monochromatic, with no change in frequency;
(3) the intensity of the beam changes by a factor $\cos^2(\theta-\phi)$.

Fact (2) implies that the beam consists of photons of energy $h\nu$ both before and after meeting the polarizer; then fact (3) can only imply that the *proportion* of photons surviving the polarizer is $\cos^2(\theta-\phi)$. This in turn suggests that we should say for a single photon

$$\left.\begin{array}{l}\text{the } \textbf{probability} \text{ that a photon in state } P_\theta \\ \text{survives a polarizer with orientation } \phi \\ \text{(thus then being in state } P_\phi)\end{array}\right\} = \cos^2(\theta-\phi)$$

Here, as always in quantum mechanics, the word *probability* really means *proportion*, where we have a large number of identical situations in mind, e.g. a large number of photons striking a polarizer. It should now be clear that it is not easy to avoid the idea of indeterminacy when we wish to describe a beam of light as a stream of photons; it is usually not possible to say with certainty what will happen to any particular photon in the beam.

A physical theory must relate and predict the results of experiments. Thus quantum mechanics must be able to deal with probabilities. We are given a clue to the appropriate formalism by noting that

$$\cos(\theta-\phi) = \cos\theta\cos\phi+\sin\theta\sin\phi, \tag{1.1}$$

suggesting that we should associate number pairs ('vectors') with the polarization states

$$P_\theta \leftarrow (\cos\theta, \sin\theta) \qquad P_\phi \leftarrow (\cos\phi, \sin\phi).$$

('P_θ is *represented* by $(\cos\theta, \sin\theta)$.') The probability may now be written in terms of the 'scalar product' (eqn (1.1)) of these two vectors. This is not an accident: in every case where quantum mechanics has been used the link, indeed the only link, between theory and experiment is through scalar product expressions of this kind.

Quantum mechanics therefore uses the techniques of vector algebra, along with extensions appropriate to more complicated situations. Accounts of these techniques are to be found in other books, and in this book we shall merely list some 'mathematical prerequisites' from time to time.

Mathematical prerequisites

An n-component **column vector** is an ordered set of n complex **components**, conventionally arranged as a vertical array

$$c = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}.$$

The **Hermitian conjugate** of a column vector is a **row vector**, whose components are the complex conjugates c_i^* of those of the column vector, arranged conventionally as a horizontal array:

$$c^+ = (c_1^*, c_2^*, \ldots, c_n^*)$$

(often called 'c-dagger').

The **sum** of two column vectors is obtained by adding corresponding components:

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix}.$$

The sum of two row vectors is similarly defined. We may multiply a vector by a (possibly complex) number, according to

$$\lambda \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} \lambda c_1 \\ \lambda c_2 \\ \vdots \\ \lambda c_n \end{pmatrix}.$$

The **zero vector** (either row or column) has all components zero and is usually written 0.

The **scalar product** of a row vector w^+ and a column vector v is

$$w^+ v \equiv (w_1^*, w_2^*, \ldots, w_n^*) \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \sum_{i=1}^{n} w_i^* v_i.$$

If $w^+ v = 0$, the column vectors v and w (or equivalently the row vectors v^+ and w^+) are **orthogonal**.

The value of $v^+ v$ for any vector v is always positive, since

$$v^+ v = \sum_{i=1}^{n} |v_i|^2,$$

and thus is a sum of positive terms. The **length** of the column vector v is the positive root $(v^+ v)^{\frac{1}{2}}$. A vector is **normalized** if its length is 1.

The nature of a quantum state

We are now ready to look at the basic structure of quantum mechanics. A difficulty which we meet at the very beginning is that the ideas are unfamiliar. We are used to Newtonian mechanics where everything is *predictable*. For example, if we have to do with the motion of a particle under the influence of a given force, and if we are given the position and the velocity of that particle at time $t = 0$, it is 'simply' a matter of solving the Newtonian equations of motion to obtain the precise position (or the velocity) at any future time—or, for that matter, any past time: it works both ways. ('Simply' in quotes, because

there may be practical numerical difficulties in solving the equations; such difficulties are not really relevant here.)

Our discussion of photon polarization, on the other hand, suggests that we are wrong to demand complete predictability in physics. To understand what a radical change in thinking is required, consider some 'quantum' system at time $t = 0$, and imagine that we have at that time subjected it to an observation O_1 so comprehensive that we have squeezed out every last drop of information about its state that we can. It is then natural to say that it is then in a *definite* **state** S which is specified by the information that we have extracted. We have not, as yet, introduced anything new; after all, if we have measured the position and velocity of a Newtonian particle (which we understand) we are equally entitled to say that it is then in a definite state.

The major change is that even though we declare that a *quantum* system is in a definite state S, a different observation O_2 does not necessarily yield a definite *result*. Sometimes we may get the result a, say, and sometimes b. Worse still, this same observation, when performed for a *different* state S', may sometimes give the result a and sometimes b. In other words, knowing the result of an observation *cannot* tell us the state that the system was in when the observation was made; and knowing the state of a system *cannot* predict with certainty what the result of any observation will be.

In such a situation it is difficult to see how one can have a theory describing the mechanics of a *single* quantum system; in fact, no-one has ever succeeded in producing a satisfactory one. The claims of quantum mechanics are much more modest, however, and can be summarized as follows.

1. A thoroughly comprehensive observation O_1 of a system will put that system in a definite quantum state S; the particular state S which results is defined by the result of the observation. (Thus in one sense, the state S is a kind of *codification* of the result of the observation.)
2. A further observation O_2—possibly a different one and not necessarily comprehensive—made on a quantum system in a definite state S may produce one of a set of possible results $a_1, a_2, \ldots$. If there is more than one possible result, *no prediction is made* about which result will occur.
3. Now imagine that we use the comprehensive observation O_1 to put the system *repeatedly* into a definite quantum state; on those occasions when the quantum state is S (as checked by the result of O_1), let us then go on to make the further observation O_2. Quantum mechanics predicts the *proportion* p_1 *of occasions* for which the result is a_1, the *proportion* p_2 *of occasions* for which the result is a_2, and so on.

We may therefore say that quantum mechanics is concerned with the *statistical* outcome of a fixed experimental procedure repeated many times; its aim is to predict the values of the *probabilities* $p_1, p_2, \ldots$. It achieves this aim with the aid of an appropriate mathematical structure, which we must now consider.

The representation of states by state vectors

Quantum systems may be quite simple, or they may be extremely complicated; therefore the same will be true of the mathematical structures we set up to represent them. To begin with we consider the very simplest systems; certain generalizations will become necessary for the more complicated systems of Chapter 4.

For the simplest systems, quantum mechanics asserts that each possible physical state is to be **represented** by a normalized column vector (the **state vector**) or (equivalently) by the Hermitian conjugate row vector

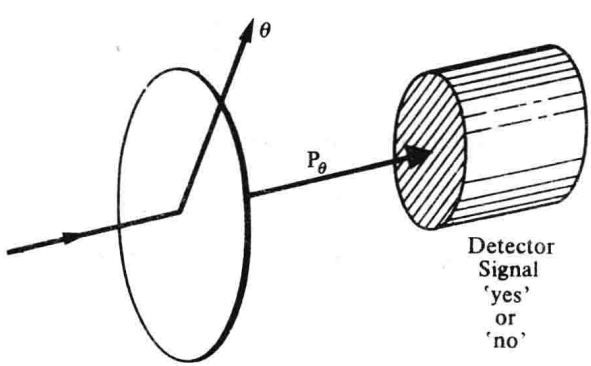
$$\text{physical state} \leftarrow \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \quad \text{or} \quad (c_1^*, c_2^*, \ldots, c_n^*) \qquad \text{(Rule I)}$$

with $\sum |c_k|^2 = 1$. The components c_k are to be allowed to be complex. How many components are needed depends on how complicated the system is; we have already hinted that photon polarization requires two (p. 4). Indeed, some important systems cannot be represented by state vectors with a finite number of components; they will not be considered until Chapter 4.

To complete the scheme of quantum mechanics we shall need some further rules to link the values of the components with the results of observation (see below), and others to provide a *dynamics*, i.e. equations of motion governing the change in the components as time passes (see Chapter 3).

The physical interpretation of the formalism

Let us begin with an illustration. Consider the simple *analyser* consisting of a plane-polarizer in front of a photon detector; the detector gives a signal ('yes') whenever a photon passes through the polarizer and meets the detector. An



observation consists in aiming a photon at the analyser and sensing a signal ('yes') or the absence of a signal ('no'). When the analyser is oriented at angle θ, any photon in polarization state P_θ will certainly reach the detector and yield a 'yes'; as it happens, this is true for no other polarization state whatever. (Other states, of course, *may* give a 'yes', but not with certainty; the state P_ϕ will give 'yes' with probability $\cos^2(\theta-\phi)$.)

It should be clear that a measurement of this kind cannot determine the state of the incident photon. In fact, it is generally true in quantum mechanics that we never 'measure' states, but physical quantities; knowledge of the state implies knowledge of the probabilities of all the various possible outcomes, according to the rules now to be outlined.

To help us in this we introduce the idea of a **'simple' observation**, whose two possible outcomes are 'yes' and 'no', and for which there is *exactly one* state (say S) yielding 'yes' with certainty; let us call this observation O(S), labelling it with the state S in a natural way. Observations of this kind are not as artificial as they may seem; the photon detector just described provides the example of the 'simple' observation $O(P_\theta)$. In any case, more complicated observations are easily expressed in terms of 'simple' observations (Chapter 2), so that the idea is in no way restrictive. By the way, the 'comprehensive' observations mentioned earlier (p. 6) are 'simple' observations in this sense; they may *not* be simple in any practical experimental sense!

The fundamental connection between theory and observation can now be given. Imagine that a system in a state represented by the state vector w is subjected to the observation O(S), where the state S is represented by the state vector v. Then

$$\left.\begin{array}{l}\text{the probability that the outcome}\\ \text{of the observation is 'yes'}\end{array}\right\} \text{ is } |w^+v|^2. \qquad \text{(Rule II)}$$

It is important to understand that this rule provides the *only* connection between the formalism and the physics. It expresses the physical relation between states in terms of a scalar product, and it is for this reason that quantum mechanics uses the methods of linear algebra—the natural mathematical apparatus for dealing with scalar products.

Let us see how Rules I and II work out for a plane-polarized photon. We have already hinted (p. 4) that the state P_θ may be represented by

$$P_\theta \leftarrow \begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} \quad \text{or} \quad (\cos\theta, \sin\theta).$$

Let us perform the 'simple' observation $O(P_\phi)$ on a photon in the state P_θ (or, in everyday language, allow a plane-polarized photon, state P_θ, to fall on the analyser described above, oriented at angle ϕ). Applying Rule II in this case leads us to write down

$$w^+ = (\cos\phi, \sin\phi) \quad \text{and} \quad v = \begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix},$$

and then to write

$$\left.\begin{matrix}\text{the probability of getting}\\ \text{a signal from the detector}\end{matrix}\right\} \text{ is } \left|(\cos\phi, \sin\phi)\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}\right|^2$$

$$= |\cos\phi\cos\theta + \sin\phi\sin\theta|^2$$

$$= \cos^2(\theta-\phi).$$

So the rules work for a plane-polarized photon, using a state vector with two components.

PROBLEMS

1. Use Rule II to confirm that making the observation O(S) on the state S gives 'yes' with certainty. (Recall that state vectors are required to be normalized.)

2. u and v are any two fixed normalized vectors, and α is a complex number.
   (a) Prove that $(u^+ + \alpha^* v^+)(u+\alpha v) \geqslant 0$ for any α. (The left side is a sum of squared moduli.)
   (b) Prove that as α varies, the minimum possible value of $(u^+ + \alpha^* v^+)(u+\alpha v)$ is $1-|u^+v|^2$, attained when $\alpha = -v^+u$. (One route to this result is to remark that the partial derivatives with respect to the real and imaginary parts of α must both vanish at the minimum. There is a short cut: differentiate with respect to α^*, keeping α 'constant'.)
   (c) Deduce that $0 \leqslant |u^+v|^2 \leqslant 1$, and thus confirm that the scalar product expression in Rule II may safely be interpreted as a probability.

3. Suppose that the state S_k is represented by the vector

$$S_k \leftarrow \begin{pmatrix}0\\0\\\vdots\\0\\1\\0\\\vdots\\0\end{pmatrix},$$

with all components zero, except $c_k = 1$. For a *general* state

$$S \leftarrow \begin{pmatrix}c_1\\c_2\\\vdots\\c_k\\\vdots\\c_n\end{pmatrix},$$

show that the probability that the observation $O(S_k)$ yields a 'yes' is $|c_k|^2$. (The components c_k are for this reason often called **probability amplitudes** in the literature. However, 'amplitude' is a much overworked word, and we shall not use it.)