

Quantum Information

Dr C. H. W. Barnes

Cavendish Laboratory, Cambridge University

Contents

1	Preface	<i>page</i> 2
1.1	Minor Option Syllabus	2
1.2	Books	3
1.3	Aims and Objectives of Course	3
1.4	Overview	4
1.4.1	Measurement	4
1.4.2	Entanglement and many-particle systems	5
1.4.3	Decoherence	5
1.4.4	Technological Applications	6
2	Introduction	7
2.1	The postulates of quantum mechanics and the Copenhagen interpretation	7
2.1.1	The Postulates of Quantum Mechanics	7
2.1.2	Young's double slit experiment from the Copenhagen viewpoint	9
2.2	Quantum Entanglement	12
2.3	The Density Matrix	13
3	Measurement 1	15
3.1	von Neumann's model of projective measurement	15
3.2	Real Measuring Instruments	19
3.3	The Quantum Eraser	20
3.4	What Constitutes a Measurement?	20
3.5	Schrödinger's Cat	21
3.6	Wigner's Friend	21
3.7	The Einstein-Podolsky-Rosen Paradox	22
4	Some alternative interpretations of quantum mechanics	25
4.1	Many-worlds interpretation	25
4.2	Bohm's guiding waves	26
4.3	Transaction interpretation	28

4.4	Histories	28
4.5	Quantum state diffusion	30
5	Hidden variables theories	31
5.1	Bell's theorem	31
5.2	Experimental tests of Bell's theorem	35
5.3	Other tests of local realism	38
6	Entanglement	40
6.1	Schmidt decomposition	40
6.2	Effect of local operations on entanglement and the Bell States	43
6.3	Entanglement concentration	44
6.4	Further examples of entanglement manipulation	45
6.5	Three and more particle entanglement	46
6.6	Push Button Entanglement	47
7	Measurement 2	49
7.1	Generalised Measurements and Positive Operator-Valued Measure	49
7.2	Implementation of POVMs	52
7.3	Weak Measurements	53
8	Decoherence	59
8.1	Measurement and Decoherence	59
8.2	Decoherence: How Long does it take?	65
8.3	Experimental tests of decoherence	69
9	Quantum Cryptography	71
9.1	The RSA public key encryption scheme	71
9.2	The no-cloning theorem	72
9.3	Quantum Cryptography	73
9.4	Error correction and privacy amplification	75
9.5	Practical QKD	76
10	Quantum Teleportation	78
10.1	The Bell states	78
10.2	Quantum Teleportation	79
10.3	Experimental implementations of teleportation	80
11	Introduction to Quantum Computing	82
11.1	The elements of a quantum computer	83
11.2	Deutsch's Problem	84
11.3	Shor's factorisation algorithm	86
11.3.1	Classical Example	86
11.3.2	Quantum algorithm	87
11.4	Grover's database search algorithm	89
11.4.1	Grover's algorithm	89
11.4.2	Grover's algorithm in the two-qubit case ($N = 4$)	90

Contents

	1
11.5 Errors	91
11.6 Quantum error-correcting codes	94
11.7 Experimental systems for implementing quantum computing.	98
11.7.1 Ion traps	99
11.7.2 Nuclear Magnetic Resonance	99
11.7.3 Superconducting systems	100
11.7.4 Semiconductor systems	100
11.7.5 Linear optics	101

1

Preface

These notes are an adapted version of those given by Prof. M. C. Payne, TCM Group, prior to 2007.

1.1 Minor Option Syllabus

- (i) **Introduction:** The postulates of quantum mechanics - the Copenhagen Interpretation. Quantum entanglement. Density matrices.
- (ii) **Measurement 1:** What constitutes a measurement? Schrödinger's cat and Wigner's friend. The Einstein-Podolsky -Rosen paradox.
- (iii) **Some alternative interpretations of quantum mechanics:** Many worlds. Bohm's guiding waves. Transaction interpretation. Histories. Quantum state diffusion.
- (iv) **Hidden variables theories:** Bell's theorem; experimental tests.
- (v) **Quantum Entanglement:** Bipartite systems: Schmidt decomposition, reduced density matrix, entanglement measures. Tripartite systems.
- (vi) **Measurement 2:** Positive Operator-Valued Measure (POVM); Weak measurements.
- (vii) **Decoherence:** Decoherence time.
- (viii) **Quantum cryptography:** The BB84 protocol. The no-cloning theorem. Eavesdropping strategies. Privacy amplification. Other protocols. Experimental realisations.
- (ix) **Quantum teleportation:** Theoretical strategy and experimental realisations.
- (x) **Quantum computing:** Qubits. Logical operations. Algorithms for quantum computers: factorisation, database searches. Error correction. Possible systems for implementing quantum computing: ion traps; nuclear magnetic resonance; semiconductor quantum dots.

1.2 Books

An easy to understand introduction to the subject can be found in the March 1998 edition of Physics World and articles on quantum information often appear in the news media.

The following books provide detailed coverage of parts of the course:

- (i) *Quantum Computation and Quantum Information*, Nielson MA, Chuang IL (CUP 2000)
- (ii) *The Physics of Quantum Information*, Bouwmeester R, Ekart A, Zeilinger A (Spring 2000)
- (iii) *Introduction to Quantum Computation and Information*, H.-K. Lo, S. Popescu and T. Spiller (World Scientific 1998).
- (iv) *Quantum Mechanics*, Rae A I M (3rd edn IOP 1992).
- (v) *The Interpretation of Quantum Mechanics*, Onnes R (Princeton 1994).
- (vi) *Quantum Theory: Concepts and Methods*, A. Peres (Kluwer 1993).

There are some very good resources on the World Wide Web such as at:

- (i) <http://www.theory.caltech.edu/preskill/ph229> - Lecture notes and examples for a course on Quantum Information taught by John Preskill at Caltech. Note however that this treatment is much more mathematical than the present course.
- (ii) <http://www.qubit.org> - The Quantum Information Research Group in Oxford.
- (iii) <http://www.cam.qubit.org> - The Quantum Information Research Group in Cambridge.

1.3 Aims and Objectives of Course

A study of quantum information could focus on one or more of many subjects ranging from the history and/or philosophy of science to mathematics, physics or technology. As this is a short course, my main aim will be to expose students to a wide range of concepts that could be viewed as coming within the realm of quantum information rather than trying to cover just one of them in great depth. Students attending this course should expect to: (i) gain a more complete understanding of the conceptual difficulties of quantum mechanics and of some of the attempts to overcome these difficulties; (ii) appreciate that many-particle quantum mechanical systems are complicated and can show behaviour that is not only non-classical but may also appear not to follow the conventional rules of quantum mechanics; (iii)

see that there are a range of technological innovations either under development or being researched that are based on the stranger aspects of quantum mechanics.

1.4 Overview

Below I provide a more detailed outline of the concepts that will be covered in the course divided into a number of themes. One of the difficulties with providing this overview is that many of the concepts are inter-related and so do not conveniently fit under a single category.

1.4.1 *Measurement*

Undergraduate courses on quantum mechanics tend to concentrate on solving the wave equations and predicting the outcome of measurement processes with only a marginal comment on the measurement process itself. This is not surprising since the solutions of these wave equations and the theoretical predictions for the results of the measurement process on these solutions have been found to be in perfect agreement with experiment. In contrast, the measurement process itself is very poorly understood. In this course we shall discuss some of the famous ideas and paradoxes, such as Schrödinger's cat, that have highlighted the conceptual difficulties associated with measurement in quantum mechanics.

One aim of this course is to provide a detailed understanding of which aspects of the measurement process are understood and which ones are not and to emphasize that measuring instruments are necessarily large and complex. These points will be of enormous importance when we study the quantum-mechanical evolution of even modestly sized systems in the section on decoherence. The course also introduces different types of measurements from the ones you have previously encountered in lecture courses on quantum mechanics - these are so-called weak measurements and measurements that project a quantum state to one of a set of non-orthogonal states, so called positive operator-valued measures or POVMs. The lecture course includes a review of some alternative formulations to 'quantum mechanics', many of which claim to overcome the problem of understanding measurement. The degree to which these competing formulations actually achieve this aim will be discussed.

1.4.2 Entanglement and many-particle systems

All of the exact solutions of quantum mechanics are for one-particle systems. The wave equations of real systems containing many particles cannot be solved exactly. Mathematically, this is primarily because the wavefunction of a many particle system cannot be written as a product of single particle wavefunctions, even if this product is correctly symmetrised. If we are using state vectors the equivalent statement is that the state vector for a many-particle system cannot be written in terms of a direct product of state vectors for each particle. Instead the many-particle wavefunction has to be expanded in a basis set that is the direct product of the basis sets for each particle and a general many-body wavefunction is a sum of products of these single-particle basis functions. Thus for spin 1/2 particles the number of basis functions required to expand the wavefunction of an N particle system is 2^N . One consequence of this scaling has been stated already - it makes solving real quantum-mechanical problems for even modest numbers of particles totally intractable. However, there are much more subtle effects in quantum mechanics that are associated with the fact that the many-particle wavefunction cannot be written as a product of single particle wavefunctions, even when the individual particles are separated by very large distances. Thus the particles are correlated in subtle ways, even if they do not interact and are separated by huge distances. This is essentially the crux of the Einstein-Podolsky-Rosen paradox (EPR Paradox) which will be discussed in the lectures. A further consequence of the vast number of basis states required to expand a many particle wavefunction is that the resulting energy spectrum of such a system is always dense since the width of the energy spectrum usually only increases at most linearly with the number of particles N . This will be of particular relevance in section on decoherence.

1.4.3 Decoherence

There are a number of very important concepts that relate to the properties of many-particle quantum mechanical systems - the culmination of all these ideas being the phenomenon of decoherence. The first of these concepts is that interacting systems always produce entangled wavefunctions - this has the consequence that the behaviour of each individual particle no longer always appears to follow the conventional quantum-mechanical rules. The second concept is that when one has access to just one part of an entangled many-particle system the rules of the measurement process are not the conventional rules that appear in the postulates of quantum mechanics. Putting all of these ideas together leads to the phenomenon of decoherence

which is the loss of coherence between, say, the alive and dead cat states of Schrödinger's cat. This removes some of the conceptual problems associated with quantum-mechanical measurement but not all of them - the cat is still half alive and half dead but is no longer in a coherent superposition of these two states and/or fluctuating between different states.

1.4.4 Technological Applications

Remarkably, over the last few years a range of technological applications have been proposed that exploit these weird or difficult aspects of quantum mechanics which for the previous sixty years had been hardly discussed. The three applications that we shall discuss in this course are quantum cryptography, quantum teleportation and quantum computing. It should be emphasized that real experimental progress has been made in all of these areas - these are not just off the wall theoretical ideas - indeed in the case of quantum cryptography this technology is now commercially available. There are other new technological applications that I shall not discuss, in particular a quantum analogue of the sort of information theory that is covered in detail in Prof. MacKay's lecture course on Information Theory. The need to cover so much of the basic ideas of this field for anyone not attending those lectures made it sensible to simply not cover this area within this lecture course.

2

Introduction

This first section of the course provides (i) a revision of the standard interpretation of quantum mechanics - the Copenhagen interpretation; (ii) a description of the difference between a product of single-particle wave functions and a true many-body wave function thus providing an introduction to the concept of quantum entanglement and the complexity of the many-body wave function (iii) a review of the use of the density operator or density matrix which we will make use of in some later sections of the lecture course.

2.1 The postulates of quantum mechanics and the Copenhagen interpretation

The standard interpretation of quantum mechanics is usually referred to as the Copenhagen interpretation and is due to the work of Bohr who did much of the pioneering work in understanding the implications of quantum mechanics following its development in the mid 1920's. The Copenhagen interpretation is usually expressed in the form of the following 5+1 postulates, which you will have seen in a slightly different form in your Quantum Mechanics 1 Course in 1B.

2.1.1 The Postulates of Quantum Mechanics

- 1 The quantum state of a system is represented by a vector in its Hilbert space.
- 2 Quantum evolutions are unitary. (e.g, generated by the time-dependent Schrödinger equation).
- 3 The immediate repetition of a measurement yields the same outcome.
- 4 Measurement outcomes are restricted to an orthonormal set $\{|s_k\rangle\}$ of eigenstates of the measurement observable.

- 5 The probability of finding a given outcome is $p_k = |\langle s_k | \psi \rangle|^2$, where $|\psi\rangle$ is the pre-existing state of the system.
- 6 The state of a composite quantum system is a vector in the tensor product of the constituent Hilbert spaces.

Although there is agreement that these are the postulates of quantum mechanics according to the Copenhagen interpretation, what is far less clear is what the Copenhagen interpretation actually is. Bohr wrote relatively little about his ideas and indeed many of the ideas were developed verbally in sparring matches between himself and Einstein. So there are probably more versions of the Copenhagen interpretation, owing to different authors propounding their own understanding of it, than there are different formulations of quantum mechanics!! However, most authors do agree on certain elements of the Copenhagen interpretation and I shall try to identify them here.

If we look at the postulates then postulate 1 defines the importance of the state vector $|\psi\rangle$. This postulate, with a few further conditions, is the starting point for the mathematical formulation of quantum mechanics using the ideas of Hilbert spaces (A Hilbert space is an inner product space an abstract vector space in which distances and angles can be measured which is ‘complete’, meaning that if a sequence of vectors approaches a limit, then that limit is guaranteed to be in the space as well (Wikipedia)). It should be noted that the axioms of quantum mechanics can be stated without any mention of Hilbert spaces and that they emerge naturally from more ‘reasonable’ axioms (Hardy quant-ph 0101012).

Postulate 6 (Zureck Phys. Rev. A **76** 052110 (2007)) is the so-called complexity postulate. It is unlikely you will have seen it before. It describes the form of the Hilbert space for a composite quantum-mechanical system made from a number of interacting component quantum-mechanical systems. It is often omitted since the direct product of a set of Hilbert spaces is again a Hilbert space and therefore postulate 1 covers this. I have included it here because it emphasizes the origin of entanglement in the postulates and it is therefore the key to the power of all quantum information technologies. If postulate 6 were found to be untrue for large systems it is also their weakness.

Postulate 2 defines how the state vector evolves in time. Postulates 3,4 and 5 all refer to the process of measurement carried out on a state vector $|\psi\rangle$. In particular, these postulates make it clear that measurement is a probabilistic process, the probability of each possible outcome being defined in postulate 5. If this probabilistic nature was not bad enough the 3rd

postulate shows that the process of measurement leads to a change of the state vector which was $|\psi\rangle$ immediately before the measurement but this is replaced by the state vector $|s_k\rangle$ as a result of the measurement. This process is sometimes referred to as wave function reduction. There is an obvious conflict between postulates 1,2 and the measurement postulates, since they define a deterministic evolution of the state vector whereas the measurement postulates define a probabilistic evolution of the state vector. Thus within the Copenhagen interpretation the things that carry out the measurement process are not quantum mechanical. Nowadays this seems improbable as we can increasingly show that all the bits of the measuring apparatus are indeed happily obeying Schrödinger's equation. However, in the Copenhagen interpretation there is a clear distinction between microscopic objects and the macroscopic objects that make measurements on the microscopic objects. We shall see parts of the debate about this separation when we discuss Schrödinger's cat and Wigner's friend later in the course.

There are also attempts being made to reduce the number of postulates. One of the most recent is by Zureck (Phys. Rev. A **76** 052110 (2007)) in which he shows that postulates 1,2,3 and 6 imply that the measured states of a system must be orthogonal - postulate 4. The 'Relative state' formulation of quantum mechanics by Hugh Everett III (Rev. Mod. Phys 29 454 (1957)) shows that a consistent form of quantum mechanics may be defined without any need for reduction. We will discuss this in more detail in section 4.1

2.1.2 *Young's double slit experiment from the Copenhagen viewpoint*

Young's double slit experiment is one of the classic experiments to demonstrate the wave properties of matter. However, suitable modifications of the experimental arrangement can also be used to address questions such as 'which slit did the particle pass through'. The basic double-slit experiment is illustrated in figure 2.1.

Figure 2.1 shows the interference pattern we expect to observe in a Young's slits experiment. However, this is very different from the actual outcome when the experiment is performed on a single particle. In this case, given the initial wave function of the particle selected by the first slit, we can use Postulate 2 to evolve this wave function in time until the particle reaches the screen where it is detected by some suitably chosen measuring apparatus, possibly a photographic plate. In the Copenhagen interpretation, the only information we can infer from this experiment is associated with the measurement of the position of the particle when it reaches the screen. Ob-

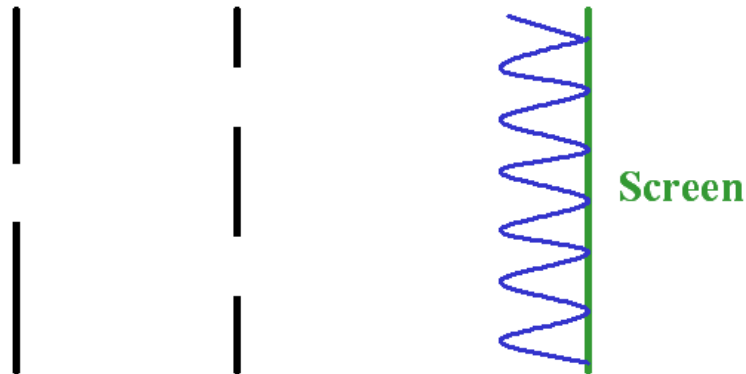


Fig. 2.1. Schematic illustration of Young's double slit experiment.

viously very limited information is obtained if we only record the positions of a few particles on the screen. Only when a very large number of particles have passed through the apparatus will our measurement tell us how the probability of finding a particle varies across the screen.

Questions such as - which of the two slits the particle passed through - have no meaning in the Copenhagen interpretation since no measurement has been carried out to answer this question. It is possible to address this question in the Copenhagen interpretation but only by inserting measuring apparatus close to the two slits. In this case the system is very different from the original two slits apparatus as can be seen in figure 2.2.

If the path taken by the particle is detected by say, a photon, then it can be shown (see Examples Sheet 1) that the momentum imparted to the particle by the photon is sufficient to destroy the interference pattern. This is represented by the uniform probability for the position of the particle across the screen in figure 2.2. However, Dürr, Nonn and Rempe have performed a 'which way' experiment in which the momentum imparted to the particle was not sufficient to destroy the interference pattern on the screen (Nature **395** 33 (1998), and News and Views p12 same issue). However, the interference pattern still disappeared in their experiment. We will examine their experiment in detail later in the course, where it will be shown that the destruction of the interference pattern is associated with quantum entanglement and measurement on only part of the entangled system. These ideas are discussed in detail in section 6 of the course.

There are many other modifications of the Young's slit experiment that

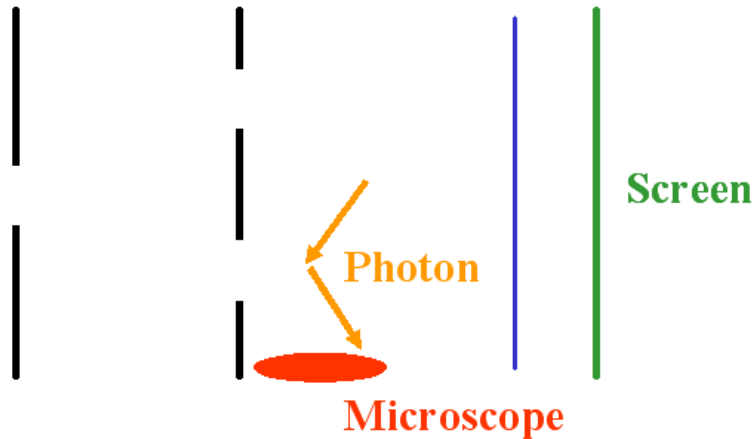


Fig. 2.2. Modification of Young's double-slit experiment to determine which slit the particle went through.

have been introduced to illustrate some of the conceptual difficulties in quantum mechanics and in alternative formulations.

One of these is Wheeler's delayed choice experiment which addresses the issue of whether particles *choose* whether to show wavelike or particle like quantum mechanical effects by knowing in advance what measurement they are *about* to encounter. In Wheeler's delayed choice experiment (Mathematical Foundations of Quantum Theory, edited by A.R. Marlow, Academic Press, 1978)the choice of whether to observe wave-like properties of the particle by observing interference effects on a screen or particle-like properties by focussing a telescope on just one of the double slits is made after the particle has passed through the slits. According to the results of the double slit experiment, if experimenters do something to learn which slit the photon goes through, they change the outcome of the experiment and the behavior of the photon. If the experimenters know which slit it goes through, the photon will behave as a particle. If they do not know which slit it goes through, the photon will behave as if it were a wave when it is given an opportunity to interfere with itself. The double-slit experiment is meant to observe phenomena that indicate whether light has a particle nature or a wave nature. The fundamental lesson of Wheeler's delayed choice experiment is that the result depends on whether the experiment is set up to detect waves or particles.

You may think that I have belaboured the importance of measurement but this element is critical in the Copenhagen interpretation. Essential

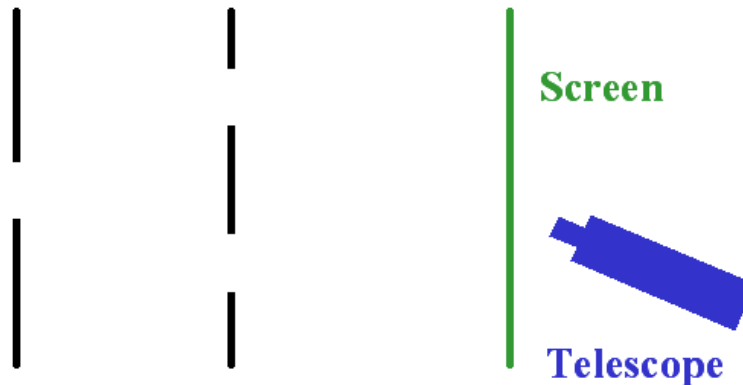


Fig. 2.3. Wheeler's delayed choice experiment. The experimenter chooses whether to observe wave-like or particle-like properties after the particle has passed through the double slits.

to the Copenhagen Interpretation is the idea that you cannot know some property of a particle without making a measurement on it. It is precisely this statement that is contradicted by the Einstein-Podolsky-Rosen paradox and so it is critical to understand this element of Copenhagenism.

2.2 Quantum Entanglement

The idea of quantum entanglement derives from postulate 6 and will be re-visited repeatedly during this course. Put most simply it is the statement that, in general, quantum-mechanical wave functions of many-particle systems cannot be written in terms of a product of single particle wave functions (or equivalently the state-vector cannot be written as a direct product of the state vectors of these single particle states)

$$\Psi(r_1, r_2, \dots, r_n) \neq \Psi(r_1)\Psi(r_2)\dots\Psi(r_n) \quad (2.1)$$

A simple example of this is the singlet spin wave function

$$\frac{1}{\sqrt{2}}(\alpha_1\beta_2 - \beta_1\alpha_2) \quad (2.2)$$

This cannot be written as a unique wave function for particle 1 times a unique wave function for particle 2, instead the wave functions of the two particles are inextricably linked together and we say that the wave functions of the two particles are entangled.

Mathematically, the reason that entangled wave functions can exist is that the Hilbert space for the many-particle wave functions is a direct product of

the Hilbert spaces for all the particles in the system. For example for three spin-half particles the basis set consists of the following 8 basis functions

$$\alpha_1\alpha_2\alpha_3 \quad \alpha_1\alpha_2\beta_3 \quad \alpha_1\beta_2\alpha_3 \quad \alpha_1\beta_2\beta_3 \quad \beta_1\alpha_2\alpha_3 \quad \beta_1\alpha_2\beta_3 \quad \beta_1\beta_2\alpha_3 \quad \beta_1\beta_2\beta_3 \quad (2.3)$$

In general, as mentioned previously, for N spin-half particles there will be 2^N basis functions. Thus even for quantum-mechanical particles with the smallest non-trivial number of basis states you can see that calculating the many-body wave function is going to be a lot of work for even modest numbers of particles. For example the many body wave function for a system of 50 spin-half particles will have $2^{50} = 10^{15}$ expansion coefficients, which is close to the memory limit of the very biggest computers presently available. For 100 particles, the number of expansion coefficients is larger than the number of particles in the universe! We shall see later in this course that this scaling of the number of basis states with the number of particles is the fundamental reason for the huge potential power of quantum computers.

The combination of entanglement and quantum-mechanical measurement leads to the Einstein-Podolsky-Rosen (EPR) paradox, which will be discussed in section 3.7. Entanglement is increasingly viewed as a resource that can be exploited to perform certain tasks. Section 6 of the course focusses on the problems of creating, manipulating and quantifying the degree of entanglement in a quantum-mechanical system.

2.3 The Density Matrix

Any quantum-mechanical system is likely to become entangled with the environment surrounding it. For example, a system of N interacting electrons trapped in some experimental system will become entangled with the atoms that make up the system. Under these circumstances it can be useful to represent our knowledge of the quantum-mechanical system in a statistical way as an ensemble of the possible states pure states of the system together with their probability weights. This can be done using a density operator or density matrix for the system.

The definition of the density operator or density matrix ρ for an ensemble in which state $|\psi_a\rangle$ has weight w_a is

$$\rho = \sum_a w_a |\psi_a\rangle \langle\psi_a| \quad (2.4)$$

To give some examples for a spin-half particle - if the particle is in the up

state then the density matrix is

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad (2.5)$$

if the particle is in the down state the density matrix is

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad (2.6)$$

and if the spin points in the $+x$ direction the density matrix is

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}. \quad (2.7)$$

If the entanglement with the environment is such that the spin has an equal probability of being up or down then

$$\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}. \quad (2.8)$$

Note that the trace of all these density matrices is one - this is basically the normalisation constraint for the wave functions combined with our definition for the weights

$$\sum_a w_a = 1. \quad (2.9)$$

Note also that for pure states the trace of ρ^2 is also 1 but that for the unpolarised ensemble $\text{Tr}\rho^2 < 1$. This result is true for all ensembles consisting of a number of different states, these are sometimes referred to as mixed states.

In order to relate the density matrix to the physical observables of the system we use the following results. The expectation value of an observable A is given by $\text{Tr}\rho\hat{A}$. If we want to find the probability of getting result a_i when we measure observable A we simply transform the density matrix so that it is written in terms of the eigenstates of A such that

$$\rho_{i,j} = \sum_a c_{j,a}^* c_{i,a} |\phi_i\rangle \langle\phi_j| \quad (2.10)$$

It can be seen that the i 'th diagonal element of the density matrix in this representation gives the probability of obtaining the result a_i from the measurement of A .

3

Measurement 1

As discussed in the previous lectures, measurement is the most poorly understood area of quantum mechanics. In this section of the course I will present von Neumann's model for the measurement process. I shall describe a number of real measuring instruments and identify the common features shared by all of them. I shall show how information about a quantum system that is, in principle, measurable can be lost if a measurement is not performed. Finally I shall discuss the difficulties associated with measurement in the Copenhagen interpretation and describe some of the paradoxes that have been devised in order to identify the problems with our understanding of quantum mechanical measurement and the Copenhagen Interpretation. In the next section of the course, I shall discuss a number of alternative formulations of quantum mechanics. These formulations must of course predict the same time-evolution of physically observable quantities as predicted in the Copenhagen Interpretation since these agree with experiment but they attempt to provide more rigorous or believable models of the measurement process. As mentioned previously in the overview, later in the course I shall discuss two other forms of measurement that are important in quantum information theory, namely weak measurement and positive operator value measures or POVMs.

3.1 von Neumann's model of projective measurement

von Neumann proposed a simple quantum-mechanical model for the measurement process. His model applies to projective measurements, which are probably all the quantum mechanical measurements you have come across in the past. These are measurements of physical properties associated with Hermitian operators where the measurement is a symmetry-breaking event that projects the initial quantum state (ie. the quantum state immediately

prior to measurement) onto one of set of eigenstates of the Hermitian operator. Later in this course we shall see that there are other possible types of measurement. A description of von Neumann's model for measurement can be found in many reference texts and here I will essentially follow the analysis in Preskill's lecture notes.

If we wish to measure an observable M , we turn on a coupling between this observable and a 'pointer' variable that represents the read-out of the apparatus. The model does not contain any of the microscopic details of the measuring apparatus - we shall return to this point later. The pointer can be thought of as some particle that propagates freely except for this coupling to the quantum system that is measured. If the reading of the pointer is to be conventionally regarded as a useful measurement the wave packet associated with the pointer must be narrow enough in position space for its own quantum fluctuations to have a negligible effect on the reading. However, equally important, it must not be too narrow that the spreading due to the uncertainty in the momentum causes appreciable broadening of the wave packet while the measuring process takes place. On these grounds alone the pointer will have to be large compared to the system being measured. When we discuss weak measurement, we shall see that this type of measurement relies on the pointer wave function being much larger than the displacement caused by the measurement.

The Hamiltonian describing the coupling of the quantum system to the pointer has the form

$$\hat{H} = \hat{H}_0 + \frac{1}{2m} \hat{P}^2 + \lambda \hat{M} \hat{P} \quad (3.1)$$

where $\hat{P}^2/2m$ is the Hamiltonian of the free pointer particle, H_0 is the unperturbed Hamiltonian of the system to be measured, and λ is a coupling constant that can be turned on and off as desired and \hat{M} is the operator corresponding to the observable to be measured M . In this model the observable to be measured is coupled to the momentum \hat{P} of the pointer. Following the argument presented above, we shall ignore the free motion of the pointer and hence neglect the term $\hat{P}^2/2m$ in the above equation from now on.

If \hat{M} does not commute with \hat{H}_0 we have to worry about how the observable evolves during the course of the measurement. To simplify the analysis we shall assume that either these operators commute or that the measurement is carried out quickly enough that the free evolution of the system can be neglected during the measurement process. With all these somewhat outrageous assumptions the Hamiltonian can be approximated as

$\hat{H} = \lambda \hat{M} \hat{P}$, where the two operators \hat{M} and \hat{P} obviously commute since they are operators for different systems. The time evolution operator for the entire system is simply

$$U(t) = \exp \left[\frac{-i\lambda t \hat{M} \hat{P}}{\hbar} \right] \quad (3.2)$$

Expanding the operator \hat{M} in the basis in which it is diagonal

$$M = \sum_a |a\rangle M_a \langle a| \quad (3.3)$$

$U(t)$ can be expressed as

$$U(t) = \sum_a |a\rangle \exp \left[\frac{-i\lambda t M_a \hat{P}}{\hbar} \right] \langle a| \quad (3.4)$$

We note that the operator \hat{P} generates a translation of the position of the pointer as follows

$$\exp \left[\frac{-ix_0 \hat{P}}{\hbar} \right] \Psi(x) = \Psi(x - x_0) \quad (3.5)$$

which can be most easily seen by expanding the exponential and remembering that

$$P = \frac{\hbar}{i} \frac{\partial}{\partial x}. \quad (3.6)$$

If our quantum system is initially in the state

$$|\phi\rangle = \sum_a \alpha_a |a\rangle, \quad (3.7)$$

which is a superposition of eigenstates of the operator \hat{M} , and it is not entangled with the wave function of the pointer then the initial state of the entire system can be written

$$|\Psi(0)\rangle = |\phi\rangle \otimes |\Psi(x)\rangle \quad (3.8)$$

where $|\Psi(x)\rangle$ is the initial state vector of the pointer. After a time t the state of the entire system has evolved to

$$U(t) \left(\sum_a \alpha_a |a\rangle \otimes |\Psi(x)\rangle \right) = \sum_a \alpha_a |a\rangle \otimes |\Psi(x - \lambda t M_a)\rangle \quad (3.9)$$

Thus it can be seen that the wave function of the pointer is now entangled with the wave function of the quantum system and that the position of

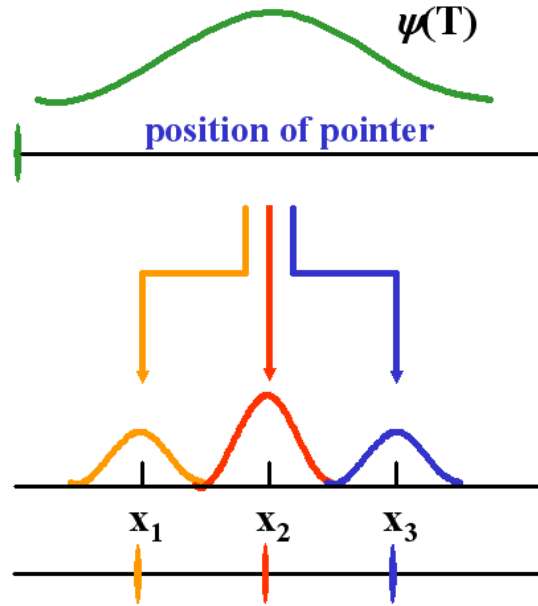


Fig. 3.1. Schematic illustration of the evolution of a quantum state and the pointer of measuring instrument from their states before the measurement to their final entangled state at the end of the measurement in von Neumann's model. Note that wave function collapse/ state reduction has not yet occurred.

the pointer is correlated with the value of the observable M as illustrated schematically in figure 3.1

If the pointer wavepacket is narrow enough for us to resolve all the values of M_a that occur, then when we observe the position of the pointer we will select an eigenstate of the operator \hat{M} . With probability $|\alpha_a|^2$ we shall detect that the pointer has shifted its position by $\lambda t M_a$ in which case we shall have chosen the eigenstate $|a\rangle$ of the operator \hat{M} . Thus as a result of the interaction between the quantum system and the pointer and our observation of the outcome we conclude that the initial state $|\phi\rangle$ of the quantum system is projected to the state $|a\rangle$ with probability $|\langle a|\phi\rangle|^2$. This is von Neumann's model of orthogonal measurement.

As an example we can consider a (neutral) spin-half particle passing through a Stern-Gerlach apparatus. To measure the spin in the z -direction the apparatus has a non-uniform magnetic field oriented in the z -direction so that $B_z = \lambda z$. If the magnetic moment of the spin half particle is $\mu\sigma$ then the Hamiltonian that describes the interaction between the spin-half

particle and the magnetic field is

$$\hat{H} = -\lambda\mu\hat{\sigma}_z z \quad (3.10)$$

If we apply the previous analysis, $\hat{\sigma}_z$ is now the observable that is measured and as it couples to the position of the particle the result of the measurement is to produce a translation in P_z , the momentum of the particle in the z -direction. Thus in this case the coupling produces an impulse on the particle in the z -direction, the direction of the impulse being correlated with the direction of the spin.

Clearly von Neumann's model does not explain all the features of the measurement process. It provides an explanation of the correlation between the position of the pointer and the different eigenstates in the wave function of the quantum mechanical system prior to measurement. However, as illustrated in figure 3.1, within the model all the different positions of the pointer are produced by the measurement process. The selection of a particular pointer reading and hence the collapse of the wave function to the state $|a_i\rangle$ is not explained by this model. Indeed these features cannot be explained by any model in which time evolution is deterministic, as it is when governed by the 2nd postulate.

3.2 Real Measuring Instruments

In lectures we shall discuss how a number of measuring instruments work. In particular we shall discuss the use of photographic film and photomultipliers for detecting photons and Geiger counters and bubble chambers for detecting charged particles. The important and unifying themes to note from the discussions of measurement devices are that (i) the devices used for detecting these particles are large and complex, hence they are necessarily macroscopic; (ii) their operation is discussed primarily in classical terms; (iii) their operation always involves a mechanism for amplification of a signal; (iv) the detection process always involves loss of energy (at least in the examples discussed though it could be some other quantity) from the particle to be detected - this is needed to take the system over a barrier that prevents spontaneous discharge of the detecting system. Obviously if we are trying to detect a single photon or a single charged particle the efficiency of the detector is of critical importance. At a basic quantum mechanical level it is difficult to think of many transitions that occur with an efficiency even close to 100% and so, even ignoring the difficulty of detecting any signal produced, a detector that contains a single quantum-mechanical detecting particle is bound to be a very inefficient detector. The way of overcoming

this problem is simply by scaling up the number of quantum mechanical elements in the detection apparatus. Then the chance of the initial transition occurring at some point in the detector is close to 100%. The amplification mechanism in real detectors ensures that the question of efficiency is only important at the initial stage of detection. It should be noted though that when the particle to be detected has a very low energy it may be difficult to ensure efficient detection of the particle because of the need to take the detecting apparatus over the energy (or other) barrier mentioned in (iv) above. The fact that real measuring apparatus contains large numbers of particles will be relevant when we discuss decoherence later in the course.

3.3 The Quantum Eraser

In von Neumann's model of measurement we emphasized that the selection of a single pointer position and the consequent reduction of the wavefunction of the quantum mechanical system to the state $|a\rangle$ only occurred when an observer selected one of the mixture of entangled states generated by the measurement. What happens if no such observer is present? In the case of a Stern-Gerlach experiment, if no detection of the spin-half particle takes place, there is no reason why we cannot recombine the two outgoing states from the Stern-Gerlach apparatus. In this case we will have reconstituted the original wavefunction of the quantum system (provided the phase changes along the two paths were the same) and we have lost any information about the effect of the original Stern-Gerlach experiment. The loss of this potential information that could have been obtained had the particle been detected is usually referred to as the quantum eraser. However, you should think further about this point when you have done question 3 on the first examples sheet

3.4 What Constitutes a Measurement?

von Neumann's model of measurement does not address the issue of why just one of the possible outcomes of the measurement is actually selected. This part of the measurement process is arbitrarily assigned to an observer whose properties and non-quantum mechanical evolution are never addressed. You no doubt feel that this is highly unsatisfactory. Indeed, many of the alternative formulations of quantum mechanics aim to get round this difficult issue as we shall see in later lectures. The distinction between quantum systems and systems that can actually perform a measurement and hence cause reduction of the wavefunction has been a thorn in the side of conventional quantum mechanics and continues to be hotly disputed. Paradoxes such as

Schrödinger's cat and Wigner's friend should be appreciated in the light of this enormous difficulty in conventional quantum mechanics, they were put forward in order to question the fundamental tenets of the theory.

3.5 Schrödinger's Cat

The point about Schrödinger's cat is simply to couple some macroscopic event, in this case the life or death of the cat, into the outcome of the measurement process. If we believe von Neumann's model the cat remains half dead and half alive until an external observer comes along and opens the box and observes the state of the cat thus selecting which of these outcomes is chosen. Remember this selection is not retrospective, as we saw with the quantum eraser, it is only when the external observer does their job that the actual reduction of the wavefunction takes place - it is a superposition up to this time. Our experience of the world does not include cats that are half dead and half alive and so this suggests that humans are the ultimate observers who are capable of causing wavefunction reduction. However, even this explanation has difficulties as the following example - Wigner's friend shows. Whether cats or any other animals are happy with half alive and half dead cats is an issue for philosopher's. Bohr was not worried about Schrödinger's cat. In his view the cat is macroscopic and it thus lies outside the realm of the quantum mechanics of microscopic systems and in the classical realm of macroscopic systems. Hence, Bohr simply said that the cat, being classical, cannot be in a mixture of states.

3.6 Wigner's Friend

This paradox adds a further layer of complexity to the measurement problem by including an exchange of information between two conscious observers.

One considers three systems Q is a quantum system on which a measurement of an observable A is performed; W is Wigner and F is either the measuring apparatus that measures A or is Wigner's friend who has been asked to make the measurement of the observable A . We shall assume that A has only two eigenvalues 1 and 0, which can be interpreted as 'yes' and 'no'. When the measurement is made F writes the message s_n on a card for W to see and thus to tell him what he has found.

If the initial state of Q is an eigenvector of \hat{A} then there is no paradox. The final state of $Q + F$ after the measurement is a single ket $|\psi_n\rangle \otimes |s_n\rangle$.

If, however, the initial state of Q is a superposition of eigenstates of \hat{A}

$$\sum_n c_n |\psi_n\rangle \quad (3.11)$$

then the final state of $Q + F$ is

$$\sum_n c_n |\psi_n\rangle \otimes |s_n\rangle \quad (3.12)$$

If W decides not to look at the messages sent by his friend then he believes that F is in a superposition of states until the point when he interrogates him about the outcome of the experiment, at which point he could infer either outcome which need not agree with the message written previously by F . From the viewpoint of F though he is a perfectly acceptable observer who recorded a unique outcome from the experiment which was the one written on the card.

This inconsistency of interpretation and in particular the possibility that the result observed by W does not agree with what is written on the card led to a suggestion that wavefunction collapse occurs when the collective human consciousness becomes aware of a fact. The universal consciousness suggestion would explain why all observers will agree about the outcome of the measurement.

3.7 The Einstein-Podolsky-Rosen Paradox

Einstein was never happy with quantum mechanics and formulated a series of thought experiments designed to undermine the fundamental tenets of the theory. Bohr managed to formulate rebuttals to most of these proposals, even though it is now known that much of Bohr's reasoning was flawed. Einstein's ultimate challenge to the Copenhagen interpretation was enshrined in the Einstein-Rosen-Podolsky (EPR) paradox. This overturned the basic tenet of Copenhagenism that only measurements could provide any information about the properties of a quantum system. The EPR paradox can arise whenever we are dealing with an entangled system but one of the simplest examples of the EPR paradox is the maximally entangled singlet spin state

$$\frac{1}{\sqrt{2}} (\alpha_1 \beta_2 - \beta_1 \alpha_2) \quad (3.13)$$

The meaning of the term 'maximally entangled' will be explained in section 5 of the course. Before we make any measurements on this system we

have no information about which direction the spin of particle 1 is pointing or which direction the spin of particle 2 is pointing - Copenhagen rules supreme! Now we measure the spin of particle 1 in the x -direction and let's assume that we find it is $+\hbar/2$. If we write the wavefunction of the system before measurement in terms of eigenstates of the x spin operator we find

$$\begin{aligned} & \frac{1}{\sqrt{2}} (\alpha_1 \beta_2 - \beta_1 \alpha_2) \\ = & \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2}} [\alpha_1^x + \beta_1^x] \frac{1}{\sqrt{2}} [\alpha_2^x - \beta_2^x] - \frac{1}{\sqrt{2}} [\alpha_1^x - \beta_1^x] \frac{1}{\sqrt{2}} [\alpha_2^x + \beta_2^x] \right) \\ = & \frac{1}{\sqrt{2}} (\alpha_1^x \beta_2^x - \beta_1^x \alpha_2^x) \end{aligned} \quad (3.14)$$

In fact, irrespective of which direction we choose to define the basis states, the wavefunction of the singlet state will always have the general form,

$$\frac{1}{\sqrt{2}} (\alpha_1 \beta_2 - \beta_1 \alpha_2) = \frac{1}{\sqrt{2}} (\alpha_1^e \beta_2^e - \beta_1^e \alpha_2^e) \quad (3.15)$$

where e is a unit vector. Now, returning to our measurement of the spin of particle 1 in the x -direction, if we find the result $+\hbar/2$ then, as a result of the measurement we have reduced the wavefunction to

$$\alpha_1^x \beta_2^x \quad (3.16)$$

Therefore, our measurement of particle 1 tells us not only the spin of particle 1 **but also** the spin of particle 2. This totally undermines the Copenhagen interpretation, which would claim that you cannot have any knowledge of the spin of particle 2 until you make a measurement on the particle. Now, we have to make some caveats about the above.

- (i) If the two particles are separated and you want to do something to particle 2 and exploit this information about the particle you have to wait until you receive a message about the outcome of the first measurement before you know the state of the particle. One's knowledge of the state of particle 2 is not only dependent on the measurement of particle 1 but also on some form of classical communication (CC). This is the reason why it is not possible to use variations of the EPR paradox to communicate faster than light.
- (ii) Consider a large ensemble of identical spin-singlet states and two observers who make measurements of the spins of the particles along a chosen axis. Both of them will record equal numbers of up and down spins along the chosen direction, irrespective of whether the other observer is making measurements or not. Basically, measuring

the spin of one particle does not change the statistical properties of the measurement of the second spin. It is only when the two observers compare the results of their measurements that they miraculously discover that their results are perfectly correlated. These correlations are usually referred to as EPR correlations.

Einstein proposed the EPR paradox saying that the real world cannot behave this way, the result of a measurement on one particle cannot be correlated with a measurement on another particle unless the measurements are causally related. Experiments show that indeed quantum mechanics does contain these correlations.

The combination of a measurement or some other manipulation applied to one particle of an entangled pair combined with classical communication is used extensively in quantum information, as we shall see in later sections of the course. It is usually referred to as LQCC, meaning ‘local quantum operations and classical communication’.

4

Some alternative interpretations of quantum mechanics

In this section we shall discuss a number of alternatives to the theory of quantum mechanics and the Copenhagen interpretation. As mentioned in the previous section of the course, these formulations must predict the same time-evolution of physically observable quantities as predicted in the Copenhagen Interpretation since these agree with experiment but they attempt to provide more rigorous or ‘believable’ models of the measurement process. You can read more in Chapter 11 of Rae’s book and in an article by John Cramer (Rev. Mod. Phys. **58** 647 (1986)). Both of these give a good overview of the problems of quantum mechanics including a discussion of the Schrödinger Cat and Wigner’s friend paradoxes.

4.1 Many-worlds interpretation

This approach to the problem of measurement in quantum mechanics was based on Everett’s PhD thesis (see Rev. Mod. Phys **29** 454 (1957)) under the supervision of Wheeler and is known as the Everett-Wheeler, Everett-Wheeler-Graham or many world’s interpretation of quantum mechanics. The basic tenet of this approach is that collapse of the state vector never occurs. Instead, all outcomes of a measurement occur and the universe splits into a number of different universes in each of which a different outcome of the measurement is recorded. Thus the entire measurement process in the many-worlds interpretation is exactly as shown in figure 3.1. The crucial progress that Everett made was that he showed that the universes produced could not interact with each other. The many-worlds interpretation overcomes many of the conceptual problems associated with measurement. It is no longer a probabilistic process since all outcomes of the measurement are recorded and it avoids the problem of state reduction. On the negative side the idea that there are lots of copies of ourselves in all these different universes is somewhat

unsettling. For instance in another universe you may have got the top 1st in Pt II physics! The theory rapidly produces vast numbers of different universes but it should be emphasized that the number is not so enormous as one might expect. In particular it is only when macroscopically different outcomes occur depending purely on one quantum-mechanical measurement that the universes split. As most of our everyday actions are governed by essentially classical systems in the sense that they are dependent on large numbers of events before crucial decisions are taken then it is clear that each individual quantum measurement in this case does not cause the universe to split. I have not come across a discussion of the possibility for universes to unbranch and this might give difficulties in situations like the quantum eraser. However, I believe that the ideas presented in the section 8 on decoherence would answer most of these problems - essentially the worlds only split once decoherence has taken place. This leaves the only objection to the many-worlds interpretation a degree of unease that there are many copies of oneself some possibly winning Olympic gold medals, others possibly coming top in Physics in Cambridge and we can't find out!

4.2 Bohm's guiding waves

The idea behind Bohm's guiding waves is that particles have a real position and velocity, and the role of the wave function is to guide the motion of the particles so that their statistical properties agree with the quantum-mechanical description. This idea was first proposed by de Broglie in 1927 but Bohm's name is strongly associated with developing the idea and it is sometimes referred to as de Broglie-Bohm theory. This theory is actually the first example of a hidden variable theory, such theories will be discussed in greater detail in section 5.

So how does it work? We start with the Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi + V(x, y, z) \Psi \quad (4.1)$$

and define quantities R and S as real functions of r such that

$$\Psi = R \exp\left(\frac{iS}{\hbar}\right) \quad (4.2)$$

Substituting this form in the Schrödinger equation and separating the real and imaginary parts gives:

$$\frac{\partial S}{\partial t} = -\left[\frac{|\nabla S|^2}{2m} + V - \frac{\hbar^2}{2m} \frac{\nabla^2 R}{R} \right]$$

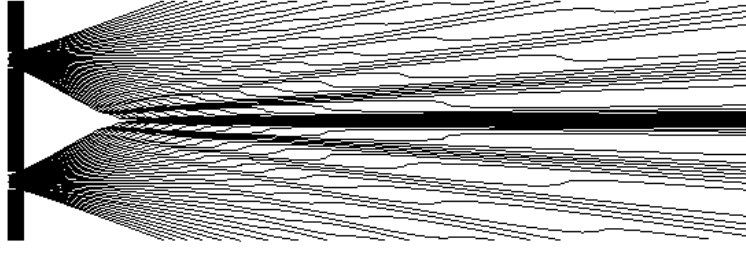


Fig. 4.1. An ensemble of trajectories for the two-slit experiment in Bohm's guiding wave description. In the absence of the quantum potential the trajectories would be straight lines. (From Quantum Theory Without Observers by Sheldon Goldstein: Physics Today, March 1998, p42-46 and April 1998, p38-42.)

$$\frac{\partial R}{\partial t} = -\frac{1}{2m} [R\nabla^2 S + 2\nabla R \cdot \nabla S] \quad (4.3)$$

If we define $P = R^2$ then we can rewrite these two equations in the form

$$\frac{\partial P}{\partial t} + \nabla \cdot \frac{P\nabla S}{m} = 0 \quad (4.4)$$

$$\frac{\partial S}{\partial t} + \frac{|\nabla S|^2}{2m} + V + Q = 0 \quad (4.5)$$

where Q is defined by

$$Q = -\frac{\hbar^2}{4m} \left[\frac{\nabla^2 P}{P} - \frac{1}{2} \frac{|\nabla P|^2}{P} \right] \quad (4.6)$$

So what is the point of all this? It can be seen that $P = |\Psi|^2$, which is the probability density associated with the particle. If we further assume that a particle at position r has a velocity v where $v = \nabla S/m$ then Eqn. 4.4 is simply the continuity equation. With this form for the particle velocity, the second and third terms of Eqn. 4.5 are now just the kinetic and potential energies of the particle. The fourth term in Eqn. 4.5, Q , has no classical analogue and according to this theory it is an additional potential known as the quantum potential - it is this potential that allows purely quantum mechanical phenomena such as tunnelling, and interference in the two slit experiment. The action of this potential can be clearly seen in figure 4.1. However, the quantum potential does not conform to our understanding of the origin of all other potentials in physics since there is no fundamental interaction that causes it and perhaps more worryingly it is non-local, requiring action at a distance to reproduce the quantum mechanical results.

4.3 Transaction interpretation

This interpretation of quantum mechanics has been pioneered by John Cramer. He has written a comprehensive review of this interpretation in *Rev. Mod. Phys.* **58** 647 (1986). The basic idea behind the transaction interpretation is that all processes in quantum mechanics take place as a result of a ‘transaction’ between an emitter and an absorber. An emitter sends out the wave Ψ , this is a normal propagating wave which is usually referred to as a retarded wave. When this wave encounters an absorber the absorber emits an advanced wave Ψ^* . This wave is a solution of the Schrödinger equation but for $-t$ rather than t since

$$i\hbar \frac{\partial \Psi^*}{\partial(-t)} = -\frac{\hbar^2}{2m} \nabla^2 \Psi^* + V \Psi^* \quad (4.7)$$

Thus represents a wave travelling back in time. All potential absorbers of the wave at all future times indicate their availability to take part in a transaction with the emitter by sending advanced waves back to the emitter. The emitter then chooses which transaction is going to take place and a single transaction takes place in which a particle is exchanged between the emitter and one of the absorbers. This sequence of events is illustrated in figure 4.2.

This interpretation overcomes the problem of which result is actually obtained in any measurement - information about all possible outcomes of the measurement process is sent back to the emitter, which decides which of the outcomes actually occurs and initiates a transaction that produces this outcome. The drawback of the interpretation is that the theory is infinitely non-local in time - the retarded wave samples all future configurations of absorbers and this information is brought back to the emitter before it chooses which transaction actually takes place.

4.4 Histories

The work on histories in quantum mechanics is due to Griffiths. You can read about this subject in chapters 5 and 6 of Omnes’s book. As its name suggests, a history of a physical system is simply a series of properties occurring at different times. Within the Copenhagen interpretation the only things that can enter a history of a physical system are the outcomes of measurements on the system between measurements. Within the Copenhagen interpretation unmeasured properties of a quantum system have no physical values associated with them. Already we have seen that the EPR paradox undermined this basic tenet of Copenhagenism. Griffiths work uses concepts

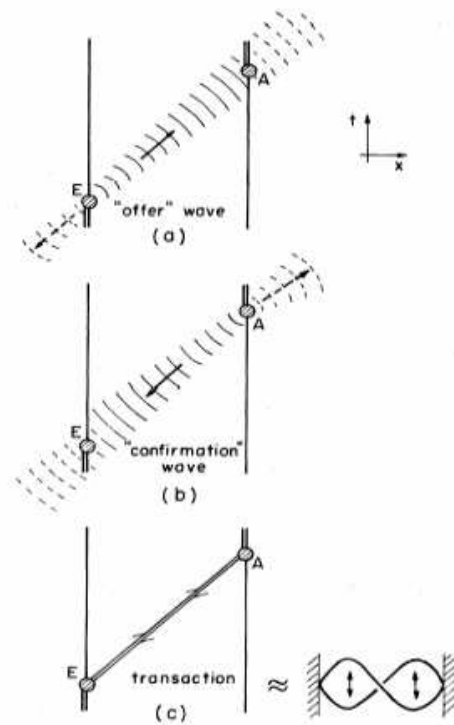


Fig. 4.2. Schematic illustration of the transaction interpretation (from Cramer's article)

from logic and probability theory to show that only certain histories are logically consistent and hence there is an extent to which the properties of a quantum system can be described between measurements. The problem is that while there are many histories that are not consistent - for instance it can be shown that a particle will not follow an erratic course through free space - there are many histories that are consistent. Thus this approach does not allow us to uniquely associate a history with a given quantum system. For instance, consider a photon emitted from an atom and detected by a detector some distance away. Given these initial and final events one consistent history is that the photon followed a straight line trajectory between the atom and the detector. However, another consistent history is that the photon was emitted as an outgoing spherical wave and was subsequently detected by the detector. Perhaps the greatest importance of histories is that it proves that histories in which the photon moves for a time along a straight line path and then abruptly reverses its direction in free space

are not logically consistent. Hence such events can not happen (or more realistically they have a vanishing small chance of happening). The idea of consistent histories does allow us to banish some of the more outlandish suggestions for what happens between measurements in quantum mechanics

4.5 Quantum state diffusion

This approach has been pioneered by Professor Ian Percival at Queen Mary and Westfield College, London. This theory is described in his book 'Quantum State Diffusion' (CUP, 1998). The basic idea of this approach is that the interaction between a quantum system and its environment introduces a stochastic dynamics (or diffusion) into the motion of a quantum system. Within the quantum state diffusion (QSD) approach the master equation for the evolution of the density operator ρ is

$$\rho = -\frac{i}{\hbar} [\hat{H}, \rho] + \sum_j \left(L_j \rho L_j^* - \frac{1}{2} L_j^* L_j \rho - \frac{1}{2} \rho L_j^* L_j \right) \quad (4.8)$$

where the L_j are the Linblad operators or Linblads which determine the stochastic dynamics of the evolution of the quantum system. The number and form of the Linblad operators will depend on the system and property of the system of interest. For instance, in the case of a measurement corresponding to the hermitian operator M there is a single Linblad operator $L = cM$. One of the strengths of the theory is that the fluctuations in the environment produce localisation, an effect which is well known in the theory of disordered systems. Processes such as wave function reduction, which require localisation of the wave function in some function space, thus appear naturally in the quantum state diffusion approach.

5

Hidden variables theories

One of the simplest ways round the problem of measurement in quantum mechanics is to postulate that the outcome of the measurement depends on the value of a variable λ that is not accessible to us. Such theories are known as hidden-variables theories. The de-Broglie-Bohm theory discussed in the section 3.2 was the earliest hidden-variables theory, a particular particle following a deterministic trajectory which would be known if the initial position of the particle had been known, thus this is the hidden variable of the theory. One might have expected that by cunning construction it would never be possible to disprove any hidden variable theory without getting access to the hidden variables. However, the extraordinary work of John Bell, fig. 5.1 gives us a way of testing a large class of hidden-variables theory by looking at the correlations between the results of measurements of the spin of EPR pairs along different axes.

5.1 Bell's theorem

The idea of hidden variables theories is that the outcome of the measurement on a particle is uniquely determined by the value of the hidden variable, λ , associated with that particle.

Bell's theorem relates to the physics of two-state systems. The example he gives [Rev. Mod. Phys. **38** 447 (1966)] of a hidden-variable theory for such a system is as follows:

A state of the system is specified by a state vector $|\psi\rangle$ **and** a variable λ . No physical meaning is ascribed to the variable λ but it is assumed to follow some dynamical law that gives it the range $-1/2 \leq \lambda \leq 1/2$ where, in time, each value of λ is equally probable.

For a two-state system, a general measurement M will have a measure-



Fig. 5.1. John Bell receiving an honorary doctorate from Queen's University Belfast in 1988 two years before his death.

ment operator \widehat{M} of the form

$$\widehat{M} = \alpha \mathbb{I} + \beta \cdot \sigma, \quad (5.1)$$

where \mathbb{I} is the unit matrix, β is a vector and σ the vector of the Pauli matrices.

In order to simplify the example, without loss of generality, Bell chooses a coordinate system in which the system state $|\psi\rangle$ has the form

$$|\psi\rangle = |1\rangle \quad (5.2)$$

The example then specifies that measurement of M on this state yields with certainty the value

$$m_\lambda = \alpha + |\beta| \text{sign}(\lambda|\beta| + |\beta_z|/2) \text{sign}(X) \quad (5.3)$$

where

$$\begin{aligned} X &= \beta_z && \text{if } \beta_z \neq 0 \\ &= \beta_x && \text{if } \beta_z = 0, \quad \beta_x \neq 0 \\ &= \beta_y && \text{if } \beta_z = 0, \quad \text{and } \beta_x = 0 \end{aligned} \quad (5.4)$$

and

$$\begin{aligned}\text{sign}(X) &= +1 & \text{if } X \geq 0 \\ &= -1 & \text{if } X < 0\end{aligned}\tag{5.5}$$

Try some different values for β to see why this choice works.

The quantum-mechanical state specified solely by the state vector $|\psi\rangle$ is obtained by averaging over λ . This gives the expectation value

$$\langle \alpha \mathbb{I} + \beta \cdot \sigma \rangle = \int_{-1/2}^{1/2} m_\lambda d\lambda = \alpha + \beta_z\tag{5.6}$$

as required. This hidden-variable example is deterministic because the outcomes of the measurement are unique and depend on λ and the theory is also local since the measurement of one particle will not be affected by the result of a measurement on any other particle.

Bell's theorem establishes an inequality between the results of measurements on EPR pairs that allows experiments to determine if there are hidden variables specifying quantum mechanical states in nature. What follows is a derivation of this inequality:

For simplicity, we will assume that each electron in the EPR pair shares the same value of the hidden variable λ but more general derivations exist that allow for the possibility of multiple different hidden variables for each particle.

The result of measuring the z component of the spin of particle 1 will be represented by $S_{z1}(\lambda)$ and we shall use plus or minus one to denote the result. The result of a measurement of the spin of the second particle in a direction at an angle ϕ to the z -axis will be represented by $S_{\phi2}(\lambda)$.

We define a quantity $p(\lambda)$ that gives the probability of an EPR pair being produced with a value of λ between λ and $\lambda + d\lambda$ as $p(\lambda)d\lambda$. We shall assume that $p(\lambda)$ is normalised so that

$$\int p(\lambda)d\lambda = 1$$

We now consider an experiment in which the quantities S_{z1} and $S_{\phi2}$ are measured on a large number of pairs. According to the hidden-variables theory the correlation coefficient between the measurements will be

$$C(\phi) = \int S_{z1}(\lambda)S_{\phi2}(\lambda)p(\lambda)d\lambda\tag{5.7}$$

Now consider a second set of measurements in which the first detector is

not changed but the second one is set at an angle θ to the z -axis. The measurements yield a second correlation coefficient $C(\theta)$. From the two measurements we can determine

$$C(\phi) - C(\theta) = \int [S_{z1}(\lambda)S_{\phi2}(\lambda) - S_{z1}(\lambda)S_{\theta2}(\lambda)]p(\lambda)d\lambda \quad (5.8)$$

Now we need to perform some manipulations on this expression in order to get to a useful result. The first thing to do is to use the fact that each individual measurement of the spin of particle 2 is perfectly anti-correlated with a measurement of the spin of particle 1 along the same axis so that

$$\begin{aligned} S_{\theta2}(\lambda) &= -S_{\theta1}(\lambda) \\ S_{\phi2}(\lambda) &= -S_{\phi1}(\lambda) \end{aligned} \quad (5.9)$$

Substituting from Eqn. 5.9 into Eqn. 5.8 gives

$$\begin{aligned} C(\phi) - C(\theta) &= - \int S_{z1}(\lambda) [S_{\phi1}(\lambda) - S_{\theta1}(\lambda)] p(\lambda) d\lambda \\ &= - \int S_{z1}(\lambda) S_{\phi1}(\lambda) [1 - S_{\phi1}(\lambda) S_{\theta1}(\lambda)] p(\lambda) d\lambda \end{aligned} \quad (5.10)$$

where the last step follows because $S_{\phi1}(\lambda) = \pm 1$ and so $[S_{\phi1}(\lambda)]^2 = 1$. If we now take the absolute values of both sides of Eqn. 5.10 we get

$$\begin{aligned} |C(\phi) - C(\theta)| &= \left| \int S_{z1}(\lambda) S_{\phi1}(\lambda) [1 - S_{\phi1}(\lambda) S_{\theta1}(\lambda)] p(\lambda) d\lambda \right| \\ &\leq \int |S_{z1}(\lambda) S_{\phi1}(\lambda) [1 - S_{\phi1}(\lambda) S_{\theta1}(\lambda)] p(\lambda)| d\lambda \end{aligned} \quad (5.11)$$

where the last line follows because the absolute value of an integral is always less than or equal to the integral of the absolute value of the function. Now neither $p(\lambda)$ nor the term in square brackets in Eqn. 5.11 can be negative and $|S_{z1}(\lambda) S_{\phi1}(\lambda)| = 1$ because each of these quantities takes the value ± 1 so Eqn. 5.11 can be written

$$\begin{aligned} |C(\phi) - C(\theta)| &\leq \int [1 - S_{\phi1}(\lambda) S_{\theta1}(\lambda)] p(\lambda) d\lambda \\ &\leq 1 + \int S_{\phi1}(\lambda) S_{\theta2}(\lambda) p(\lambda) d\lambda \end{aligned} \quad (5.12)$$

We shall only consider cases where the z -axis and the directions defined by θ and ϕ are in the same plane. Furthermore the measurements depend only on the relative orientations of the axes and not the absolute orientation. Thus the integral in Eqn. 5.12 measures the correlation between two spin

components at an angle of $(\theta - \phi)$ which is, by definition, $C(\theta - \phi)$. Hence we can write Eqn. 5.12 as

$$|C(\phi) - C(\theta)| - C(\theta - \phi) \leq 1 \quad (5.13)$$

This inequality is known as Bell's theorem and, given the conditions we imposed at the beginning of the derivation, it applies to any local deterministic hidden variables theory.

The crucial issue is whether quantum mechanics obeys Bell's theorem. The easiest choice for θ to investigate this is to choose $\theta = 2\phi$. In this case the quantum mechanical results are

$$C(\phi) = -\cos(\phi) \quad (5.14)$$

$$C(2\phi) = -\cos(2\phi) \quad (5.15)$$

Hence, quantum mechanics can only be consistent with a local deterministic hidden-variables theory if

$$|\cos(2\phi) - \cos(\phi)| + \cos(\phi) \leq 1 \quad (5.16)$$

This inequality is violated for $0 \leq \phi \leq \pi/2$ with the maximal violation for $\phi = \pi/3$ when the left hand side of Eqn. 5.16 takes the value 1.5. Quantum mechanics as formulated is therefore **not** consistent with a local deterministic hidden-variables theory.

5.2 Experimental tests of Bell's theorem

In order to provide a definitive test of Bell's theorem it is necessary to reformulate the expression to take account of finite efficiency of the measuring apparatus, the form of Bell's theorem given in Eqn. 5.16 can only be applied if all particles are detected. It also turns out to be difficult to generate EPR pairs for spin-half particles so experiments are generally performed using EPR photon pairs that are generated in an atomic cascade process during where two photons are emitted. If the atoms start and end in $j = 0$ states, the two photons emitted must have zero total angular momentum and will behave in the same way as the singlet state of two spin 1/2 particles. The experiment thus measures the correlation between the polarization directions of the two photons. A number of experiments have been performed over a number of years to test Bell's theorem. While early experiments gave results that were consistent with quantum mechanics the limitations of the experiments meant they did not definitively show a violation of the Bell

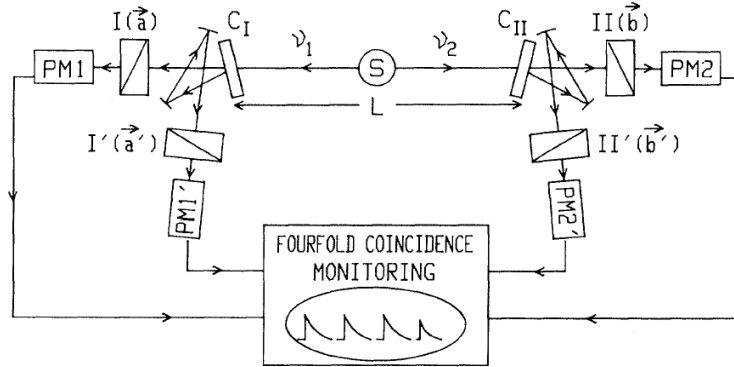


Fig. 5.2. The Aspect experiment. The optical switches C_I and C_{II} direct the outgoing photons onto one of two polarisers, I or I' for photon 1 and II or II' for photon 2. The photons are detected using photomultipliers (PM).

inequality. The first definitive experiment was performed by Alain Aspect in 1982 (Phys. Rev. Lett. **49**, p1804 (1982)) and is illustrated in figure 5.2

One feature of the Aspect experiment was the use of fast optical switches to direct each photon onto one of two possible polarisers set along direction \vec{a} or \vec{a}' for photon 1 and \vec{b} or \vec{b}' for photon 2. This prevents any loophole in the test of the Bell inequality that is based on the EPR source knowing what detectors the photons will encounter, this is essentially the issue addressed by Wheeler's delayed choice experiment. However, the low efficiency of the photomultipliers prevented a direct test of the Bell inequality. Instead the experiment tested the Clauser-Horne-Shimony-Holt inequality (Clauser *et al.*, Phys. Rev. Lett. **23**, 880 (1969)) which is a 'Bell-like' inequality in the sense that it also provides a test of local, deterministic hidden-variables theory but does this by comparing the coincidence rates, such as those measured in the Aspect experiment, with the polarisers present and with the polarisers removed. After carrying out a series of experiments to determine the coincidence rates, the quantity S is calculated from

$$S = \frac{N(\vec{a}, \vec{b})}{N(\infty, \infty)} - \frac{N(\vec{a}, \vec{b}')}{N(\infty, \infty')} + \frac{N(\vec{a}', \vec{b})}{N(\infty', \infty)} + \frac{N(\vec{a}', \vec{b}')}{N(\infty', \infty')} - \frac{N(\vec{a}', \infty)}{N(\infty', \infty)} - \frac{N(\infty, \vec{b})}{N(\infty, \infty)} \quad (5.17)$$

where $N(\vec{a}, \vec{b})$ is the coincidence rate with the polarisers along a and b , $N(\infty, \infty)$ is the coincidence rate with both of these polarisers removed, $N(\vec{a}, \infty)$ is the coincidence rate with the polariser along \vec{a} present but the polariser along \vec{b} removed, etc

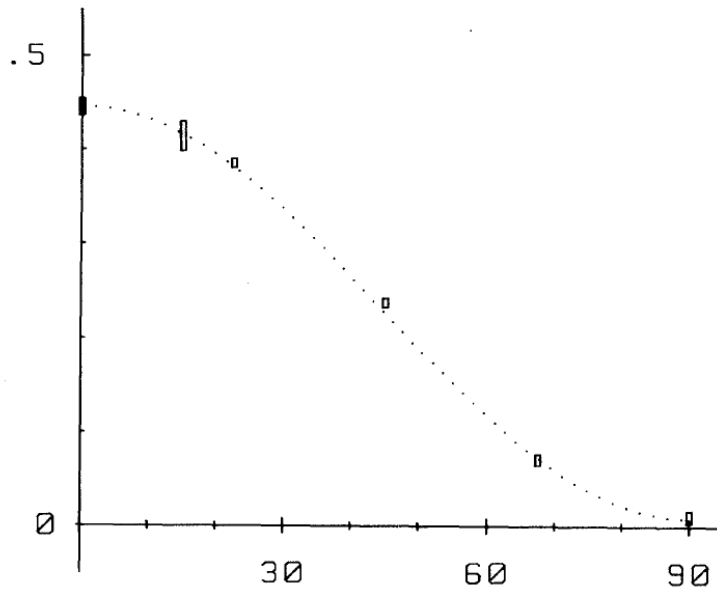


Fig. 5.3. Average normalised coincidence rates as a function of the relative orientation of the polarisers measured in the Aspect experiment. Error bars show ± 1 standard deviation. The quantum mechanical prediction is shown by the dotted line.

The Clauser-Horne-Shimony-Holt inequality shows that for any local deterministic hidden-variables theory $-1 \leq S \leq 0$.

The results of the Aspect experiment for the coincidence rates are shown in figure 5.3. The results are consistent with the quantum mechanical predictions and gave a violation of the Clauser-Horne-Shimony-Holt inequality of 5 standard deviations, the experiment was absolutely definitive.

Recently, similar experiments to the Aspect experiment have been repeated with much larger separation between the detectors using techniques developed for quantum cryptography. The first of these experiments was performed by Weihs *et al.* (Phys. Rev. Lett. **81**, 5039 (1998)). These experiments provide strict relativistic separation between the measurements thus overcoming one of the loopholes in the test of Bell's theorem. The final loophole for the definitive test of Bell's theorem is the so-called *detection* loophole. As mentioned previously, if the detectors have detection efficiencies much less than 1 then the experiments cannot directly test Bell's theorem. However, a recent experiment by Rowe *et al.* (Nature **409**, 791 (2001)) has closed the detection loophole by using beryllium ions which can

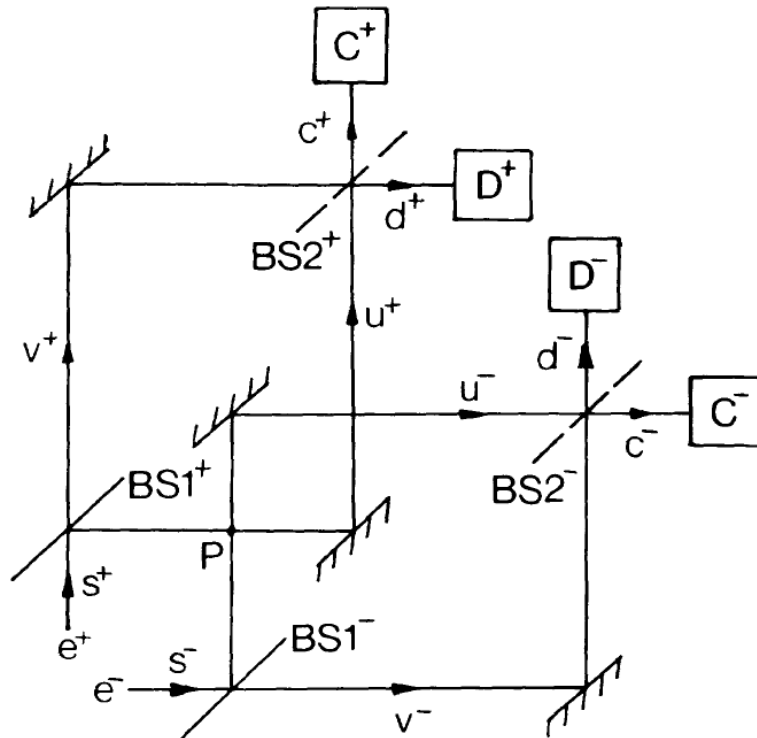


Fig. 5.4. Experimental arrangement for demonstrating Hardy's paradox. $BS1^\pm$ and $BS2^\pm$ are beamsplitters and C^\pm and D^\pm are detectors.

be detected with very high efficiency. This experiment was also the first to use massive particles rather than photons for a test of Bell's theorem.

5.3 Other tests of local realism

There are now many other tests of whether quantum mechanics is a local realistic theory. Greenberger, Horne and Zeilinger have derived a Bell-like theorem that involves a direct contradiction rather than the inequality that appears in Bell's theorem (in *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, ed. M. Kafatos (Kluwer, 1989) p74 and in *Am. J. Physics.* **58**, 1131 (1990)). However, this test requires at least three entangled particles. The Hardy paradox (*Phys. Rev. Lett.* **68**, 2981 (1992)) demonstrates Bell's theorem without inequalities using just two particles.

Hardy's paradox considers the experimental arrangement illustrated in figure 5.4 which consists of 2 overlapping Mach-Zehnder interferometers, one for an electron one for a positron. The electron interferometer is arranged so

that there is complete destructive interference along path d^- for an electron propagating on its own through the apparatus so that when only electrons pass through the apparatus they are all detected by detector C^- and there are no counts at detector D^- . Similarly, the positron interferometer is arranged so that there is complete destructive interference along path d^+ for a positron propagating on its own through the apparatus so that when only positrons pass through the apparatus they are all detected by detector C^+ and there are no counts at detector D^+ . When an electron and a positron are simultaneously at point P they annihilate each other with probability 1. When an electron and positron pass through the apparatus the annihilation at point P reduces the wave amplitudes along paths u^- and u^+ so that there is no longer complete destructive interference along paths d^- and d^+ . Thus, when an electron and a positron pass through the apparatus simultaneously it is possible that they can be detected at detectors D^- and D^+ . Hardy's paradox occurs when we try to associate an 'element of reality' to the paths of the particles and the paradox occurs only in the experiments in which the electron is detected by detector D^- AND the positron is detected by D^+ . From the arrangement of the interferometers we know that the electron could only reach detector D^- if the positron went along path u^+ . Similarly, we know that the positron could only reach detector D^+ if the electron went along path u^- . However, if we know that the positron went along path u^+ the electron went along path u^- they must have annihilated at point P and could not have reached the detectors. Thus our attempt to add an element of reality to the paths of the particles creates a paradox.

Hardy's paradox is very closely related to the subject of interaction-free measurement which is the basis of a question in the examples sheets. For an excellent discussion of these (and other related) subjects see Phys. Lett. A **301**, 130 (2002).

6

Entanglement

We have come across the concept of entanglement already in these lectures and you will remember that entanglement was crucial to the Einstein-Podolsky-Rosen paradox. For a long time entanglement was regarded as a bit of an embarrassment with, perhaps, the feeling that a complete theory of quantum mechanics might provide a more believable explanation of things like EPR correlations and thus remove some of the more disturbing consequences of entanglement. The rise of quantum information has led to a new view in which entanglement is treated as a resource that can be exploited to do useful things. Indeed, we shall see in section 10 how entanglement is exploited to perform teleportation. This new view of entanglement naturally leads to questions such as how entangled is a set of particles, how do I create entangled particles and can I transfer entanglement from one set of particles to another? We shall explore these questions in this section of the lectures. Many of these topics are covered in the book by Nielsen and Chuang.

6.1 Schmidt decomposition

There is, as yet, no general answer to the question ‘how entangled is a set of particles?’. However, in the case of just two particles, sometimes referred to as a bipartite system the answer to this question is known. The question can be answered by writing the Schmidt decomposition of the state of the two particle system. In general, for bipartite systems, the Schmidt decomposition allows any state to be written in the following form

$$|\psi\rangle = \sum_i \sqrt{\lambda_i} |i_A\rangle \otimes |i_B\rangle \quad (6.1)$$

where the λ_i (and hence also the $\sqrt{\lambda_i}$) are real, non-negative numbers and

the subscripts A and B indicate states of particles A and B respectively. The sum of λ_i is 1 due to the normalisation constraint. **BEWARE**, some sources use λ_i rather than $\sqrt{\lambda_i}$ for their definition of the Schmidt decomposition so you should take care to check the fundamental definition and convert results from one definition to the other if necessary.

If particles A and B have state spaces of the same dimension, it is easy to show that any arbitrary state of the two-particle system can be written in this form. A general state of the system can be written in terms of orthonormal bases $|i\rangle$ and $|k\rangle$ for systems A and B, respectively,

$$|\psi\rangle = \sum_{j,k} a_{j,k} |j\rangle \otimes |k\rangle \quad (6.2)$$

The coefficients $a_{i,j}$ can be treated as the elements of a matrix a . Note that a need not be Hermitian. It is known that any square matrix can be expressed in the form $a = udv$, where d is a diagonal matrix and v and u are unitary matrices. Thus

$$|\psi\rangle = \sum_{i,j,k} u_{j,i} d_{i,i} v_{i,k} |j\rangle \otimes |k\rangle \quad (6.3)$$

Defining

$$\begin{aligned} |i_A\rangle &= \sum_j u_{j,i} |j\rangle \\ |i_B\rangle &= \sum_k v_{i,k} |k\rangle \\ \sqrt{\lambda_i} &= d_{i,i} \end{aligned} \quad (6.4)$$

we see that this gives

$$|\psi\rangle = \sum_i \sqrt{\lambda_i} |i_A\rangle \otimes |i_B\rangle \quad (6.5)$$

or using the notation $|ii\rangle$ to represent $|i_A\rangle \otimes |i_B\rangle$ this state can also be written

$$|\psi\rangle = \sum_i \sqrt{\lambda_i} |ii\rangle \quad (6.6)$$

The Schmidt number of state $|\psi\rangle$ is the number of non-zero λ_i in these expressions and the bases $|i_A\rangle$ and $|i_B\rangle$ are called the Schmidt bases for A and B respectively. The state is entangled if more than one λ_i is non-zero. If one of the λ_i is nearly one and the others are all small or zero (provided that they not all zero) then the state is only weakly entangled. The state

becomes more entangled as all the non-zero become equal. The easiest case to quantify is the system of two spin 1/2 particles. In this case the Schmidt decomposition of the state can be written

$$|\psi\rangle = \cos(\theta) |\alpha_A\rangle \otimes |\alpha_B\rangle + \sin(\theta) |\beta_A\rangle \otimes |\beta_B\rangle \quad (6.7)$$

where $0 \leq \theta \leq \pi/4$. Note that the axes for the up and down spins for particles A and B will not, in general, be in the same direction. The degree of entanglement for this state is 0 when $\theta = 0$ (ie the particles are not entangled) and is maximum for $\theta = \pi/4$. This explains why we used the term maximally entangled to describe the singlet spin wavefunction when we discussed the EPR paradox previously.

There are a number of quantities that that can all be computed easily from the Schmidt decomposition. In many situations even though we know that a system is entangled we may only have access to one of the particles to perform measurements. The quantity that determines the outcomes of such measurements is the reduced density matrix. The reduced density matrix for particle A, ρ_A , is defined as follows

$$\rho_A = \text{Tr}_B |\psi(A, B)\rangle\langle\psi(A, B)| \quad (6.8)$$

where Tr_B is a partial trace over the states of system B. If we use the Schmidt basis then the reduced density matrix for particle A is given by

$$\rho_A = \text{Tr}_B |\psi(A, B)\rangle\langle\psi(A, B)| = \sum_i \lambda_i |i_A\rangle\langle i_A| \quad (6.9)$$

It can be seen that ρ_A is diagonal in the Schmidt basis for system A. Also the reduced density matrix for particle B, ρ_B , in the Schmidt basis for system B is identical to ρ_A . These reduced density matrices can be used to determine expectation values and measurement probabilities for particles A and B using the methods outlined in section 2 of the notes. We shall discuss this further in section 8.

The degree of entanglement of a pure state can be naturally parameterised by the entropy of the entanglement, E , which is defined as the von Neumann entropy of either ρ_A or ρ_B

$$E = -\text{Tr}_A \rho_A \log_2 \rho_A = -\text{Tr}_B \rho_B \log_2 \rho_B = -\sum \lambda_i \log_2 \lambda_i \quad (6.10)$$

Those of you attending Prof. Mackays's Information Theory course will recognise the above formula as the Shannon entropy of the Schmidt coefficients.

6.2 Effect of local operations on entanglement and the Bell States

How is the entanglement of a system of particles affected if we perform purely local operations on one of the particles? If the effect of the local operation can be written in terms of a unitary transformation U_A then the operation simply rotates the Schmidt basis $|i_A\rangle$ to $|i'_A\rangle$ where

$$|i'_A\rangle = U_A |i_A\rangle \quad (6.11)$$

The Schmidt decomposition of the transformed state $|\psi\rangle$ is

$$|\psi\rangle = \sum_i \sqrt{\lambda_i} |i'_A\rangle \otimes |i_B\rangle \quad (6.12)$$

Hence it can be seen that the entropy of entanglement and the degree of entanglement are not affected by such local operations on one particle. Clearly if any two entangled states have the same Schmidt coefficients they can be transformed into each other by using only local unitary operations. An example of a set of such states is the so-called Bell states.

The Bell states are a complete set of maximally entangled states for two spin-half particles (or any other particles that have two degrees of freedom). For generality I shall use the states $|0\rangle$ and $|1\rangle$ to represent the two basis states for each particle. From now on, I shall use the term qubit to describe any quantum system with two degrees of freedom. This term refers to a quantum bit - a system that can be in an arbitrary superposition of two states as opposed to the classical bit that can only be in one or other of two states. Also to adopt the terminology universally applied in quantum information Alice has access to particle A and Bob has access to particle B.

The Bell states are

$$\begin{aligned} |\psi^-\rangle &= \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle) \\ |\psi^+\rangle &= \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle) \\ |\phi^-\rangle &= \frac{1}{\sqrt{2}} (|00\rangle - |11\rangle) \\ |\phi^+\rangle &= \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \end{aligned} \quad (6.13)$$

where I have used the notation of the second form of the Schmidt decomposition given above.

All of the Bell states may be obtained by operating on only one of the qubits of the EPR state, which in this notation is the state $|\psi^-\rangle$. Even if

the particles in the EPR pair are spatially separated any of the Bell states may be generated by Alice performing local operations on particle A. The following unitary operations applied to particle A of $|\psi^-\rangle$ will generate the 4 Bell states given in Eqn. 6.13, their qubit mappings are also given for completeness.

$$\begin{aligned}
 U_{0,0} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & |0\rangle \rightarrow |0\rangle & |1\rangle \rightarrow |1\rangle \\
 U_{0,1} &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} & |0\rangle \rightarrow |0\rangle & |1\rangle \rightarrow -|1\rangle \\
 U_{1,0} &= \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} & |0\rangle \rightarrow -|1\rangle & |1\rangle \rightarrow -|0\rangle \\
 U_{1,1} &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} & |0\rangle \rightarrow |1\rangle & |1\rangle \rightarrow -|0\rangle
 \end{aligned} \tag{6.14}$$

Local operations cannot increase the degree of entanglement of a set of particles. However in the next section we shall see how we can distill a larger degree of entanglement for a small number of sets of particles if we have a larger number of less entangled particles. This is the process of entanglement concentration.

6.3 Entanglement concentration

As we shall see in section 10, teleportation requires a supply of maximally entangled EPR states. The question is, if we have a large number, N , of partially entangled pairs with entanglement ϵ is there a method for producing a set of maximally entangled EPR pairs from them. Given the constraints on the entropy of entanglement it is clear that we cannot produce more than $N\epsilon$ EPR pairs. You can read about techniques for achieving this with optimal efficiency in the limit when $N \rightarrow \infty$ in an article by Bennett *et al.* (Phys. Rev. A, **53**, p2046, (1996)).

There is a simple method that demonstrates entanglement concentration. It is referred to as the *Procrustean* method in Bennett *et al.*'s paper. It actually works rather efficiently and is only surpassed by other methods for large values of N . This is how it works. Assume that we have a source of partially entangled pairs. The state of each pair can be written in the Schmidt basis as

$$|\psi\rangle = \cos(\theta) |0_A\rangle \otimes |0_B\rangle + \sin(\theta) |1_A\rangle \otimes |1_B\rangle \tag{6.15}$$

where $0 < \theta < \pi/4$.

If particle A is passed through a polarisation dependent absorber (or a polarisation dependent deflector) so that a fraction $1 - \tan^2(\theta)$ of the $|0_A\rangle$ state is either absorbed (or deflected). Alice does nothing with the particles that are absorbed (or deflected) and simply tells Bob who discards his particles as well. As a result of the loss of the amplitude of the $|0_A\rangle$ state the unabsorbed (or undeflected) state is

$$|\psi'\rangle = \sqrt{2} \sin(\theta) [\cos(\pi/4) |0_A\rangle \otimes |0_B\rangle + \sin(\pi/4) |1_A\rangle \otimes |1_B\rangle] \quad (6.16)$$

ie the unabsorbed (or undeflected) state is now maximally entangled. The $\sqrt{2} \sin(\theta)$ factor simply reflects that fact that only a fraction of the originally partially entangled pairs become maximally entangled states as a result of this process.

6.4 Further examples of entanglement manipulation

Nielson has derived a general condition to determine whether one state may be transformed into another using only local operations and classical communication. The local operations can be both unitary transformations and non-unitary transformations such as measurements. Full details can be found in Phys. Rev. Lett. **83**, 436 (1999) and I shall only give the conclusions of his work here. Consider particles that have Hilbert spaces of dimension d . If we have the two-particle states $|\psi\rangle$ and $|\phi\rangle$ which have Schmidt decompositions

$$\begin{aligned} |\psi\rangle &= \sum_i \sqrt{\lambda_i} |i_A\rangle \otimes |i_B\rangle \\ |\phi\rangle &= \sum_i \sqrt{\lambda'_i} |i'_A\rangle \otimes |i'_B\rangle. \end{aligned} \quad (6.17)$$

If the Schmidt coefficients are ordered from the largest to the smallest so that λ_1 and λ'_1 are the largest and λ_d and λ'_d are the smallest then $|\psi\rangle$ can be transformed into $|\phi\rangle$ using local operations if and only if for each k in the range $1 \leq k \leq d$

$$\sum_{i=1}^k \lambda_i \leq \sum_{i=1}^k \lambda'_i. \quad (6.18)$$

Obviously the equality holds for $k = d$ due to the normalisation constraint. The mathematical term for the relationship defined by the above equation is majorisation and is indicated by the symbol \prec . Thus a succinct statement of the theorem is

$$|\psi\rangle \rightarrow |\phi\rangle \text{ iff } \lambda \prec \lambda'. \quad (6.19)$$

Clearly this criterion is more restrictive than the criterion that the transformation cannot increase the degree of entanglement.

It is quite possible that $|\psi\rangle$ cannot be transformed into $|\phi\rangle$ and that $|\phi\rangle$ cannot be transformed into $|\psi\rangle$ because neither set of Schmidt coefficients majorizes the other. In fact for large d if all the Schmidt coefficients are non-zero the probability of the set of Schmidt coefficients for one random state majorising the set for any other random state becomes negligible. Thus for large d the probability that any arbitrary entangled state can be converted into any other entangled state reduces to zero.

In a remarkable piece of work, Jonathan and Plenio have shown that for some states $|\psi\rangle$ and $|\phi\rangle$, where $|\psi\rangle$ cannot be transformed into $|\phi\rangle$ it is possible to perform the transformation $|\psi\rangle \rightarrow |\phi\rangle$ by borrowing entanglement, which is returned after the transformation has been completed. This process has similarities to the action of catalysts in chemistry and so, not surprisingly, this term has been applied to this process. This work is described in Phys. Rev. Lett. **83**, 3566 (1999).

6.5 Three and more particle entanglement

Quantifying the degree of entanglement for systems consisting of more than two particles is an intense area of research at present. It is easy to show (see question on Examples Sheet 2) that a general state of a three or more particle system cannot be written in the form of a Schmidt decomposition

$$|\psi\rangle = \sum_i \sqrt{\lambda_i} |i_A\rangle \otimes |i_B\rangle \otimes |i_C\rangle \dots \quad (6.20)$$

Acin *et al.* have shown that a general state of a three qubit system may be written in the following form

$$|\psi\rangle = \sqrt{\lambda_1} |000\rangle + \sqrt{\lambda_2} \exp(i\phi) |100\rangle + \sqrt{\lambda_3} |101\rangle + \sqrt{\lambda_4} |110\rangle + \sqrt{\lambda_5} |111\rangle \quad (6.21)$$

where the λ_i (and hence also the $\sqrt{\lambda_i}$) are real, non-negative numbers. The sum of λ_i is 1 due to the normalisation constraint and the phase ϕ can be chosen in the range $0 \leq \phi \leq \pi$. The phase could be associated with any of the coefficients. Note that this form differs from that given by Acin *et al.* in Phys. Rev. Lett. **85** 1560 (2000) because I have used a form that is consistent with my chosen definition of the Schmidt coefficients.

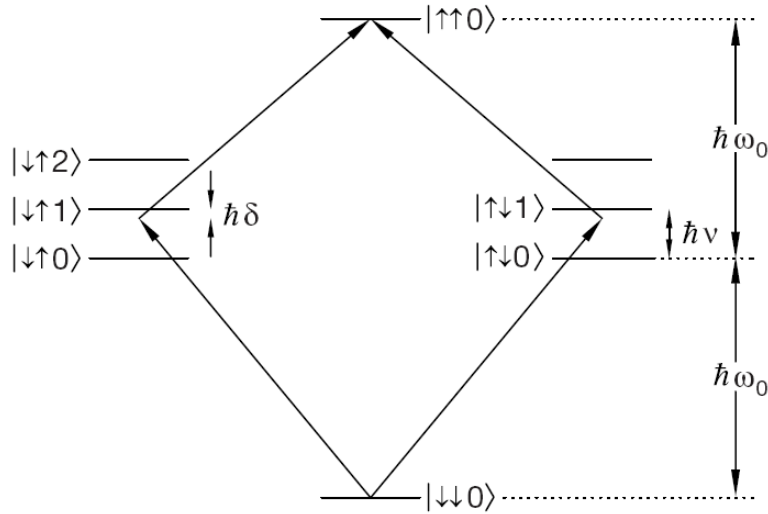


Fig. 6.1. Energy level diagram for push button entanglement for 2 ions in an ion trap. The up and down arrows represent the two internal energy levels of the ions and the numbers represent the number of center of mass excitations.

6.6 Push Button Entanglement

Earlier in the course I mentioned some of the difficulties involved in creating entangled pairs of particles. Waiting for a random process to produce a pair of entangled particles requires coincidence monitoring to infer that an entangled pair was measured and makes it hard to discriminate against noise. If we do not know when one of the particles from an entangled pair will arrive it is difficult to envisage ever implementing processes such as entanglement concentration or implementing efficient techniques to store entangled particles until they are required for teleportation or some other application. These problems become much more severe as the number of entangled particles increase. Sackett *et al.* have implemented a technique proposed by Mølmer and Sørensen that can, in principle, create entangled states of arbitrary large numbers of particles. The experimental details can be found in Nature **404** p256 (2000) (see also News and Views p231 of the same volume) and the theoretical technique can be found in Phys. Rev. Lett. **82**, 1835 (1999). The method exploits the phenomenon of Rabi oscillations but, by using the technique of detuning attempts to ensure that only N -particle spin excitations are generated rather than single-particle, 2-particle... and $(N - 1)$ -particle spin excitations. The energy level diagram for the 2 particle system used by Sackett *et al* is illustrated in figure 6.1

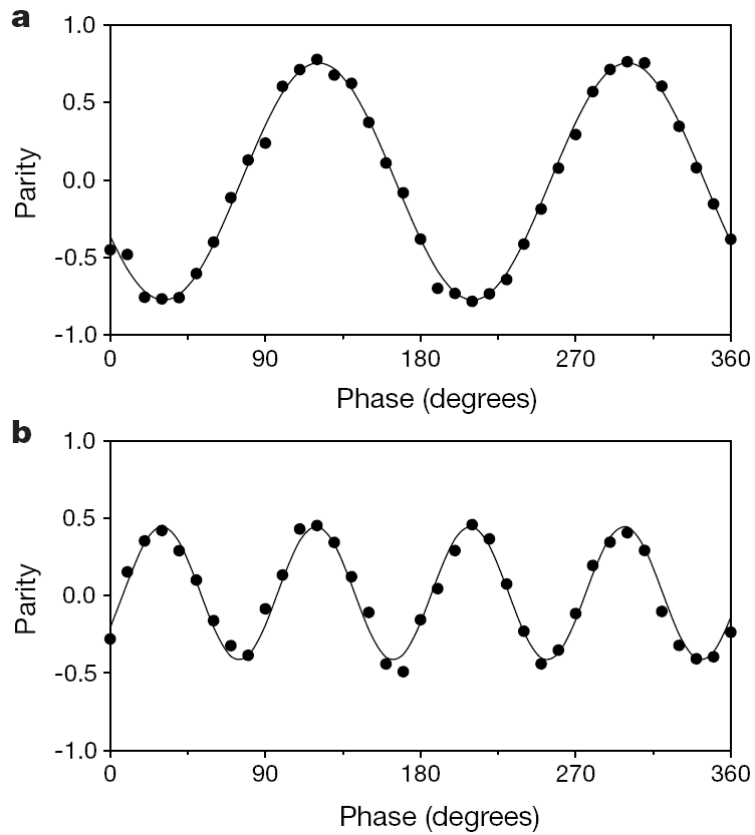


Fig. 6.2. Interference signals indicating the degree of entanglement for (a) system of 2 ions and (b) system of 4 ions.

Care has to be taken to ensure that the spin excitations do not become entangled with center-of-mass excitations as this would generate a diagonal reduced density matrix for the spin system. The implementation was successful in creating both 2-particle and 4-particle entanglement as can be seen from figure 6.2 but the degree of entanglement needs to be improved before this technique can be widely exploited.

7

Measurement 2

In most courses on quantum mechanics, only projective measurements are discussed. These are measurements of quantities corresponding to Hermitian operators and always result in reduction of the original state to one of the eigenstates of the operator corresponding to the measurement. However, there are more general forms of measurement in quantum mechanics and they are important for performing processes such as entanglement manipulation and entanglement concentration. In this section of the lectures we shall discuss two forms of measurement, so-called generalised measurements which correspond to positive operator-valued measures and weak measurements.

7.1 Generalised Measurements and Positive Operator-Valued Measure

In the general theory of measurement in quantum mechanics measurements are described by a set of measurement operators $\{M_m\}$ which are operators acting on the state space of the system being measured. The index m refers to the measurement outcomes that may occur in the experiment. If the state of the quantum system is $|\psi\rangle$ immediately before the measurement then the probability that the result m is measured is given by

$$p(m) = \langle\psi| M_m^\dagger M_m |\psi\rangle \quad (7.1)$$

and the state of the system after measurement is

$$\frac{M_m |\psi\rangle}{\sqrt{\langle\psi| M_m^\dagger M_m |\psi\rangle}} \quad (7.2)$$

The measurement operators satisfy the completeness relation

$$\sum_m M_m^\dagger M_m = 1 \quad (7.3)$$

which ensure that the probabilities sum to one

$$1 = \sum_m p(m) = \sum_m \langle \psi | M_m^\dagger M_m | \psi \rangle. \quad (7.4)$$

It is clear that projective measurements obey all these rules but in this general definition of measurement there is no need for all the projectors to be orthogonal. Any partition of unity by non-negative operators is called a ‘positive operator-valued measure’ or POVM. Thus generalised measurements in quantum mechanics correspond to POVMs.

We made the point previously that no projective measurement can distinguish non-orthogonal states $|\psi_i\rangle$ and $|\psi_j\rangle$. It is in the postulates of quantum mechanics. In lectures, I have made this explicit by considering the work of Zureck Phys. Rev. A **76** 052110 (2007). He proves this result as follows: he considers a system consisting of a measurement apparatus with a set of measurement states $\{|A_i\rangle\}$ and a quantum mechanical system that is spanned by a set of basis states $\{|\psi_i\rangle\}$. If initially, the joint system is in the separable state

$$|\phi_I\rangle = \left(\sum_i \alpha_i |\psi_i\rangle \right) |A_0\rangle, \quad (7.5)$$

and during the measurement process the state evolves to the entangled state

$$|\phi_F\rangle = \sum_i \alpha_i |\psi_i\rangle |A_i\rangle, \quad (7.6)$$

then, by considering the quantity

$$0 = \langle \phi_I | \phi_I \rangle - \langle \phi_F | \phi_F \rangle \quad (7.7)$$

it can be shown that in general

$$\langle \psi_i | \psi_j \rangle (1 - \langle A_i | A_j \rangle) = 0. \quad (7.8)$$

This expression indicates that unless $|\psi_i\rangle$ and $|\psi_j\rangle$ are orthogonal the measurement apparatus will not be able to distinguish between them. That is, if $\langle \psi_i | \psi_j \rangle \neq 0$ we must have $\langle A_i | A_j \rangle = 1$.

It is possible however to use a POVM to distinguish non-orthogonal states, at least part of the time. Consider the case of a spin 1/2 particle. Assume that we are randomly sent one of two states, either $|\psi_1\rangle$ which is a spin in the $+z$ direction or $|\psi_2\rangle$ which is a spin at 120° to the $+z$ axis in the

plane containing the x and z axes. No single orthogonal measurement can tell us whether then spin we just received was either $|\psi_1\rangle$ or $|\psi_2\rangle$. A single orthogonal measurement of the spin in the z -direction would tell us that whenever we measured a $-z$ spin that the state must have been $|\psi_2\rangle$. If we want to be able to assign both spin states at least some of the time then we can use a POVM. To do this we use a POVM consisting of three spin operators, the first in the $-z$ direction and the other two at $\pm 60^\circ$ to the $+z$ direction in the plane containing the x and z axes. If we measure the spin to be pointing in the $-z$ direction we know the state was $|\psi_2\rangle$. If we measure the spin to be pointing in the direction -60° to the $+z$ direction we know the state was $|\psi_1\rangle$. If we measure the spin to be pointing in the direction 60° to the $+z$ direction we do not know whether the state was $|\psi_1\rangle$ or $|\psi_2\rangle$.

The formal way approach to the idea of POVMs is to consider the Hilbert space of the quantum system we have access to, H_A , is part of a larger Hilbert space, H , so that

$$H = H_A \oplus H_A^\perp \quad (7.9)$$

Our observers who live in H_A have access only to observables M with corresponding operators M_A that lie with the Hilbert space H_A such that

$$M_A |\psi^\perp\rangle = 0 \quad (7.10)$$

The easiest way of understanding the implications of the restriction of the observers to H_A is to think of applying a measurement in the entire space H . From the postulates of quantum mechanics we know that the measurement will yield one of the eigenstates of the corresponding operator, which will have a unique decomposition in H_A and H_A^\perp , thus the eigenstates can be written

$$|u_a\rangle = |\tilde{\psi}_a\rangle + |\tilde{\psi}_a^\perp\rangle \quad (7.11)$$

where $|\tilde{\psi}_a\rangle$ and $|\tilde{\psi}_a^\perp\rangle$ are unnormalised vectors in H_A and H_A^\perp , respectively. Consider a situation where the initial density matrix of the system only occupies the accessible space so that it can be written ρ_A . After the measurement, the density matrix of the entire system will be $|u_a\rangle\langle u_a|$ with probability $\langle u_a|\rho_A|u_a\rangle$ but this probability is equal to $\langle \tilde{\psi}_a|\rho_A|\tilde{\psi}_a\rangle$ since ρ_A only exists in H_A . However, as seen by the observers who live in H_A , who know nothing about H_A^\perp , there is no distinction between $|u_a\rangle$ and $|\tilde{\psi}_a\rangle$ apart from the normalisation. If we write $|\tilde{\psi}_a\rangle = \sqrt{\lambda_a}|\psi_a\rangle$, where $|\psi_a\rangle$ is a normalised state, then for an observer limited to observations in H_A we

might as well say that the outcome of the measurement is the density matrix $|\psi_a\rangle\langle\psi_a|$ with probability $\langle\tilde{\psi}_a|\rho_A|\tilde{\psi}_a\rangle$.

Now consider the operators

$$F_a = |\tilde{\psi}_a\rangle\langle\tilde{\psi}_a| = \lambda_a |\psi_a\rangle\langle\psi_a| \quad (7.12)$$

It is clear that each F_a is Hermitian and non-negative. Furthermore

$$\sum_a F_a = 1 \quad (7.13)$$

Thus our sum of projectors is a POVM. These results show that measurements on part of a quantum system correspond to POVMs. Neumark's theorem proves that any POVM can be realised by extending the Hilbert space of the system and performing orthogonal measurements in the larger space. Thus, we can prepare non-orthogonal states in the space H_A using a single measurement, albeit in a larger Hilbert space. (Remember that any projective measurement performed purely in the space of H_A will always yield orthogonal states). Even more importantly, specific states in the space of can be prepared by a measurement performed purely in the space H_A^\perp . This may seem surprising but we have come across this concept already. This is precisely what happens in the EPR paradox – a specific state of particle 2 is produced as a result of a measurement performed on particle 1.

7.2 Implementation of POVMs

Despite the frequent mention of POVMs in the field of quantum information until recently the only experimental realisation of a POVM was the example of measurement performed on one particle of an entangled pair. Recently, Ahnert and Payne have devised methods for performing POVMs on a single particle with unit probability of success. The first implementation (Phys. Rev. A 69, 012312 (2004)) was for a specific POVM – the POVM consisting of the three polarisation measurement operators oriented at 120° with respect to each other which was mentioned above. In more recent work (Phys. Rev. A 71, 012330 (2005)) this has been generalised to any arbitrary POVM which can be performed using the module shown in figure 7.1 which implements one specific element of the POVM. By arranging a sequence of such modules any arbitrary POVM can be implemented. One such example, again for the case of three polarisation measurement operators oriented at 120° with respect to each other, is shown in figure 7.2 but it should be remembered that this technique can be used for an arbitrary number of measurement operators.

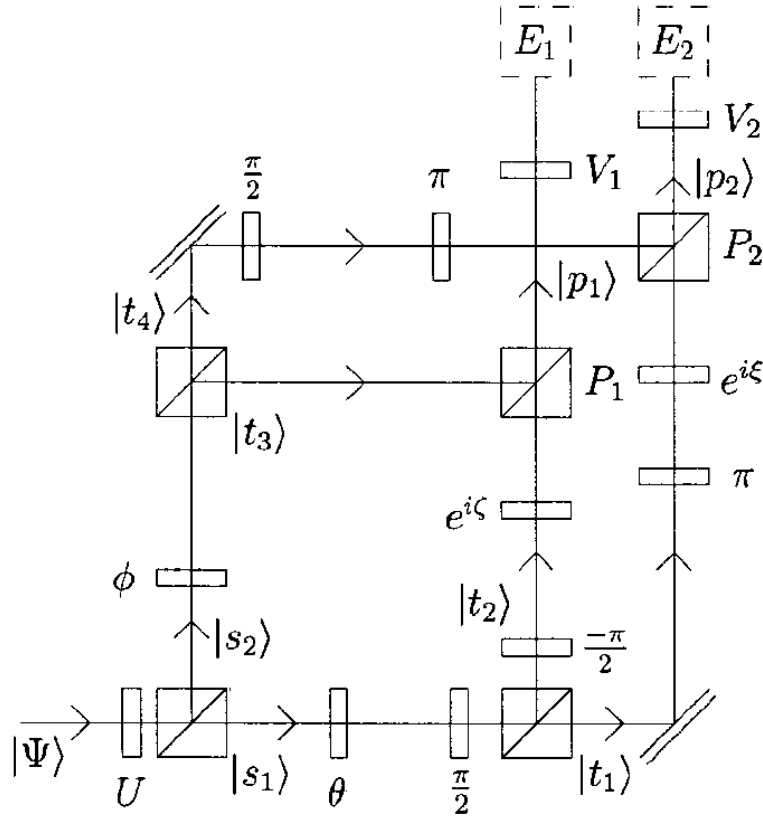


Fig. 7.1. Implementation of one measurement operator, E_1 , of a POVM. The squares with diagonal lines are polarizing beam splitters, the angles $\theta, \phi, \pi/2, \pi$ indicate polarization rotations by these angles, U, V_1, V_2 are unitary operators and $e^{i\xi}, e^{i\zeta}$ represent phase shifts.

7.3 Weak Measurements

The concept of weak measurements was introduced in a paper by Aharonov *et al.* (Phys. Rev. Lett. **60**, 1351 (1988)). They considered the example of the weak measurement of the spin of a particle. Most people thought that the idea was nonsensical but remarkably some of the most vociferous opponents proceeded to convince themselves that the effect was real (Duck *et al.* Phys. Rev. D. **40**, 2112 (1989)) and the effect has been observed experimentally (Ritchie *et al.* Phys. Rev. Lett. **66**, 1107 (1991)). This example which is shown in figure 7.3 will form the basis of our discussion of weak measurement.

The simplest starting point for considering weak measurements is to apply von Neumann's model of orthogonal measurement by choosing the coupling

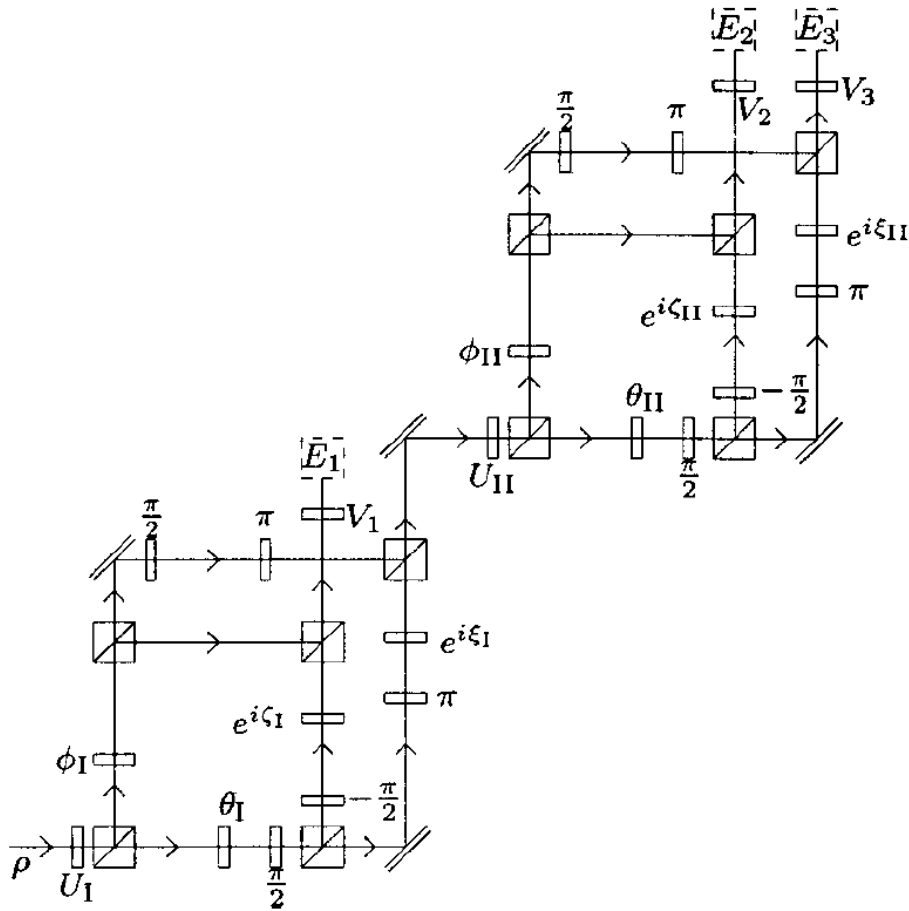


Fig. 7.2. Implementation of the POVM for three polarisation measurement operators oriented at 120° with respect to each other.

strength between the quantum system and the measurement instrument or the time that the coupling acts over to be small enough that the displacement of the position of the pointer due to the measurement is much smaller than the width of the pointer wavepacket. If you remember our discussion of von Neumann's model in section 3 of the lectures, we chose parameters such as the mass of the pointer so that the width of the pointer wavepacket was much smaller than the different displacements of the pointer associated with each of the eigenvalues of the measurement. At first glance it might be thought that the opposite regime where the width of the pointer is much larger than the displacements is a worthless form of measurement as it would seem unlikely that it could provide any useful information at all. However, as

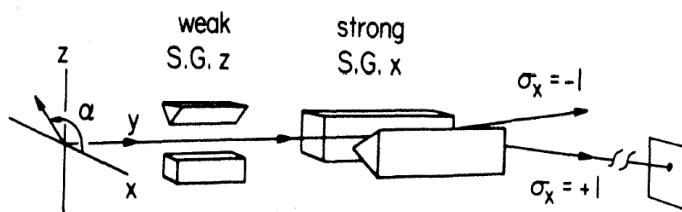


Fig. 7.3. Experimental arrangement for the measurement of the spin of spin-half particle. Particles in a specific spin state pass first through a Stern-Gerlach experiment with a weak magnetic field in the z -direction and then through a second Stern-Gerlach experiment with a strong magnetic field in the x -direction.

the following example shows, these so-called weak measurements can result in measured values that are much larger than any of the eigenvalues of the operator corresponding to the measurement.

To generate a weak measurement we assume that the quantum system has been generated in a specific initial state. We shall consider a spin-half particle and take the initial state to be a spin pointing at an angle α to the x axis in the $x-z$ plane. The initial spin state is thus the $+1$ eigenstate of $(\cos \alpha) \sigma_x + (\sin \alpha) \sigma_z$,

$$|\phi_{in}\rangle = \frac{1}{\sqrt{2}} \left[\cos \frac{\alpha}{2} + \sin \frac{\alpha}{2} \right] |\uparrow\rangle + \frac{1}{\sqrt{2}} \left[\cos \frac{\alpha}{2} - \sin \frac{\alpha}{2} \right] |\downarrow\rangle \quad (7.14)$$

and in contrast to the case of the von Neumann model the 'pointer' wave function will be a broad Gaussian

$$\psi(z) \propto \exp\left(-\frac{z^2}{4\Delta^2}\right) \quad (7.15)$$

or equivalently in the momentum representation

$$\psi(p) \propto \exp(-\Delta^2 p^2) \quad (7.16)$$

where the momentum is in the z -direction. In fact in this example of the measurement of spin using a Stern-Gerlach apparatus, the pointer is actually the momentum of the quantum particles themselves. As discussed previously in section 3, in this experiment the position the particles end up on a screen provides the measurement of the spin of the particles. To perform the weak

measurement, we first pass the particle through a Stern-Gerlach apparatus oriented in the z -direction so, as discussed in section 3, the Hamiltonian describing the motion of the particles in the inhomogenous magnetic field is

$$H = -\lambda\mu\sigma_z z \quad (7.17)$$

As discussed previously the fact that this interaction Hamiltonian contains the position of the particle z implies that the momentum of the particles in the z -direction will be changed as a result of the interaction. Following the arguments of section 3 after the particle has passed through the magnetic field the state of the particles will be

$$\frac{1}{\sqrt{2}} \left[\cos \frac{\alpha}{2} + \sin \frac{\alpha}{2} \right] |\uparrow\rangle \otimes |\psi(p - \lambda\mu t)\rangle + \frac{1}{\sqrt{2}} \left[\cos \frac{\alpha}{2} - \sin \frac{\alpha}{2} \right] |\downarrow\rangle \otimes |\psi(p + \lambda\mu t)\rangle \quad (7.18)$$

In weak measurement we consider the limit where $\lambda\mu t \ll \Delta p$. In this case any single measurement gives very little information about the value of $\langle\sigma_z\rangle$ although repeating the measurements on a large number of particles would allow this expectation value to be determined. However, interesting things happen when we apply a postselection to the quantum state in addition to the preselection which chose the initial state shown above. If we postselect the state with spin pointing along the x -direction by performing a strong measurement and selecting the $\sigma_x = +1$ state we project the quantum state onto the final state

$$|\phi_{fin}\rangle = |\leftarrow\rangle = \frac{1}{\sqrt{2}} [|\uparrow\rangle + |\downarrow\rangle] \quad (7.19)$$

From the rules of measurement the wave function of the system after the postselection is

$$\frac{1}{\sqrt{2}} \left[\left[\cos \frac{\alpha}{2} + \sin \frac{\alpha}{2} \right] \exp(-\Delta^2(p - \lambda\mu t)^2) + \left[\cos \frac{\alpha}{2} - \sin \frac{\alpha}{2} \right] \exp(-\Delta^2(p + \lambda\mu t)^2) \right] |\leftarrow\rangle \quad (7.20)$$

If we consider the case when the angle $\alpha = \pi - 2\epsilon$ with $\epsilon \ll 1$, this expression reduces to

$$\frac{1}{\sqrt{2}} \left[(1 + \epsilon) \exp(-\Delta^2(p - \lambda\mu t)^2) - (1 - \epsilon) \exp(-\Delta^2(p + \lambda\mu t)^2) \right] |\leftarrow\rangle \quad (7.21)$$

We should first remember that changes of the momentum of the quantum particles in the z -direction of $\pm\lambda\mu t$ represent measurements of spin (measured in terms of the eigenvalues of the σ_z Pauli spin operator) of

± 1 . However, in the above equation there is a cancellation between the peaks of the two Gaussians centred at these values so that the momentum distribution instead has peaks roughly at the values $\pm \lambda \mu t / \epsilon$. Hence, this measurement yields values for the spin of the particle that are of the order of $1/\epsilon$, which with suitable choice of ϵ , can be arbitrarily large! In a more detailed model, it can be seen that there are limits to how large the values this measurement produces can be. However, the basic correctness of the concept has been demonstrated in the experiments of Ritchie *et al.* (Phys. Rev. Lett. **66**, 1107 (1991)) shown in figure 7.4.

The field of weak measurements is very much in its infancy. It is now accepted that these measurements can generate apparently unphysical results, such as the unphysical value of the spin of an electron in the example above. Another example is being able to measure a negative kinetic energy (Aharonov *et al.* Phys. Rev. A **48**, 4084 (1993)). Advocates of weak measurement believe that this approach may provide the key to interrogating a quantum state without destroying it and thus possibly measuring a variety of properties, even those corresponding to non-commuting operators, without destroying the state.

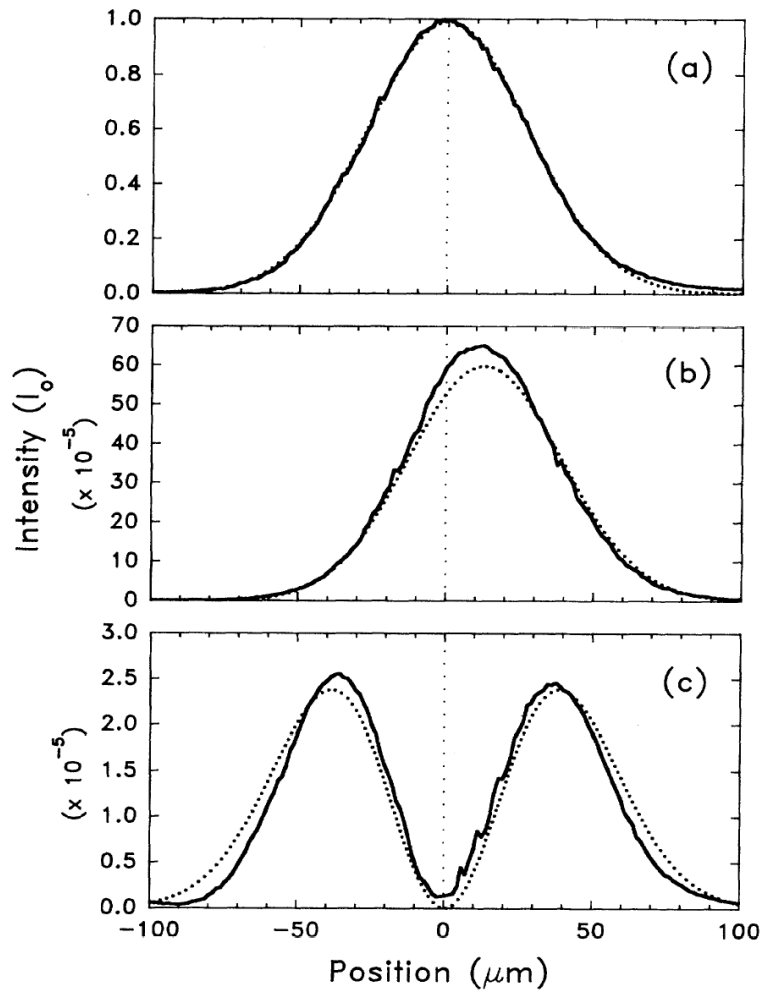


Fig. 7.4. Results of weak measurement of the spin of a particle. (a) Initial state polarised along the x -axis. (b) Initial state polarised at an angle of 2.2×10^{-2} to the z -axis. (c) Initial state polarised along the z -axis (for which the theoretical value of the weak spin measurement would be infinite)

8

Decoherence

Decoherence is a universal property of any quantum system coupled to a large number of other quantum mechanical degrees of freedom. Basically, the process causes the off-diagonal elements of the density matrix to fall rapidly to zero. Hence, the coherence between the different states of the quantum system disappears and it is no longer possible to produce interference effects between these states. This occurs even though the overall evolution of the entire system is unitary. The point is that the coherence now involves all the additional quantum-mechanical degrees of freedom and if we only look at the quantum system the coherence has been lost. Before we can really address the question of decoherence we need to develop our understanding of how to deal with quantum-mechanical systems where an observer only has access to part of the entire system. We shall see that we have already discussed some of the consequences of this separation between accessible and inaccessible parts of a quantum mechanical system in previous lectures. We begin with a discussion of the measurement problem from Zureck *Physics Today* (1991). A modified and updated version can be found on the web if you search for Zureck *Los Alamos Science* **27** 1-25 (2002).

8.1 Measurement and Decoherence

The text in this and the next section is taken from Zureck *Physics Today* (1991). The full article is well worth reading.

A convenient starting point for the discussion of the measurement problem and, more generally, of the emergence of classical behavior from quantum dynamics is the analysis of quantum measurements due to John von Neumann (1932). In contrast to Bohr, who assumed at the outset that measurement apparatus must be classical (thereby forfeiting claims that quantum theory is universal), von Neumann analyzed the case of a quantum apparatus. Here



Fig. 8.1. Zurek from his *Physics Today* article.

we reproduce his analysis for the simplest case: a measurement on a two-state system S (which can be thought of as an atom with spin $1/2$) in which a quantum two-state (one bit) detector records the result.

The Hilbert space H_S of the system is spanned by the orthonormal states $|\uparrow\rangle$ and $|\downarrow\rangle$, while the states $|d_\uparrow\rangle$ and $|d_\downarrow\rangle$ span the H_D of the detector Hilbert space. A two-dimensional H_D is the absolute minimum needed to record the possible outcomes. One can devise a quantum detector that *clicks* only when the spin is in the state $|\uparrow\rangle$, that is,

$$|\uparrow\rangle |d_\downarrow\rangle \rightarrow |\uparrow\rangle |d_\uparrow\rangle \quad (8.1)$$

and remains unperturbed otherwise (Zeh 1970, Wigner 1963, Scully *et al.* 1989).

We assume that, before the interaction, the system was in a pure state $|\uparrow\rangle$ given by

$$|\Psi_S\rangle = \alpha |\uparrow\rangle + \beta |\downarrow\rangle \quad (8.2)$$

with the complex coefficients satisfying $|\alpha|^2 + |\beta|^2 = 1$. The composite

system starts as

$$|\Phi^i\rangle = |\Psi_S\rangle |d_\downarrow\rangle \quad (8.3)$$

Interaction results in the evolution of $|\Phi^i\rangle$ into a correlated state $|\Phi^C\rangle$:

$$|\Phi^i\rangle = (\alpha |\uparrow\rangle + \beta |\downarrow\rangle) |d_\downarrow\rangle \rightarrow \alpha |\uparrow\rangle |d_\uparrow\rangle + \beta |\downarrow\rangle |d_\downarrow\rangle = |\Phi^C\rangle \quad (8.4)$$

This essential and uncontroversial first stage of the measurement process can be accomplished by means of a Schrödinger equation with an appropriate interaction as we have seen earlier in the lecture course. It might be tempting to halt the discussion of measurements with eqn. 8.4. After all, the correlated state vector $|\Phi^C\rangle$ implies that, if the detector is seen in the state $|d_\uparrow\rangle$, the system is guaranteed to be found in the state $|\uparrow\rangle$. Why ask for anything more?

The reason for dissatisfaction with $|\Phi^C\rangle$ as a description of a completed measurement is simple and fundamental: In the real world, even when we do not know the outcome of a measurement, we do know the possible alternatives, and we can safely act as if only one of those alternatives has occurred. Such an assumption is not only unsafe but also simply wrong for a system described by $|\Phi^C\rangle$.

How then can an observer (who has not yet consulted the detector) express our ignorance about the outcome without giving up his certainty about the ‘menu’ of possibilities? Quantum theory provides the right formal tool for the occasion: A density matrix can be used to describe the probability distribution over the alternative outcomes.

Von Neumann was well aware of these difficulties. Indeed, he postulated (1932) that, in addition to the unitary evolution given by the Schrödinger equation, there should be an ad hoc ‘process 1’ a nonunitary reduction of the state vector that would take the pure, correlated state $|\Phi^C\rangle$ into an appropriate mixture: This process makes the outcomes independent of one another by taking the pure-state density matrix:

$$\begin{aligned} \rho^C &= |\Phi^C\rangle \langle \Phi^C| \\ &= |\alpha|^2 |\uparrow\rangle |d_\uparrow\rangle \langle d_\uparrow| \langle \uparrow| + |\beta|^2 |\downarrow\rangle |d_\downarrow\rangle \langle d_\downarrow| \langle \downarrow| \\ &\quad + \alpha\beta^* |\uparrow\rangle |d_\uparrow\rangle \langle d_\downarrow| \langle \downarrow| + \alpha^*\beta |\downarrow\rangle |d_\downarrow\rangle \langle d_\uparrow| \langle \uparrow| \end{aligned} \quad (8.5)$$

and canceling the off-diagonal terms that express purely quantum correlations (entanglement) so that the reduced density matrix with only classical correlations emerges:

$$\rho^R = |\alpha|^2 |\uparrow\rangle |d_\uparrow\rangle \langle d_\uparrow| \langle \uparrow| + |\beta|^2 |\downarrow\rangle |d_\downarrow\rangle \langle d_\downarrow| \langle \downarrow|$$

(8.6)

Why is the reduced density matrix ρ^R easier to interpret as a description of a completed measurement than ρ^C ? After all, both ρ^C and ρ^R contain identical diagonal elements. Therefore, both outcomes are still potentially present. So what –if anything– was gained at the substantial price of introducing a nonunitary process 1?

The key advantage of ρ^R over ρ^C is that its coefficients may be interpreted as classical probabilities. The density matrix ρ^R can be used to describe the alternative states of a composite spin-detector system that has classical correlations. Von Neumann's process 1 serves a similar purpose to Bohr's 'border' even though process 1 leaves all the alternatives in place. When the off-diagonal terms are absent, one can nevertheless safely maintain that the apparatus, as well as the system, is each separately in a definite but unknown state, and that the correlation between them still exists in the preferred basis defined by the states appearing on the diagonal. By the same token, the identities of two halves of a split coin placed in two sealed envelopes may be unknown but are classically correlated. Holding one unopened envelope, we can be sure that the half it contains is either heads or tails (and not some superposition of the two) and that the second envelope contains the matching alternative.

By contrast, it is impossible to interpret ρ^C as representing such classical ignorance. In particular, even the set of the alternative outcomes is not decided by ρ^C . This circumstance can be illustrated in a dramatic fashion by choosing $\alpha = -\beta = 1/\sqrt{2}$ so that the density matrix ρ^C is a projection operator constructed from the correlated state

$$|\Phi^C\rangle = \frac{1}{\sqrt{2}} [|\uparrow\rangle |d_\uparrow\rangle - |\downarrow\rangle |d_\downarrow\rangle] \quad (8.7)$$

This state is invariant under the rotations of the basis. For instance, instead of the eigenstates $|\uparrow\rangle$ and $|\downarrow\rangle$ of σ_z one can rewrite $|\Phi^C\rangle$ in terms of the eigenstates of σ_x :

$$\begin{aligned} |\leftarrow\rangle &= \frac{1}{\sqrt{2}} [|\uparrow\rangle + |\downarrow\rangle] \\ |\rightarrow\rangle &= \frac{1}{\sqrt{2}} [|\uparrow\rangle - |\downarrow\rangle] \end{aligned} \quad (8.8)$$

This representation immediately yields

$$|\Phi^C\rangle = \frac{1}{\sqrt{2}} [|\leftarrow\rangle |d_{\leftarrow}\rangle + |\rightarrow\rangle |d_{\rightarrow}\rangle] \quad (8.9)$$

where

$$\begin{aligned} |d_{\leftarrow}\rangle &= \frac{1}{\sqrt{2}} [|\uparrow\rangle + |\uparrow\rangle] \\ |d_{\rightarrow}\rangle &= \frac{1}{\sqrt{2}} [|\uparrow\rangle - |\uparrow\rangle] \end{aligned} \tag{8.10}$$

are, as a consequence of the superposition principle, perfectly legal states in the Hilbert space of the quantum detector. Therefore, the density matrix

$$\rho^C = |\Phi^C\rangle\langle\Phi^C|$$

could have many (in fact, infinitely many) different states of the subsystems on the diagonal.

This freedom to choose a basis should not come as a surprise. Except for the notation, the state vector $|\Phi^C\rangle$ is the same as the wave function of a pair of maximally correlated (or entangled) spin-1/2 systems in David Bohm's version (1951) of the Einstein-Podolsky-Rosen (EPR) paradox (Einstein *et al.* 1935). And the experiments that show that such nonseparable quantum correlations violate Bell's inequalities (Bell 1964) are demonstrating the following key point: The states of the two spins in a system described by $|\Phi^C\rangle$ are not just unknown, but rather they cannot exist before the 'real' measurement (Aspect *et al.* 1981, 1982). We conclude that when a detector is quantum, a superposition of records exists and is a record of a superposition of outcomes a very nonclassical state of affairs.

Unitary evolution condemns every closed quantum system to purity. Yet, if the outcomes of a measurement are to become independent events, with consequences that can be explored separately, a way must be found to dispose of the excess information and thereby allow any orthogonal basis—any potential events and their superpositions—to be equally correlated. We have just analyzed quantum correlation from the point of view of its role in acquiring information. Here, I shall discuss the flip side of the story: Quantum correlations can also disperse information throughout the degrees of freedom that are, in effect, inaccessible to the observer. Interaction with the degrees of freedom external to the system—which we shall summarily refer to as the environment—offers such a possibility.

Reduction of the state vector, $\rho^C \rightarrow \rho^R$, decreases the information available to the observer about the composite system SD . The information loss is needed if the outcomes are to become classical and thereby available as initial conditions to predict the future. The effect of this loss is to increase

the entropy $E = -\text{Tr}\rho \ln \rho$ by an amount

$$\Delta E = E(\rho^R) - E(\rho^C) = -(|\alpha|^2 \ln |\alpha|^2 + |\beta|^2 \ln |\beta|^2) \quad (8.11)$$

Entropy must increase because the initial state described by ρ^C was pure, $E(\rho^C) = 0$, and the reduced state is mixed. Information gain –the objective of the measurement– is accomplished only when the observer interacts and becomes correlated with the detector in the already precollapsed state ρ^R .

To illustrate the process of the environment-induced decoherence, consider a system S , a detector D , and an environment E . The environment is also a quantum system. Following the first step of the measurement process establishment of a correlation as shown in eqn. 8.4 the environment similarly interacts and becomes correlated with the apparatus:

$$\begin{aligned} |\Phi^C\rangle |E_0\rangle &= (\alpha |\uparrow\rangle |d_\uparrow\rangle + \beta |\downarrow\rangle |d_\downarrow\rangle) |E_0\rangle \\ &\rightarrow \alpha |\uparrow\rangle |d_\uparrow\rangle |E_\uparrow\rangle + \beta |\downarrow\rangle |d_\downarrow\rangle |E_\downarrow\rangle \\ &= |\Psi\rangle \end{aligned} \quad (8.12)$$

The final state of the combined SDE ‘von Neumann chain’ of correlated systems extends the correlation beyond the SD pair. When the states of the environment $|E_i\rangle$, corresponding to the states $|d_\uparrow\rangle$ and $|d_\downarrow\rangle$ of the detector are orthogonal, $\langle E_i | E_j \rangle = \delta_{i,j}$, the density matrix for the detector-system combination is obtained by ignoring (tracing over) the information in the uncontrolled (and unknown) degrees of freedom

$$\begin{aligned} \rho_{SD} &= \text{Tr} |\Psi\rangle \langle \Psi| \\ &= \sum_i \langle E_i | \Psi \rangle \langle \Psi | E_i \rangle \\ &= |\alpha|^2 |\uparrow\rangle |d_\uparrow\rangle \langle d_\uparrow| \langle \uparrow| + |\beta|^2 |\downarrow\rangle |d_\downarrow\rangle \langle d_\downarrow| \langle \downarrow| \\ &= \rho^R \end{aligned} \quad (8.13)$$

The resulting ρ^R is precisely the reduced density matrix that von Neumann called for. Now, a superposition of the records of the detector states is no longer a record of a superposition of the state of the system. A preferred basis of the detector, sometimes called the ‘pointer basis’ for obvious reasons (Zurek Phys. Rev. D **24** 1516 (1981)), has emerged. Moreover, we have obtained it –or so it appears– without having to appeal to von Neumann’s nonunitary process 1 or anything else beyond the ordinary, unitary Schrödinger evolution. The preferred basis of the detector –or for that matter, of any open quantum system is selected by the dynamics.

Not all aspects of this process are completely clear. It is, however, certain that the detector-environment interaction Hamiltonian plays a decisive

role. In particular, when the interaction with the environment dominates, eigenspaces of any observable Λ that commutes with the interaction Hamiltonian,

$$[\Lambda, H_{int}] = 0, \quad (8.14)$$

invariably end up on the diagonal of the reduced density matrix (Zurek 1981, 1982). This commutation relation has a simple physical implication: It guarantees that the pointer observable Λ will be a constant of motion, a conserved quantity under the evolution generated by the interaction Hamiltonian. Thus, when a system is in an eigenstate of Λ , interaction with the environment will leave it unperturbed.

In the real world, the spreading of quantum correlations is practically inevitable. For example, when in the course of measuring the state of a spin-1/2 atom, if a photon had scattered from the atom while it was traveling along one of its two alternative routes, this interaction would have resulted in a correlation with the environment and would have necessarily led to a loss of quantum coherence. The density matrix of the SD pair would have lost its off-diagonal terms. Moreover, given that it is impossible to catch up with the photon, such loss of coherence would have been irreversible. Irreversibility could also arise from more familiar, statistical causes: Environments are notorious for having large numbers of interacting degrees of freedom, making extraction of lost information as difficult as reversing trajectories in the Boltzmann gas.

8.2 Decoherence: How Long does it take?

A tractable model of the environment is afforded by a collection of harmonic oscillators (Feynman and Vernon 1963, Dekker 1981, Caldeira and Leggett 1983a, 1983b, 1985, Joos and Zeh 1985, Paz et al. 1993) or, equivalently, by a quantum field (Unruh and Zurek 1989). If a particle is present, excitations of the field will scatter off the particle. The resulting ripples will constitute a record of its position, shape, orientation, and so on, and most important, its instantaneous location and hence its trajectory.

A boat traveling on a quiet lake or a stone that fell into water will leave such an imprint on the water surface. Our eyesight relies on the perturbation left by the objects on the pre-existing state of the electromagnetic field. Hence, it is hardly surprising that an imprint is left whenever two quantum systems interact, even when ‘nobody is looking,’ and even when the lake is stormy and full of pre-existing waves, and the field is full of excitations –that is, when the environment starts in equilibrium at some finite temperature.

‘Messy’ initial states of the environment make it difficult to decipher the record, but do not preclude its existence. A specific example of decoherence—a particle at position x interacting with a scalar field ϕ (which can be regarded as a collection of harmonic oscillators) through the Hamiltonian

$$H_{int} = \epsilon x \frac{d\phi}{dt} \quad (8.15)$$

has been extensively studied by many, including the investigators just referenced. The conclusion is easily formulated in the so-called ‘high-temperature limit’, in which only thermal-excitation effects of the field ϕ are taken into account and the effect of zeropoint vacuum fluctuations is neglected. In this case, the density matrix $\rho(x, x')$ of the particle in the position representation evolves according to the master equation

$$\frac{d\rho}{dt} = \frac{1}{i\hbar} [H, \rho] - \gamma(x - x') \left(\frac{\partial}{\partial x} - \frac{\partial}{\partial x'} \right) \rho - \frac{2m\gamma k_B T}{\hbar^2} (x - x')^2 \rho \quad (8.16)$$

where H is the particle’s Hamiltonian (although with the potential $V(x)$ adjusted because of H_{int}), γ is the relaxation rate, k_B is the Boltzmann constant, and T is the temperature of the field. Equation 8.16 is obtained by first solving the Schrödinger equation exactly for a particle plus the field and then tracing over the degrees of freedom of the field. It naturally separates into three distinct terms, each of them responsible for a different aspect of the effectively classical behavior. The first term—the von Neumann equation (which can be derived from the Schrödinger equation)—generates reversible classical evolution of the expectation value of any observable that has a classical counterpart regardless of the form of ρ (Ehrenfest’s theorem). The second term causes dissipation. The relaxation rate $\gamma = \eta/2m$ is proportional to the viscosity $\epsilon^2/2$ due to the interaction with the scalar field. That interaction causes a decrease in the average momentum and loss of energy. The last term also has a classical counterpart: It is responsible for fluctuations or random ‘kicks’ that lead to Brownian motion.

For our purposes, the effect of the last term on quantum superpositions is of greatest interest. It destroys quantum coherence, eliminating off-diagonal terms responsible for quantum correlations between spatially separated pieces of the wave packet. It is therefore responsible for the classical structure of the phase space, as it converts superpositions into mixtures of localized wave packets which, in the classical limit, turn into the familiar points in phase space. This effect is best illustrated by an example. Consider the ‘cat’ state shown in fig. 8.2, where the wave function of a particle is given by a coherent superposition of two Gaussians: $\phi(x) = (\chi^+(x) + \chi^-(x))/\sqrt{2}$

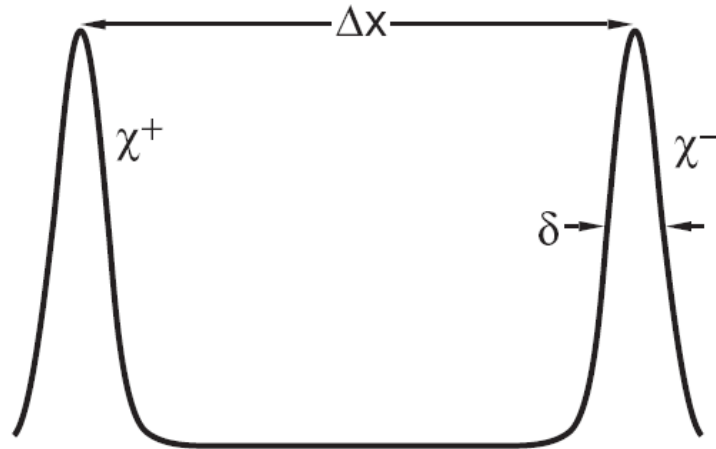


Fig. 8.2. A ‘Schrödinger Cat’ State or a Coherent Superposition This cat state $\phi(x)$, the coherent superposition of two Gaussian wave packets, could describe a particle in a superposition of locations inside a Stern-Gerlach apparatus or the state that develops in the course of a double-slit experiment. The phase between the two components has been chosen to be zero.

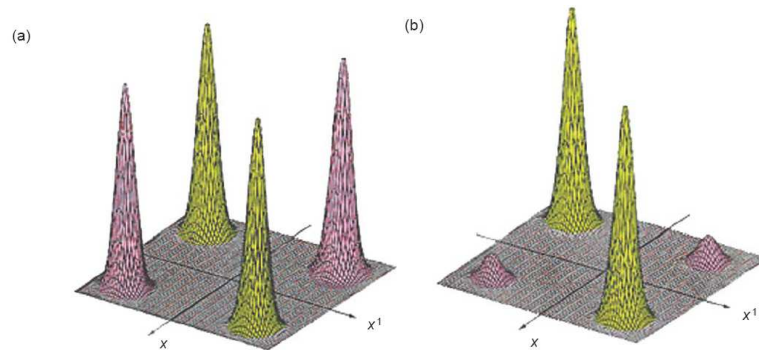


Fig. 8.3. Evolution of the Density Matrix for the Schrödinger-Cat State in fig. 8.2 (a) This plot shows the density matrix for the cat state in fig. 8.2 in the position representation $\rho(x, x') = \phi(x)\phi^*(x')$. The peaks near the diagonal (green) correspond to the two possible locations of the particle. The peaks away from the diagonal (red) are due to quantum coherence. Their existence and size demonstrate that the particle is not in either of the two approximate locations but in a coherent superposition of them. (b) Environment-induced decoherence causes decay of the off-diagonal terms of $\rho(x, x')$. Here, the density matrix in (a) has partially decohered. Further decoherence would result in a density matrix with diagonal peaks only. It can then be regarded as a classical probability distribution with an equal probability of finding the particle in either of the locations corresponding to the Gaussian wave packets.

and the Gaussians are

$$\chi^\pm = \langle x|\pm\rangle \sim \exp\left[-\frac{(x \pm \frac{\Delta x}{2})^2}{4\delta^2}\right] \quad (8.17)$$

For the case of wide separation ($x \gg \delta$), the corresponding density matrix $\rho(x, x') = \phi(x)\phi^*(x')$ has four peaks: Two on the diagonal defined by $x = x'$, and two off the diagonal for which x and x' are very different (see Figure 8.3). Quantum coherence is due to the off-diagonal peaks. As those peaks disappear, position emerges as an approximate preferred basis. The last term of eqn. 8.16, which is proportional to $(x - x')^2$, has little effect on the diagonal peaks. By contrast, it has a large effect on the off-diagonal peaks for which $(x - x')^2$ is approximately the square of the separation $(\Delta x)^2$. In particular, it causes the off-diagonal peaks to decay at the rate

$$\frac{d\rho^\pm}{dt} \sim \frac{2\gamma k_B T}{\hbar^2 (\Delta x)^2} \rho^\pm = \tau_D^{-1} \rho^\pm \quad (8.18)$$

It follows that quantum coherence will disappear on a decoherence time scale (Zurek 1984).

$$\tau_D \simeq \gamma^{-1} \left(\frac{\lambda_{dB}}{\Delta x}\right)^2 = \tau_R \left(\frac{\hbar}{\Delta x \sqrt{2mk_B T}}\right)^2 \quad (8.19)$$

where $\lambda_{dB} = \hbar/(2mk_B T)^{1/2}$ is the thermal de Broglie wavelength. For macroscopic objects, the decoherence time τ_D is typically much less than the relaxation time $\tau_R = \gamma^{-1}$.

For a system at temperature $T = 300\text{K}$ with mass $m = 1$ gram and separation $\Delta x = 1$ centimeter, the ratio of the two time scales is $\tau_D/\tau_R \sim 10^{-40}$! Thus, even if the relaxation rate were of the order of the age of the Universe, 10^{17} seconds, quantum coherence would be destroyed in $\tau_D \sim 10^{-23}$ seconds.

For microscopic systems and, occasionally, even for very macroscopic ones, the decoherence times are relatively long. For an electron ($m_e = 10^{-27}$ grams), τ_D can be much larger than the other relevant time scales on atomic and larger energy and distance scales. For a massive Weber bar, tiny Δx ($\sim 10^{-17}$ centimeter) and cryogenic temperatures suppress decoherence. Nevertheless, the macroscopic nature of the object is certainly crucial in facilitating the transition from quantum to classical.

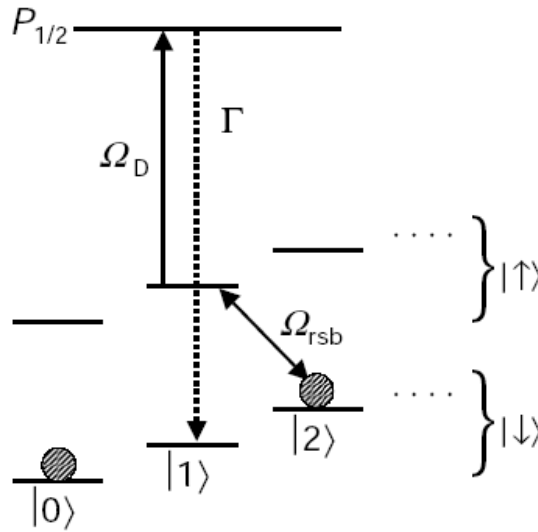


Fig. 8.4. Energy level diagram of ions used to test decoherence. A coherent superposition of states (in this case $|0\rangle$ and $|2\rangle$ as indicated by the solid circles) is initially created. Adding adding noise to the system that couples one of these states (in this case state $|2\rangle$) to other states decoherence is introduced.

8.3 Experimental tests of decoherence

Decoherence times for macroscopic systems are so short that they cannot be measured using existing technology. Myatt *et al.* Nature **403** 269 (2000) took a different approach to investigate the variation of decoherence time with separation between the states and with the magnitude of the noise. They used ions in an ion trap with the set of energy levels shown in figure 8.4. The ion trap system naturally has very weak coupling between the ions and the environment and so naturally has very long coherence times. This property makes this system ideal for implementing quantum computation. They then produced decoherence in the system by applying a noisy voltage source to the ions. The magnitude of the fluctuations in the voltage takes the place of temperature in the decoherence formula shown above. The results of their experiments shown in figure 8.5 confirmed the variation in the decoherence time with separation of initial states and with the magnitude of the fluctuations in the voltage predicted by the formula above.

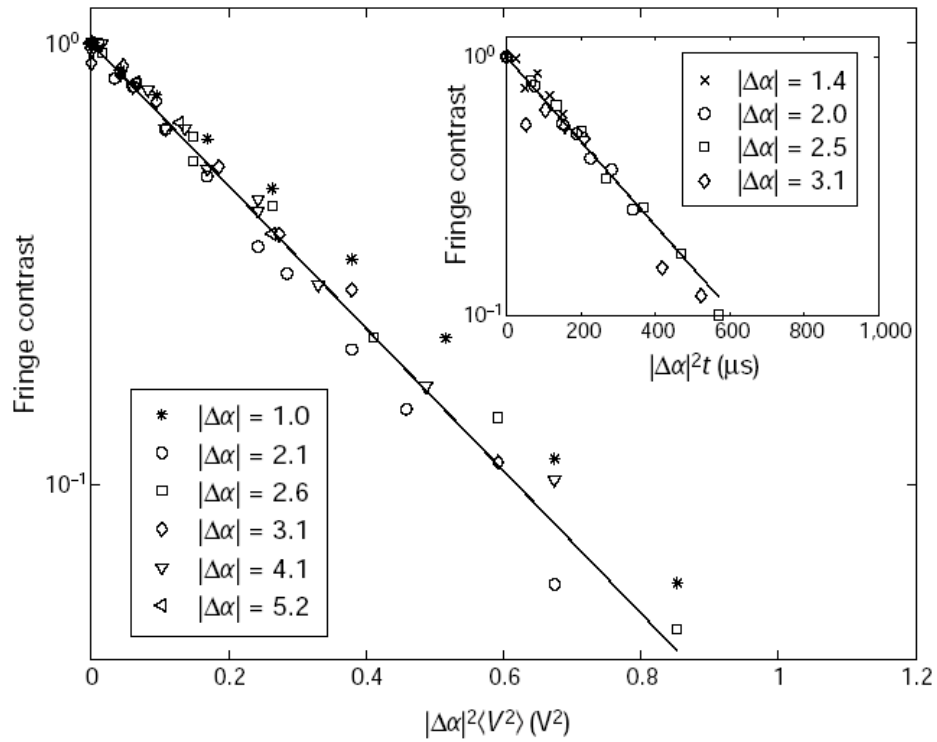


Fig. 8.5. Decoherence of Schrödinger cat states coupled to an amplitude reservoir. $\Delta\alpha$ indicates separation of initial states and V is the applied voltage that induces decoherence.

9

Quantum Cryptography

As we shall see in later lectures, quantum computers can easily break popular encryption schemes such as the RSA Public Key Crypto-system and the Data Encryption Standard (DES). However another branch of quantum information overcomes this potentially serious problem. It has been proved that quantum cryptography can provide absolute security for communications between two users. Quantum Cryptography provides techniques for creating a secure one-time pad, where the protocol for creating the one-time pad allows one to check for attack by eavesdroppers. The basis of the security of quantum cryptography is the inability to ‘clone’ quantum states. This is proved by the no-cloning theorem which is presented in section 9.2. In the following section we briefly describe the RSA encryption scheme which is widely used for internet security.

9.1 The RSA public key encryption scheme

This scheme was invented by Rivest, Shamir and Adleman (Communications of the ACM, **21** 120 (1978)) and its security is based on the difficulty of factoring large numbers. To use the scheme a user, Bob, chooses two large primes p and q and computes $N = pq$. He then randomly chooses the encryption key e such that e and $(p - 1)(q - 1)$ have no common factors. Then he computes the unique decryption key, d , such that

$$ed = 1 \pmod{(p - 1)(q - 1)} \quad (9.1)$$

This computation can be done efficiently using the Euclidean algorithm. Now e and N are made public but d must be kept secret. p and q are no longer needed so they can be discarded but they must not be revealed. Now, suppose Alice wants to send a message $m \pmod{N}$ to Bob. She can do this

securely by calculating

$$c = m^e \pmod{N} \quad (9.2)$$

and sending c to Bob. Bob recovers the message m by raising c to the power d . From number theory it is known that $m^{(p-1)(q-1)} = 1 \pmod{N}$ for any $m \pmod{N}$ and therefore Bob's operation yields

$$c^d = m^{ed} = m^{k(p-1)(q-1)+1} = m^{k(p-1)(q-1)} \times m = m \pmod{N} \quad (9.3)$$

An eavesdropper Eve who does not know d or the factorization of N will have difficulty deducing m from c, e and N alone. On the other hand if Eve can factor N into p times q then she can easily find the decryption key d by using the Euclidean algorithm with e and $(p-1)(q-1)$ as the inputs.

9.2 The no-cloning theorem

The no-cloning theorem shows that in general it is not possible to make clones (copies) of unknown quantum states. Suppose we have a quantum cloning machine with an initial state

$$|\psi\rangle \otimes |0\rangle \quad (9.4)$$

where $|\psi\rangle$ is the state we want to clone and $|0\rangle$ the state we wish to become the clone. When we operate the machine a unitary operation U acts so that

$$|\psi\rangle \otimes |0\rangle \rightarrow U |\psi\rangle \otimes |0\rangle = |\psi\rangle \otimes |\psi\rangle \quad (9.5)$$

Suppose we now want to make a clone of a second state $|\phi\rangle$. Then we have

$$\begin{aligned} U |\psi\rangle \otimes |0\rangle &= |\psi\rangle \otimes |\psi\rangle \\ U |\phi\rangle \otimes |0\rangle &= |\phi\rangle \otimes |\phi\rangle \end{aligned} \quad (9.6)$$

Taking the inner product of these two equations gives

$$\langle\psi|\phi\rangle = (\langle\psi|\phi\rangle)^2 \quad (9.7)$$

which implies the either $\langle\psi|\phi\rangle = 0$ so that $|\psi\rangle$ and $|\phi\rangle$ must be orthogonal or $\langle\psi|\phi\rangle = 1$ so the $|\psi\rangle$ and $|\phi\rangle$ are the same state. Thus a cloning machine can only clone states that are orthogonal to one another, and therefore a general cloning machine is impossible. A potential cloning machine cannot, for example, clone qubit states $|\psi\rangle = |0\rangle$ and $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, since these states are not orthogonal. You will note that fundamentally, this proof is the same as that used by Zureck to show that a measurement apparatus cannot distinguish between non-orthogonal states.

9.3 Quantum Cryptography

It has long been known that the most secure form of cryptography is the ‘one-time pad’. This is simply a random sequence of bits that only the sender and receiver of the message have access to. The sender adds the random sequence of bits to the message to be protected before it is sent. The receiver then removes the random bits from the received signal to recover the original message. Provided that the key is only known to the sender and receiver this technique is perfectly secure but the key must only be used once for the system to remain secure. If it is used repeatedly an eavesdropper will be able to build up information about the key. Hence the name ‘one-time pad’. In order to make the system totally secure we need to find a way of ensuring that only the sender and receiver have information about sequence of bits in the ‘one time pad’ and it would also be advantageous if the pad could be transmitted over large distances on demand so that the sender and receiver always have access to new secure sequences for sending secure information.

The no-cloning theorem shows that it is not possible to copy quantum states. Thus if an eavesdropper attempts to gain information about a cryptographic key distributed using a quantum state there is no way that they can simply copy the state. They will instead have to measure the state if they are to gain any information about it. However, we know from the postulates of quantum mechanics that measurement, in general, perturbs the quantum state. The various protocols for quantum cryptography ensure that the activities of the eavesdropper are detectable. They do not prevent the eavesdropping taking place but they do provide a technique for ensuring the security of the key. Providing the key is distributed using single quantum states these protocols ensure absolute security. These schemes are usually referred to as quantum key distribution or QKD.

All of the quantum cryptography protocols work by sending and receiving single quantum states in two or more incompatible bases. These states are transmitted via a quantum channel. For instance in the BB84 protocol devised by Bennett and Brassard (Bennett C H and Brassard G 1984 Proc. IEEE Int. Conf. on Computers, Systems, and Signal Processing (Bangalore) (New York: IEEE) pp 175-9)) one basis is horizontal and vertical photon polarizations and the other is a diagonal basis with photon polarizations at 45° and 135° to the vertical. Individual photons are sent by Alice in one of 4 polarisation states available in the two bases – ie the planes of polarization are either at 0° , 45° , 90° or 135° to the vertical.

$$\begin{aligned} \{0 \rightarrow |\uparrow\rangle, 1 \rightarrow |\rightarrow\rangle\} \\ \{0 \rightarrow |\nearrow\rangle, 1 \rightarrow |\searrow\rangle\} \end{aligned} \tag{9.8}$$

Alice makes this choice randomly. Bob detects the individual photons by setting his polariser to detect one of the two bases, either to detect photons polarized either horizontally or vertically OR to detect photons with the plane of polarization in the diagonal basis with polarizations at 45° and 135° to the vertical. Bob also makes this choice randomly. Alice and Bob agree on a code where photons with their planes of polarization at either 0° or 45° to the vertical are represented by the number 1 and those with their planes of polarization at either 90° or 135° to the vertical are represented by 0. When Alice and Bob choose compatible directions for the photon sent and the detector they record the same number. When they choose incompatible directions from the postulates of quantum mechanics half the time their numbers will be the same and half the time they will be different. After sending a sequence of photons and recording the polarizations sent and detected, Alice and Bob announce their choice of basis for each photon. This information can be announced publicly and this communication is usually described as taking place via a public channel. Alice and Bob then discard all their results for the times when they were using incompatible bases, thus half of their data is discarded. The sequence of 1s and 0s they are left with forms their 'one-time pad'. They check the security of their code by publishing a section of it. If an eavesdropper, Eve, had intercepted the photons sent by Alice, she would have had to make random choices for the basis to detect the photons. Half the time this would have the same as that used by Alice and her activities would not have affected Bob's results. However, the other times, when Eve chooses an incompatible basis the photon she sends on to Bob has a different polarization from the one originally sent by Alice. Thus for half of these Bob will record a different number to the one recorded by Alice. Thus Eve's presence will cause a 25% error rate between Alice's and Bob's records. Hence, when Alice and Bob publish a section of their code they will be able to detect the activities of Eve. If they find an unacceptable rate of errors they have to discard their one-time pads and start all over again. If the error rate is below a tolerable threshold then they can conclude that their one time pad is secure and use it when needed. To make this protocol secure it is important that only single photons are sent by Alice. At present there are few reliable single-photon sources (see later) and so in practice one has to send packets whose average occupancy is much less than one photon in order to minimize the possibility of any packets containing two photons which would destroy the security of the protocol. Thus most of the packets sent in present protocols contain no photons. The following protocol, the BB92 protocol, is particularly robust against problems caused

by empty packets and requires fewer photon polarisations to be sent and detected.

The BB92 protocol is a variation of the BB84 protocol in which only two polarizations are used for sending and detecting photons. Alice sends photons either polarized vertically, which she labels 1, or polarized at 45° to the vertical, which she labels 0. She chooses the directions randomly. Bob detects only photons which are polarized horizontally, which he labels 0, or at 135° to the vertical, which he labels 1. He also chooses these direction randomly. Alice sends a sequence of photons recording the 1s and 0s corresponding to the polarizations sent and Bob records the polarisation directions chosen and also whether a photon was detected. Bob then publishes a record of which times he detected a photon. Whenever Bob detects a photon, he and Alice have recorded the same number and thus this sequence of numbers is their one-time pad. This protocol has the advantage of requiring fewer polarization directions at both sending and receiving stages of the process and it is also robust against empty photon packets since data is retained only when a photon is detected, in practice as long as polarisers that detect both polarization states are used, such as those employed in the Aspect experiment, null packets are not a problem in the BB84 protocol.

9.4 Error correction and privacy amplification

Due to the finite efficiency of polarisers and the randomization of the plane of polarization of light travelling through any medium, the sequences that Alice and Bob retain will not be identical. They can remove the errors in their one time pads using a variety of schemes. One simple one is based on parity detection. A block of bits of length, k , is chosen such that the probability of more than one error is below some required tolerance. Then Alice and Bob announce the parities of each of the blocks. If these are the same then the blocks can be assumed to contain no errors. If they do not agree they announce the parities of sub-blocks of the data until the error is detected. In order to maintain the total security of the protocols Alice and Bob discard the last (or any other suitably chosen) bit of each block of data announced.

Even if Eve eavesdrops on every photon sent between Alice and Bob it still only introduces an error rate of 25% into the code. If the polarisers and all the other sources of noise introduce an intrinsic error rate ϵ between Alice's and Bob's codes, then Alice can intercept 4ϵ of the photons without being detected since her activities only cause errors at the intrinsic error rate. Given the finite efficiencies of polarisers and the desire to use QKD

over appreciable distances Eve could gain access to an appreciable part of the key. Thus for the foreseeable future privacy amplification will be an important element of QKD. The simplest privacy amplification scheme is simply to create a shorter key by taking the parity of strings of bits from the original key, ideally chosen randomly. While the length of the key is reduced, say by a factor n , the amount of information that Eve retains about the key is more rapidly reduced by $4\epsilon^n$. Clearly, the amount of information available to any eavesdropper can be made arbitrarily small by a suitable choice of n . This is the critical element in the proof that QKD is absolutely secure. It is important to use parity rather than majority voting for privacy amplification since Eve retains more information about the key if majority voting is used.

9.5 Practical QKD

QKD using the BB92 protocol has been achieved through optical fibres and through free space. QKD has been implemented over tens of km using commercial optical fibres and over several hundred metres in free space. There are even plans to attempt it using satellites. The plane of polarisation of the light leaving an optical fibre will not generally be the same as that of the light entering the fibre. There are many mechanisms that cause rotation of the plane of polarisation. A constant angle of rotation is no problem as it can be ascertained once for all at the beginning of the process. However, additional time varying changes in the plane of polarisation are caused by thermal and mechanical fluctuations in the fibre and these have to be minimised for QKD to be successful. In practice, these fluctuations are found to be relatively slow and can be removed by resetting the relative polarisation directions between the sender and the receiver from time to time during the transfer. Rates of key transfer are relatively slow at present as it is usual to run the system with only a single photon in the system at one time, this is to prevent spurious signals caused on reflection of the photon at interfaces along the path. The biggest problem with present implementations is the lack of a reliable single photon source. This then requires the protocols to be run with packets whose average occupancy is much smaller than one, typically using a value of 0.1, in order to ensure that the probability of a packet containing two photons is below a required tolerance. The provision of a reliable one photon source would thus increase key distribution rates by a factor of 10. A single photon source has now been developed Kurtseifer *et al.* (Phys. Rev. Lett. **85**, 290 (2000)) and many groups are developing such sources.

QKD has been proven under real world conditions and it is likely that this element of quantum information technology will be used practically within a short period of time. However, Brassard *et al* (Phys. Rev. Lett. **85**, 1330 (2000)) argue that various imperfections in the elements of QKD systems will prevent implementation over long distances. Very recent work by F Grosshans *et al.* Nature **421** 238 (2003) demonstrates QKD using pulses containing large numbers of photons but using the amplitude-phase uncertainty relation as the basis of the security of the technique. For a comprehensive review of this field see the article by Gisin, Ribordy, Tittel and Zbinden [quant-ph. 0101098.pdf](#).

10

Quantum Teleportation

The no-cloning theorem shows that we cannot duplicate an arbitrary quantum state. We also know that if we measure a quantum state we irrevocably destroy the state. These were the two basic facts that underlie the security of quantum cryptography. However, these two facts do not preclude the process of recreating an unknown quantum state at some distance point in space as long as the initial quantum state is destroyed as part of this process. This entire process is referred to as quantum teleportation. The steps needed to achieve teleportation (Bennett *et al.* Phys. Rev. Lett., **70**, 1895 (1993)) will be described in section 10.2 but first of all I shall just briefly remind you about the Bell States, which, as we shall see, feature strongly in teleportation.

10.1 The Bell states

The Bell states were introduced in the lectures on entanglement. They are the complete set of maximally entangled states for two spin half particles (or any other particles that have two degrees of freedom). For generality I shall use the states a and b to represent the two basis states for each particle. The Bell states are

$$\begin{aligned} |\psi^-\rangle &= \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle) \\ |\psi^+\rangle &= \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle) \\ |\phi^-\rangle &= \frac{1}{\sqrt{2}} (|00\rangle - |11\rangle) \\ |\phi^+\rangle &= \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \end{aligned} \tag{10.1}$$

All of these states may be obtained by operating on only one of the qubits of the EPR state $|\psi^-\rangle$. Thus even if the particles in the EPR are spatially separated any of the Bell states may be generated by just Alice, say, performing local operations on her qubit. The following unitary operations applied to $|\psi^-\rangle$ generate the 4 Bell states given above, their qubit mappings are also given for completeness.

$$\begin{aligned}
 U_{0,0} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & |0\rangle \rightarrow |0\rangle & |1\rangle \rightarrow |1\rangle \\
 U_{0,1} &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} & |0\rangle \rightarrow |0\rangle & |1\rangle \rightarrow -|1\rangle \\
 U_{1,0} &= \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} & |0\rangle \rightarrow -|1\rangle & |1\rangle \rightarrow -|0\rangle \\
 U_{1,1} &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} & |0\rangle \rightarrow |1\rangle & |1\rangle \rightarrow -|0\rangle
 \end{aligned} \tag{10.2}$$

10.2 Quantum Teleportation

Suppose Alice has an unknown qubit in state $|\psi\rangle$ which she wishes to send to Bob. This could be done by carefully packaging the state and sending it to Bob, hoping that it is not destroyed in the process. However, there is a more subtle way of achieving the transfer. Suppose that Bob and Alice also share an EPR pair (the state $|\psi^-\rangle$ above). We shall use subscripts A and B to identify the components of the EPR pair that Alice and Bob have and the subscript C to identify the unknown qubit, initially in Alice's possession. The unknown qubit can be written

$$|\psi\rangle_C = a|0\rangle_C + b|1\rangle_C \tag{10.3}$$

The state of the entire system can be written

$$\begin{aligned}
 |\psi\rangle_{ABC} &= |\psi\rangle_C \otimes |\psi^-\rangle_{AB} \\
 &= \frac{1}{\sqrt{2}} (a|0\rangle_C + b|1\rangle_C) \otimes (|0\rangle_A \otimes |1\rangle_B - |1\rangle_A \otimes |0\rangle_B)
 \end{aligned} \tag{10.4}$$

A simple rearrangement of the above equation in order to write it in terms of the Bell states of CA gives

$$\begin{aligned}
 |\psi\rangle_{ABC} = \frac{1}{2} \{ & |\psi^-\rangle_{CA} \otimes (-a|0\rangle_B - b|1\rangle_B) \\
 & + |\psi^+\rangle_{CA} \otimes (-a|0\rangle_B + b|1\rangle_B) \\
 & + |\phi^-\rangle_{CA} \otimes (b|0\rangle_B + a|1\rangle_B) \\
 & + |\phi^+\rangle_{CA} \otimes (-b|0\rangle_B + a|1\rangle_B) \quad \}
 \end{aligned} \tag{10.5}$$

If Alice carries out a Bell state measurement on particles A and C, her measurement will select one of the Bell states. The probability of each outcome of her measurement is $1/4$, irrespective of the state $|\psi\rangle_C$. Hence, Alice's measurement gives her no information about the state $|\psi\rangle_C$. The resulting state of Bob's particle after the measurement will be

$$\begin{aligned}
 -a|0\rangle_B - b|1\rangle_B &= -U_{00}|\psi\rangle_C \\
 -a|0\rangle_B + b|1\rangle_B &= -U_{01}|\psi\rangle_C \\
 b|0\rangle_B + a|1\rangle_B &= -U_{10}|\psi\rangle_C \\
 -b|0\rangle_B + a|1\rangle_B &= -U_{11}|\psi\rangle_C
 \end{aligned} \tag{10.6}$$

depending on whether Alice's measurement yields one of

$$|\psi^-\rangle_{CA}, |\psi^+\rangle_{CA}, |\phi^-\rangle_{CA}, |\phi^+\rangle_{CA}. \tag{10.7}$$

The important point is that in each case the state of Bob's particle is related to the original state $|\psi\rangle_C$ by a fixed unitary transformation $U_{i,j}$, independent of the identity of the original state $|\psi\rangle_C$. Thus if Alice transmits the values of i and j to Bob, communicating just 2 bits of classical information, Bob can transform his particle so that its final state is identical to $|\psi\rangle_C$. This procedure can still be followed even if Alice and Bob are spatially separated, thus allowing an unknown quantum state to be recreated at a point an arbitrary distance away. This process does not violate the no-cloning theorem since the initial quantum state is destroyed when Alice carries out her Bell state measurement. Thus at no time do we have a copy of the unknown quantum state.

10.3 Experimental implementations of teleportation

The first experiments on teleportation were performed by Bouwmeester *et al.* (Nature **390**, 575 (1997) and Boschi *et al.* (Phys. Rev. Lett. **80**, 1121, (1998)). Neither of these experiments implemented a full Bell state measurement and thus were not complete realizations of the quantum teleportation

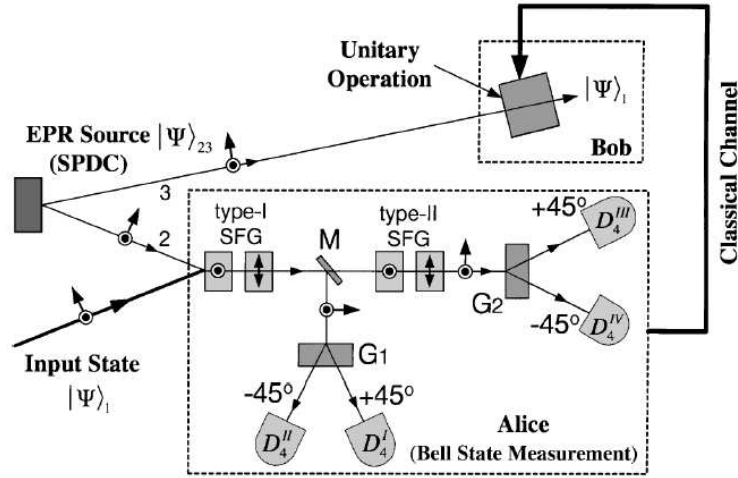


Fig. 10.1. Schematic illustration of teleportation experiment with full Bell State Measurement achieved using non-linear optical crystals (SFG in figure). Taken from paper by Kim *et al.*

protocol described above. Both experiments were performed using photons. In Bouwmeester *et al.*'s experiment only the Bell state $|\psi^-\rangle$ was detected and hence teleportation was only achieved 25% of the time. In Boschi *et al.*'s experiment to avoid performing a Bell state measurement, the photon polarization was entangled with the path of the photon in order to carry out a Bell-state like measurement. Although this experiment could, in principle, allow teleportation 100% of the time, the experiment did not teleport an unknown polarization of the photon but only the polarization imparted by an optical element within the teleportation set-up. Recently Kim, Kulik and Shih have implemented a full Bell state measurement and hence realized the protocol described above for photons (Kim *et al.* Phys. Rev. Lett. **86**, 1370, 2001).

11

Introduction to Quantum Computing

The basic idea that underlies the potential power of quantum computers is very simple. In contrast to a classical computer that can only perform a sequence of operations on a single input string of bits, a quantum computer can perform the same sequence of operations on a superposition of many input strings. If the quantum computer operates on an input of N qubits it is capable of performing 2^N simultaneous operations. However, as a result of the calculation the output register of the quantum computer is left in a superposition of all the possible results generated by all 2^N inputs. Feynman was the first person to point out the immense potential computing power of quantum computers but he did not discover a method for overcoming the measurement problem. We know that if we try to measure the state of the output registers we will cause state reduction and observe a single result of the computation and thus gain nothing over the classical computer. This obstacle to the successful exploitation of quantum computers was overcome by a number of workers including Deutsch and Jozsa. Deutsch discovered the first algorithm that would allow a particular calculation to be performed using fewer calculations on a quantum computer than a classical computer. This problem will be discussed in detail in section 11.1.

The critical conceptual step taken by Deutsch was to show that the potential power of quantum computers required a change in the paradigm for computation from the one applied in classical computing. In particular he showed that quantum computers might be most effective when searching for collective or global properties of systems and he also showed that quantum computers may not give deterministic answers, so that several runs of an algorithm might be needed to solve a problem. Deutsch's algorithm solved a trivial problem. The major leap in the field of quantum computation occurred in 1994 when Shor developed of a factorization algorithm for quantum computers. This was the first non-trivial algorithm for quantum

computers and would allow factorization of large numbers in a time that was exponentially smaller than the time required on classical computers. Shor's factorization algorithm will be discussed in section 11.3. More recently, Grover published an algorithm for quantum computers to search unstructured databases. For a database containing M entries a classical computer would take an average of $M/2$ operations to find a particular entry, essentially by using a trial and error approach. Using Grover's algorithm a quantum computer would complete this task in \sqrt{M} operations. In this case the speed advantage of the quantum computer over the classical computer is not as extreme as it is in for Shor's factorization algorithm but clearly for large values of M the quantum computer will be orders of magnitude faster than the classical computer. Grover's algorithm is described in section 11.4.

Error correction is a critical element for successful computation using classical computers. It is at least as important in quantum computers as the number of errors that can occur in quantum computers is significantly higher than in a classical computer. Ultimately, the parallel processing capacity of quantum computers is destroyed by decoherence and the only method of allowing calculations to be performed for times longer than the characteristic decoherence time of the system is by using error correction. The problem of errors and error correction in quantum computers will be discussed in section 11.5. A description of systems currently being investigated for potential use as quantum computers is contained in section 11.7.

11.1 The elements of a quantum computer

It has been shown by DiVincenzo that quantum computers are as universal as classical computers in that they can run any algorithmic task. However, there is a serious practical difference between quantum and classical computers that may limit the potential usefulness of such existence theorems for quantum computing. In a classical computer, implementing any algorithm on a machine that is not specifically designed for it will simply increase the time and memory space required for a calculation. While this may make the calculation intractable in an acceptable time or cost it does not make the calculation impossible. However, in the case of quantum computation there is a constant battle against decoherence which requires much more sophisticated error correction schemes than required on classical computers and any increase in computational time will make an algorithm much more expensive to implement.

A number of the one-qubit operations required for quantum computation have already been introduced in the section on the Bell states. There are a

number of other one-qubit operations that are used in quantum computation such as

$$\begin{aligned}
 U_{NOT} &: |0\rangle \rightarrow |1\rangle, & |1\rangle &\rightarrow |0\rangle \\
 U_\phi &: |0\rangle \rightarrow |0\rangle, & |1\rangle &\rightarrow e^{i\phi} |1\rangle \\
 H &: |0\rangle \rightarrow \frac{1}{\sqrt{2}} [|0\rangle + |1\rangle], & |1\rangle &\rightarrow \frac{1}{\sqrt{2}} [|0\rangle - |1\rangle]
 \end{aligned}
 \tag{11.1}$$

The last of these operations is often referred to as the Hadamard transform. Applying this operation to all N qubits initially in the state $|0000\dots\rangle$ generates a superposition of all possible states of the N qubits. The implementation of these one qubit operations is the subject of question 3 on examples sheet 3. The biggest challenge in quantum computation is the implementation of controllable two-qubit logic gates that are fast enough to make practical calculations feasible within the decoherence time of the system. An example of a two-qubit gate is the ‘controlled not’ gate where the first qubit is the control and the second qubit is the target, whose state is changed if the first qubit is $|1\rangle$.

$$U_{C-NOT} : |00\rangle \rightarrow |00\rangle, \quad |01\rangle \rightarrow |01\rangle, \quad |10\rangle \rightarrow |11\rangle, \quad |11\rangle \rightarrow |10\rangle
 \tag{11.2}$$

A related two-qubit gate is the SWAP operation.

$$U_{SWAP} : |00\rangle \rightarrow |00\rangle, \quad |01\rangle \rightarrow |10\rangle, \quad |10\rangle \rightarrow |01\rangle, \quad |11\rangle \rightarrow |11\rangle
 \tag{11.3}$$

The final ingredient for our quantum computer is the measurement gate, which reads 0 if the qubit is in state $|0\rangle$ and 1 if the qubit is in state $|1\rangle$. If the qubit is in state $\alpha|0\rangle + \beta|1\rangle$ then the measurement becomes probabilistic and is 0 with probability $|\alpha|^2$ and is 1 with probability $|\beta|^2$.

11.2 Deutsch’s Problem

Following Deutch, imagine we have a black box that computes a function that takes a single bit x to a single bit $f(x)$. We don’t know what is happening inside the box, but it must be something complicated because the computation takes 24 hours. There are four possible functions $f(x)$ (because each of $f(0)$ and $f(1)$ can be either 0 or 1) and we’d like to know which of them the box is computing. It would take 48 hours to find out both $f(0)$ and $f(1)$ but we don’t have that much time –we need the result in 24hours.

It turns out though that we would be satisfied to know whether $f(x)$ is constant ($f(0) = f(1)$) or balanced ($f(0) \neq f(1)$). Even so, it would take 48 hours to get the answer.

Now suppose we have a quantum black box that computes $f(x)$. Of course $f(x)$ might not be invertible, so we will need a transformation U_f that takes two qubits to two:

$$U_f : |x\rangle|y\rangle \rightarrow |x\rangle|y \oplus f(x)\rangle. \quad (11.4)$$

(This transformation flips the second qubit if f acting on the first qubit is 1, and doesn't do anything if f acting on the first qubit is 0.) We can determine if $f(x)$ is constant or balanced by using the quantum black box twice. But it still takes a day for it to produce one output, so that won't do. Can we get the answer in 24 hours by running the quantum black box *just once*. (This is 'Deutsch's problem'.)

Because the black box is a quantum computer we can choose the input state to be a *superposition* of $|0\rangle$ and $|1\rangle$. If the second qubit is initially prepared in the state $(|0\rangle - |1\rangle)/\sqrt{2}$, then

$$\begin{aligned} U_f : |x\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) &\rightarrow |x\rangle \frac{1}{\sqrt{2}}(|f(x)\rangle - |1 \oplus f(x)\rangle) \\ &= |x\rangle (-1)^{f(x)} \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \end{aligned} \quad (11.5)$$

so we have isolated the function f in an x -dependent phase. Now suppose we prepare the first qubit in the state $(|0\rangle + |1\rangle)/\sqrt{2}$. Then the black box acts as

$$U_f : \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \rightarrow \frac{1}{\sqrt{2}}((-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \quad (11.6)$$

Finally, we can perform a measurement that projects the first qubit onto the basis

$$|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle). \quad (11.7)$$

Evidently, we will always obtain $|+\rangle$ if the function is balanced, and $|-\rangle$ if the function is constant.

So we have solved Deutsch's problem and we have found a separation between what a classical computer and a quantum computer can achieve. The classical computer has to run the black box twice to distinguish a balanced function from a constant function, but a quantum computer does the job in one go!

This is possible because the quantum computer is not limited to computing either $f(0)$ or $f(1)$. It can act on a superposition of $|0\rangle$ and $|1\rangle$ and thereby extract ‘global’ information about the function, information that depends on both $f(0)$ and $f(1)$. This is quantum parallelism.

11.3 Shor’s factorisation algorithm

Factorisation is an inefficient process on classical computers because it relies on a trial and error approach. To factorise a number N ($= pq$ where p and q are primes) which has M digits requires of the order $10^{M/2}$ attempts (trial division with every number between 1 and \sqrt{N}). Shor discovered a much more efficient algorithm that could be implemented on quantum computers (P.W. Shor in Proceedings of the 35th Annual Symposium on the Foundations of Computer Science p124 (IEEE Computer Society Press, Los Alamos, California) 1994). The starting point for Shor’s algorithm is the result from number theory that finding the two prime factors p and q of $N = pq$ can be done efficiently if one knows the period of the function

$$f_{a,N}(x) = a^x \pmod{N} \quad (11.8)$$

where $x = 0, 1, 2, 3, \dots$ and a is *any* randomly chosen number smaller than N which is coprime with N (i.e has no common factors with N , if it has and you know that $N = pq$ where p and q are primes then you have solved the problem by a lucky guess –this is simply checked by application of the Euclidean algorithm, an efficient algorithm that has been known since 300 B.C.!).

The function given above is periodic with period r , which depends on a and N . Once r is known we can factor N provided that r is even and $a^{r/2} \pmod{N} \neq -1$. When a is chosen randomly between 2 and N these two conditions are satisfied with probability greater than 0.5 so that even if we fail to find a suitable r with our first choice of a we will not need many other choices of a before we succeed. Once we have a suitable value of r the factors of N are given by the greatest common divisors of $(a^{r/2} \pm 1, N)$. This last calculation can be performed efficiently using the Euclidean algorithm.

11.3.1 Classical Example

We shall use this method to factor the number $N = 15$.

First we select a such that $\gcd(a, N) = 1$ so a could be any number from the set $\{2, 4, 7, 8, 11, 13, 14\}$ and we shall choose $a = 7$.

The values of $f_{7,15} = 7^x \pmod{15}$ for $x = 0, 1, 2, 3, 4, 5, 6, 7$, are 1,7,4,13,1,7,4,13,.

Note that once a number is repeated in the series the sequence repeats from that point.

Here r , the period of the function, is 4 and by computing the largest common factors of $(7^2 \pm 1, 15)$ we find $\gcd(50, 15) = 5$ and $\gcd(48, 15) = 3$, the two factors of 15.

The periods of $f_{a,15}(x)$ for all the possible values of a in the set $\{2, 4, 7, 8, 11, 13, 14\}$ are respectively $\{4, 2, 4, 4, 2, 4, 2\}$ and in this case the method is successful for all values of a except $a = 14$. For $a = 14$ we find $r = 2$ and then $a^{r/2} \pmod{15} = -1$ so the method fails.

11.3.2 Quantum algorithm

We use $2L$ qubits for the entire calculation, the first L qubits will be used for the first register and the second L qubits will be used for the second register. We initially set the two registers in the state

$$|0\rangle|0\rangle \quad (11.9)$$

and then apply the Hadamard transform (H) to the first L qubits in the first register to place them in a superposition of all possible states between 0 and $2^L - 1$ as follows

$$|0\rangle|0\rangle \rightarrow \frac{1}{\sqrt{2^L}} \sum_{x=0}^{2^L-1} |x\rangle|0\rangle \quad (11.10)$$

We then apply the operator $U_{f_{a,N}}$ defined by

$$\begin{aligned} \frac{1}{\sqrt{2^L}} \sum_{x=0}^{2^L-1} |x\rangle|0\rangle &\rightarrow \frac{1}{\sqrt{2^L}} \sum_{x=0}^{2^L-1} |x\rangle|a^x \pmod{N}\rangle \\ &= \frac{1}{\sqrt{2^L}} \sum_{s=0}^{r-1} |\phi_s\rangle|a^s \pmod{N}\rangle \end{aligned} \quad (11.11)$$

where, as before, r is the periodicity of the function $a^s \pmod{N}$ and the first register functions are

$$|\phi_s\rangle = \sqrt{\frac{r}{2^L}} \sum_{q=0}^{\frac{2^L}{r}-1} |rq + s\rangle \quad (11.12)$$

clearly, these first register functions are periodic with period r .

Taking the same values for N and a as in the previous example the state

of the registers after applying this operation is the following superposition of states of the two registers

$$\begin{aligned} \frac{1}{4} (& |0\rangle|1\rangle + |1\rangle|7\rangle + |2\rangle|4\rangle + |3\rangle|13\rangle + |4\rangle|1\rangle + |5\rangle|7\rangle + |6\rangle|4\rangle + |7\rangle|13\rangle + \\ & |8\rangle|1\rangle + |9\rangle|7\rangle + |10\rangle|4\rangle + |11\rangle|13\rangle + |12\rangle|1\rangle + |13\rangle|7\rangle + |14\rangle|4\rangle + |15\rangle|13\rangle \\ & + \dots \end{aligned} \quad (11.13)$$

rearranging this state we find

$$\begin{aligned} |\phi_0\rangle &= \frac{1}{2}(|0\rangle + |4\rangle + |8\rangle + |12\rangle + \dots) |1\rangle \\ |\phi_1\rangle &= \frac{1}{2}(|1\rangle + |5\rangle + |9\rangle + |13\rangle + \dots) |7\rangle \\ |\phi_2\rangle &= \frac{1}{2}(|2\rangle + |6\rangle + |10\rangle + |14\rangle + \dots) |4\rangle \\ |\phi_3\rangle &= \frac{1}{2}(|3\rangle + |7\rangle + |11\rangle + |15\rangle + \dots) |13\rangle \end{aligned} \quad (11.14)$$

Simply measuring the value of register 1 at this point is not efficient since the combination of measuring both registers 1 and 2 gives only one piece of information, which offers no advantage over a classical computer. However, by applying a discrete Fourier transform to the first register at this point

$$\begin{aligned} |\phi_s\rangle &\rightarrow \sqrt{\frac{r}{2^L}} \sum_{q=0}^{\frac{2^L}{r}-1} \frac{1}{\sqrt{2^L}} \sum_{p=0}^{2^L-1} e^{2\pi i(rq+s)p/2^L} |p\rangle \\ &= \frac{1}{\sqrt{2^L}} \sum_{p=0}^{2^L-1} e^{2\pi isp/2^L} \sqrt{\frac{r}{2^L}} \left[\sum_{q=0}^{\frac{2^L}{r}-1} e^{2\pi irqp/2^L} \right] |p\rangle \\ &= \frac{1}{\sqrt{2^L}} \sum_{p=0}^{2^L-1} F(p) |p\rangle \end{aligned} \quad (11.15)$$

measurements of the first register will produce values $p = m2^L/r$, where m is an integer, irrespective of the value of the second register, provided that r is a factor of 2^L . If r is not a factor of 2^L then the measured value of p will be close to the values $m2^L/r$, where m is an integer. In the simple case where r is a factor of 2^L , just a few measurements of the first registers following the calculation will be sufficient to deduce the value of r . In the second case more measurements are needed. Here it is also important that the number of qubits used for the entire calculation, $2L$, is chosen so that

2^L is much larger than N so that the peaks in the discrete Fourier transform are quite narrow. Since one does not know in advance whether r will be a factor of 2^L Shor suggests that the algorithm should be run using $2L$ qubits where 2^L is of the order of N^2 .

Shor's algorithm has been implemented on an NMR quantum computer by Vandersypen *et al* published in Nature, **414** 883 (2001).

11.4 Grover's database search algorithm

If we know someone's telephone number and have to find their address by searching for this number in a telephone directory containing N entries we will have to look up $N/2$ entries on average before we find the number we want. If we were searching a sorted database of telephone numbers the time would be much smaller but if the database is unsorted (or unstructured) we have to take a trial and error approach. Grover showed that the same search could be performed on a quantum computer using of the order of \sqrt{N} operations. Grover's algorithm is described in Phys. Rev. Lett. **79**, 325 (1997). The technique works by increasing the amplitude of the target state vector in the superposition of all the states in the database. After $O(\sqrt{N})$ iterations of Grover's algorithm the probability of obtaining the required target state on measurement is greater than 0.5. Clearly we will sometimes need to apply this search technique a number of times before we can be certain of success. Interestingly, if we carry out twice as many iterations of Grover's algorithm before measuring the state of the system the probability of finding the target state falls almost to zero.

11.4.1 Grover's algorithm

For a search of a database of size 2^n for an entry x_0 , n qubits are required. The initial state is set to be

$$|0\rangle^{\otimes n} \quad (11.16)$$

A Hadamard transformation is carried out on all n qubits so that the state becomes

$$|0\rangle^{\otimes n} \rightarrow H^{\otimes n} |0\rangle^{\otimes n} = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \quad (11.17)$$

The notation x is used as a label for each of the possible n qubit states $|x\rangle$ (e.g for $n = 4$ the states could be labeled $|0\rangle = |0000\rangle$, $|1\rangle = |0001\rangle$, $|2\rangle =$

$|0011\rangle \dots$). The next step is called ‘Grover iteration’ where the operator

$$G = H^{\otimes n}(2|0\rangle\langle 0| - \mathbb{I})H^{\otimes n}O \quad (11.18)$$

is applied repeatedly. O is the so called ‘oracle’ which performs the transformation

$$O|x\rangle = (-1)^{f(x)}|x\rangle, \quad (11.19)$$

where $f(x) = 0$ for all $0 \leq x \leq 2^n$ except x_0 , for which $f(x_0) = 1$. The state then becomes

$$\begin{aligned} &\rightarrow G^R|\psi\rangle \\ &\simeq |x_0\rangle \end{aligned} \quad (11.20)$$

where for an optimal solution $R \simeq \pi\sqrt{2^n}/4$. The first n qubits are measured to find x_0 .

11.4.2 Grover’s algorithm in the two-qubit case ($N = 4$)

This case is special in that the correct result is obtained after a single Grover iteration $R = 1$. We will look at the case where the target state is $|01\rangle$. The first Hadamard transformation gives

$$|00\rangle \rightarrow H^{\otimes 2}|00\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle) \quad (11.21)$$

We then apply the oracle which inverts the coefficient of the target only

$$\begin{aligned} &\rightarrow \frac{1}{2}(|00\rangle - |01\rangle + |10\rangle + |11\rangle) \\ &= H^{\otimes 2}|00\rangle - |01\rangle \end{aligned} \quad (11.22)$$

Then apply $H^{\otimes 2}$, noting that $(H^{\otimes 2})^2 = 1$

$$\begin{aligned} &\rightarrow H^{\otimes 2}H^{\otimes 2}|00\rangle - H^{\otimes 2}|01\rangle \\ &= |00\rangle - \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle) \\ &= \frac{1}{2}(|00\rangle + |01\rangle - |10\rangle + |11\rangle) \end{aligned} \quad (11.23)$$

Now apply $2|00\rangle\langle 00| - \mathbb{I}$ which inverts the sign of the coefficients of all of the states except $|00\rangle$

$$\begin{aligned} &\rightarrow \frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle) \\ &= H^{\otimes 2}|00\rangle - |01\rangle - |11\rangle \end{aligned} \quad (11.24)$$

Finally, apply $H^{\otimes 2}$ again,

$$\begin{aligned}
 &\rightarrow H^{\otimes 2} H^{\otimes 2} |00\rangle - H^{\otimes 2} |01\rangle - H^{\otimes 2} |11\rangle \\
 &= |00\rangle - \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle) - \frac{1}{2}(|0\rangle - |1\rangle)(|0\rangle - |1\rangle) \\
 &= |01\rangle
 \end{aligned} \tag{11.25}$$

Therefore, the target is found, with certainty, with just one application of G , (i.e. one ‘recognition’ operation) for the $N = 4$ case. Grover’s algorithm has been implemented on an NMR quantum computer by Jones *et al.*, Nature **393**, 344 (1998).

11.5 Errors

This next section is taken directly from Preskill’s lecture notes.

As significant as Shor’s factoring algorithm may prove to be, there is another recently discovered feature of quantum information that may be just as important: the discovery of quantum error correction. Indeed, were it not for this development, the prospects for quantum computing technology would not seem bright.

As we have noted, the essential property of quantum information that a quantum computer exploits is the existence of nonlocal correlations among the different parts of a physical system. If I look at only part of the system at a time, I can decipher only very little of the information encoded in the system.

Unfortunately, these nonlocal correlations are extremely fragile and tend to decay very rapidly in practice. The problem is that our quantum system is inevitably in contact with a much larger system, its environment. It is virtually impossible to perfectly isolate a big quantum system from its environment, even if we make a heroic effort to do so. Interactions between a quantum device and its environment establish nonlocal correlations between the two. Eventually the quantum information that we initially encoded in the device becomes encoded, instead, in correlations between the device and the environment. At that stage, we can no longer access the information by observing only the device. In practice, the information is irrevocably lost. Even if the coupling between device and environment is quite weak, this happens to a macroscopic device remarkably quickly. Erwin Schrödinger chided the proponents of the mainstream interpretation of quantum mechanics by observing that the theory will allow a quantum state of a cat of the form

$$|\text{cat}\rangle = \frac{1}{\sqrt{2}}(|\text{dead}\rangle + |\text{alive}\rangle). \tag{11.26}$$

To Schrödinger, the possibility of such states was a blemish on the theory, because every cat he had seen was either dead or alive, not half dead and half alive.

One of the most important advances in quantum theory over the past 15 years is that we have learned how to answer Schrödinger with growing confidence. The state $|\text{cat}\rangle$ is possible in principle, but is rarely seen because it is extremely unstable. The cats Schrödinger observed were never well isolated from the environment. If someone were to prepare the state $|\text{cat}\rangle$, the quantum information encoded in the superposition of $|\text{dead}\rangle$ and $|\text{alive}\rangle$ would immediately be transferred to correlations between the cat and the environment, and become completely inaccessible. In effect, the environment continually measures the cat, projecting it onto either the state $|\text{alive}\rangle$ or $|\text{dead}\rangle$. This process is called decoherence. We have looked at decoherence earlier in the course.

Now, to perform a complex quantum computation, we need to prepare a delicate superposition of states of a relatively large quantum system (though perhaps not as large as a cat). Unfortunately, this system cannot be perfectly isolated from the environment, so this superposition, like the state $|\text{cat}\rangle$ decays very rapidly. The encoded quantum information is quickly lost, and our quantum computer crashes.

To put it another way, contact between the computer and the environment (decoherence) causes errors that degrade the quantum information. To operate a quantum computer reliably, we must find some way to prevent or correct these errors.

Actually, decoherence is not our only problem. Even if we could achieve perfect isolation from the environment, we could not expect to operate a quantum computer with perfect accuracy. The quantum gates that the machine executes are unitary transformations that operate on a few qubits at a time, let's say 4×4 unitary matrices acting on two qubits. Of course, these unitary matrices form a continuum. We may have a protocol for applying U_0 to 2 qubits, but our execution of the protocol will not be flawless, so the actual transformation

$$U = U_0(1 + O(\epsilon)) \tag{11.27}$$

will differ from the intended U_0 by some amount of order ϵ . After about $1/\epsilon$ gates are applied, these errors will accumulate and induce a serious failure. Classical analog devices suffer from a similar problem, but small errors are much less of a problem for devices that perform discrete logic.

In fact, modern digital circuits are remarkably reliable. They achieve such high accuracy with help from *dissipation*. We can envision a classical gate

that acts on a bit, encoded as a ball residing at one of the two minima of a double-lobed potential. The gate may push the ball over the intervening barrier to the other side of the potential. Of course, the gate won't be implemented perfectly; it may push the ball a little too hard. Over time, these imperfections might accumulate, causing an error.

To improve the performance, we cool the bit (in effect) after each gate. This is a dissipative process that releases heat to the environment and compresses the phase space of the ball, bringing it close to the local minimum of the potential. So the small errors that we may make wind up heating the environment rather than compromising the performance of the device.

But we can't cool a quantum computer this way. Contact with the environment may enhance the reliability of classical information, but it would destroy encoded quantum information. More generally, accumulation of error will be a problem for classical reversible computation as well. To prevent errors from building up we need to discard the information about the errors, and throwing away information is always a dissipative process.

Still, let's not give up too easily. A sophisticated machinery has been developed to contend with errors in classical information, the theory of error correcting codes. To what extent can we coopt this wisdom to protect quantum information as well?

How does classical error correction work? The simplest example of a classical error-correcting code is a repetition code: we replace the bit we wish to protect by 3 copies of the bit,

$$\begin{aligned} 0 &\rightarrow (000), \\ 1 &\rightarrow (111). \end{aligned} \tag{11.28}$$

Now an error may occur that causes one of the three bits to flip; if it's the first bit, say,

$$\begin{aligned} (000) &\rightarrow (100), \\ (111) &\rightarrow (011). \end{aligned} \tag{11.29}$$

Now in spite of the error, we can still decode the bit correctly, by majority voting.

In the 50's John Von Neumann showed that a classical computer with noisy components can work reliably, by employing sufficient redundancy. He pointed out that, if necessary, we can compute each logic gate many times, and accept the majority result. (Von Neumann was especially interested in how his brain was able to function so well in spite of the unreliability of neurons. He was pleased to explain why he was so smart.)

But now we want to use error correction to keep a quantum computer on track, and we can immediately see that there are difficulties:

- 1 **Phase errors.** With quantum information, more things can go wrong. In addition to bit-flip errors

$$\begin{aligned} |0\rangle &\rightarrow |1\rangle, \\ |1\rangle &\rightarrow |0\rangle. \end{aligned} \tag{11.30}$$

there can also be phase errors

$$\begin{aligned} |0\rangle &\rightarrow |0\rangle, \\ |1\rangle &\rightarrow -|1\rangle. \end{aligned} \tag{11.31}$$

A phase error is serious, because it makes the state $(|0\rangle + |1\rangle)/\sqrt{2}$ flip to the orthogonal state $(|0\rangle - |1\rangle)/\sqrt{2}$. But the classical coding provided no protection against phase errors.

- 2 **Small errors.** As already noted, quantum information is continuous. If a qubit is intended to be in the state

$$a|0\rangle + b|1\rangle \tag{11.32}$$

an error might change a and b by an amount of order ϵ and these small errors can accumulate over time. The classical method is designed to correct large (bit flip) errors.

- 3 **Measurement causes disturbance.** In the majority voting scheme, it seemed that we needed to measure the bits in the code to detect and correct the errors. But we can't measure qubits without disturbing the quantum information that they encode.
- 4 **No cloning.** With classical coding, we protected information by making extra copies of it. But we know that quantum information cannot be copied with perfect fidelity.

11.6 Quantum error-correcting codes

Despite these obstacles, it turns out that quantum error correction really is possible. The first example of a quantum error-correcting code was constructed by Peter Shor. This discovery ushered in a new discipline that has matured remarkably quickly—the theory of quantum error-correcting codes.

Probably the best way to understand how quantum error correction works is to examine Shor's original code. It is the most straightforward quantum generalization of the classical 3-bit repetition code.

Let's look at that 3-bit code one more time, but this time mindful of the

requirement that, with a quantum code, we will need to be able to correct the errors without measuring any of the encoded information.

Suppose we encode a single qubit with 3 qubits:

$$\begin{aligned} |0\rangle &\rightarrow |\bar{0}\rangle \equiv |000\rangle \\ |1\rangle &\rightarrow |\bar{1}\rangle \equiv |111\rangle, \end{aligned} \quad (11.33)$$

or, in other words, we encode a superposition

$$a|0\rangle + b|1\rangle \rightarrow a|\bar{0}\rangle + b|\bar{1}\rangle = a|000\rangle + b|111\rangle \quad (11.34)$$

We would like to be able to correct a bit-flip error without destroying this superposition.

Of course, it won't do to measure a single qubit. If I measure the first qubit and get the result $|0\rangle$, then I have prepared the state $|\bar{0}\rangle$ of all three qubits, and we have lost the quantum information encoded in the coefficients a and b .

But there is no need to restrict our attention to single-qubit measurements. I could also perform collective measurements on two-qubits at once, and collective measurements suffice to diagnose a bit-flip error. For a 3-qubit state $|x, y, z\rangle$ I could measure, say, the two-qubit observables $y \oplus z$, or $x \oplus z$ (where \oplus denotes addition modulo 2). For both $|x, y, z\rangle = |000\rangle$ and $|111\rangle$ these would be 0, but if any one bit flips, then at least one of these quantities will be 1. In fact, if there is a single bit flip, the two bits

$$(y \oplus z, x \oplus z) \quad (11.35)$$

just designate in binary notation the position (1,2,3), of the bit that flipped. These two bits constitute a *syndrome* that diagnoses the error that occurred.

For example, if the first bit flips,

$$a|000\rangle + b|111\rangle \rightarrow a|100\rangle + b|011\rangle \quad (11.36)$$

then the measurement of $(y \oplus z, x \oplus z)$ yields the result (0, 1) which instructs us to flip the first bit; this indeed repairs the error.

Of course, instead of a (large) bit flip there could be a small error:

$$\begin{aligned} |000\rangle &\rightarrow |000\rangle + \epsilon|100\rangle \\ |111\rangle &\rightarrow |111\rangle + \epsilon|011\rangle \end{aligned} \quad (11.37)$$

But even in this case the above procedure would work fine. In measuring $(y \oplus z, x \oplus z)$ we would project out an eigenstate of this observable. Most of the time (probability $1 - |\epsilon|^2$) we obtain the result (0,0) and project the damaged state back to the original state, and so correct the error. Occasionally (probability $|\epsilon|^2$) we obtain the result (0,1) and project the state

onto eqn. 11.37. But then the syndrome instructs us to flip the first bit, which restores the original state. Similarly, if there is an amplitude of order ϵ for each of the three qubits to flip, then with a probability of order $|\epsilon|^2$ the syndrome measurement will project the state to one in which one of the three bits is flipped, and the syndrome will tell us which one.

So we have already overcome 3 of the 4 obstacles cited earlier. We see that it is possible to make a measurement that diagnoses the error without damaging the information (answering (3)), and that a quantum measurement can project a state with a small error to either a state with no error or a state with a large discrete error that we know how to correct (answering (2)). As for (4) the issue didn't come up, because the state $a|\bar{0}\rangle + b|\bar{1}\rangle$ is not obtained by cloning – it is not the same as $(a|\bar{0}\rangle + b|\bar{1}\rangle)^{\otimes 3}$; that is, it differs from three copies of the unencoded state.

Only one challenge remains: (1) phase errors. Our code does not yet provide any protection against phase errors, for if any one of the three qubits undergoes a phase error then our encoded state $a|\bar{0}\rangle + b|\bar{1}\rangle$ is transformed to $a|\bar{0}\rangle - b|\bar{1}\rangle$ and the encoded quantum information is damaged. In fact, phase errors have become three times more likely than if we hadn't used the code. But with the methods in hand that conquered problems (2)-(4) we can approach problem (1) with new confidence. Having protected against bit-flip errors by encoding bits redundantly, we are led to protect against phase-flip errors by encoding phases redundantly.

Following Shor, we encode a single qubit using nine qubits, according to

$$\begin{aligned} |0\rangle &\rightarrow |\bar{0}\rangle \equiv \frac{1}{2^{3/2}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle), \\ |1\rangle &\rightarrow |\bar{1}\rangle \equiv \frac{1}{2^{3/2}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle) \end{aligned} \quad (11.38)$$

Both $|\bar{0}\rangle$ and $|\bar{1}\rangle$ consist of three clusters of three qubits each, with each cluster prepared in the same quantum state. Each of the clusters has triple bit redundancy, so we can correct a single bit flip in any cluster by the method discussed above.

Now suppose that a phase flip occurs in one of the clusters. The error changes the relative sign of $|000\rangle$ and $|111\rangle$ in that cluster so that

$$\begin{aligned} |000\rangle + |111\rangle &\rightarrow |000\rangle - |111\rangle \\ |000\rangle - |111\rangle &\rightarrow |000\rangle + |111\rangle \end{aligned} \quad (11.39)$$

This means that the relative phase of the damaged cluster differs from the phases of the other two clusters. Thus, as in our discussion of bit-flip correction, we can identify the damaged cluster, not by measuring the relative

phase in each cluster (which would disturb the encoded information) but by comparing the phases of pairs of clusters. In this case, we need to measure a six-qubit observable to do the comparison, e.g., the observable that flips qubits 1 through 6. Since flipping twice is the identity, this observable squares to 1, and has eigenvalues ± 1 . A pair of clusters with the same sign is an eigenstate with eigenvalue $+1$, and a pair of clusters with opposite sign is an eigenstate with eigenvalue -1 . By measuring the six-qubit observable for a second pair of clusters, we can determine which cluster has a different sign than the others. Then, we apply a unitary phase transformation to one of the qubits in that cluster to reverse the sign and correct the error. Now suppose that a unitary error $U = 1 + O(\epsilon)$ occurs for each of the 9 qubits. The most general single-qubit unitary transformation (aside from a physically irrelevant overall phase) can be expanded to order ϵ as

$$U = 1 + i\epsilon_x\sigma_x + i\epsilon_y\sigma_y + i\epsilon_z\sigma_z \quad (11.40)$$

the three terms of order ϵ in the expansion can be interpreted as a bit-flip operator, a phase flip operator, and an operator in which both a bit flip and a phase flip occur. If we prepare an encoded state $a|\bar{0}\rangle + b|\bar{1}\rangle$, allow the unitary errors to occur on each qubit, and then measure the bit-flip and phase-flip syndromes, then most of the time we will project the state back to its original form, but with a probability of order $|\epsilon|^2$, one qubit will have a large error: a bit flip, a phase flip, or both. From the syndrome, we learn which bit flipped, and which cluster had a phase error, so we can apply the suitable one-qubit unitary operator to fix the error.

Error recovery will fail if, after the syndrome measurement, there are two bit flip errors in each of two clusters (which induces a phase error in the encoded data) or if phase errors occur in two different clusters (which induces a bit-flip error in the encoded data). But the probability of such a double phase error is of order $|\epsilon|^4$. So for $|\epsilon|$ small enough, coding improves the reliability of the quantum information. The code also protects against decoherence. By restoring the quantum state irrespective of the nature of the error, our procedure removes any entanglement between the quantum state and the environment. Here as always, error correction is a dissipative process, since information about the nature of the errors is flushed out of the quantum system. In this case, that information resides in our recorded measurement results, and heat will be dissipated when that record is erased.

Let's summarize the essential ideas that underlie our quantum error correction scheme:

- 1 We digitized the errors. Although the errors in the quantum informa-

tion were small, we performed measurements that projected our state onto either a state with no error, or a state with one of a discrete set of errors that we knew how to convert.

- 2 We measured the errors without measuring the data. Our measurements revealed the nature of the errors without revealing (and hence disturbing) the encoded information.
- 3 The errors are local, and the encoded information is nonlocal. It is important to emphasize the central assumption underlying the construction of the code –that errors affecting different qubits are, to a good approximation, uncorrelated. We have tacitly assumed that an event that causes errors in two qubits is much less likely than an event causing an error in a single qubit. It is of course a physics question whether this assumption is justified or not –we can easily envision processes that will cause errors in two qubits at once. If such correlated errors are common, coding will fail to improve reliability.

The code takes advantage of the presumed local nature of the errors by encoding the information in a nonlocal way –that is the information is stored in correlations involving several qubits. There is no way to distinguish $|\bar{0}\rangle$ from $|\bar{1}\rangle$ by measuring a single qubit of the nine. If we measure one qubit we will find $|0\rangle$ with probability $1/2$ and $|1\rangle$ with probability $1/2$ irrespective of the value of the encoded qubit. To access the encoded information we need to measure a 3-qubit observable (the operator that flips all three qubits in a cluster can distinguish $|000\rangle + |111\rangle$ from $|000\rangle - |111\rangle$).

The environment might occasionally kick one of the qubits, in effect ‘measuring’ it. But the encoded information cannot be damaged by disturbing that one qubit, because a single qubit, by itself, actually carries no information at all. Nonlocally encoded information is invulnerable to local influences –this is the central principle on which quantum error-correcting codes are founded.

Error correction has been implemented on an NMR quantum computer by Cory *et al.* Phys. Rev. Lett. **81**, 2152 (1998).

11.7 Experimental systems for implementing quantum computing.

The practical requirements for quantum computing have been outlined in section 11.1. These are controllable qubits, at least one controllable two-qubit logic gate, long decoherence times to allow computation and measurement to take place before the system decoheres. A set of 5 require-

ments for practical quantum computation has been developed by DiVincenzo (Fortshritte der Physik –Progress of Physics **48**: 771 (2000)). In this section I give very brief descriptions of experimental systems that have been investigated as potential systems for implementing quantum computing.

11.7.1 Ion traps

Ion traps were the first systems to be investigated for use in quantum computation. Wineland's group uses this system and a number of his papers have been referenced in this course. In the ion trap a combination of electric and magnetic fields allows ions to be trapped virtually isolated from the external environment. This system thus offers long decoherence times. Two internal (electronic) states of the ions are used as the qubits and single qubit operations are implemented via the phenomenon of Rabi oscillations using lasers addressing each ion. Two qubit operations are implemented via phonon modes of the system. The drawback of this is that the phonon timescale is very slow compared to the electronic timescale and so two qubit operations are slow in this system. However, the real problem of this system is scaling to large numbers of qubits. Present ion traps are one-dimensional and cannot accommodate more than a few ions. This system is fine for proof of principle demonstrations but in its present form could never be scaled to an interesting number of qubits for practical quantum computing.

11.7.2 Nuclear Magnetic Resonance

The use of NMR to test the concepts of quantum computing came as a bit of a surprise to many people. In NMR quantum computers the qubits are nuclear spins on different atoms in a molecule and the molecules are in a large magnetic field which splits the energies of the spin states. One qubit operations are performed using microwave pulses tuned to the energy splitting (which is different for each distinct atom in the molecule) and the two qubit operations are performed using the intrinsic spin-spin interactions between the nuclei within the molecule. The surprising thing about NMR systems is that they are at high temperature, which would be expected to cause decoherence, yet they can still be used for quantum computing. The reason is that the nuclear spins are very weakly coupled to the environment. One drawback of this approach is that the measured signals result from the ensemble of nuclear spins and there is intense argument about whether this system really is a genuine quantum computer.

The high degree of sophistication of NMR techniques has allowed this

approach to be used to implement both Grover's algorithm (Jones *et al.* Nature **393**, 344 (1998)) and quantum error correction (Cory *et al.* Phys. Rev. Lett. **81** 2152 (1998)). The problem with this approach is that the measured signal strength decreases exponentially with the number of qubits in the molecule and so these systems are no use as practical quantum computers.

11.7.3 Superconducting systems

Superconducting systems offer the obvious advantage for quantum computing that they have enormously long decoherence times. There are proposals for superconducting quantum computers based on exploiting the Josephson effect (see for example Makhlin *et al.* Nature **398**, 305 (1999)). These systems are relatively easy to control and it is believed that even with existing technology it should be possible to create a system with many tens of qubits. This number of qubits would offer the possibility of demonstrating many ideas in quantum computing but is still not large for useful real world calculations - primarily due to the redundancy required for error correction. Very recently, Nakamura and co-workers have demonstrated 2-qubit gates in superconducting systems (Nature **421**, 823 (2003), Nature **425**, 941 (2003)). This is, of course, significant progress but it has taken a considerable time to achieve control over just two qubits in these systems.

11.7.4 Semiconductor systems

The use of semiconductor based systems as practical quantum computers is being very actively researched at present by many groups, including the Semiconductor Physics Group in the Cavendish. A particularly large collaborative effort is in progress in Australia and America under the direction of Prof. R.G. Clark. There are many proposals for the qubits in a semiconductor based quantum computer. Suggestions include the spins on individual impurity atoms grown in the lattice, the spins on individual electrons confined in quantum dots or pairs of energy levels for electrons confined in quantum dots. There is a belief that with the present rate of progress in fabricating semiconductor devices it will be possible to create systems with the required characteristics for quantum computing within the next 20 years. The main driving force behind this approach is that it appears that this technology could be scaled to very large numbers of qubits.

11.7.5 *Linear optics*

Linear optical techniques are a relatively recent arrival to experimental efforts on quantum computation. One of the earliest proposals for implementing quantum computation was based on encoding each qubit in two optical modes, each containing exactly one photon. However it is extremely difficult to unitarily couple two optical modes containing few photons. Knill, Laflamme and Milburn (Nature **409**, 46-52 (2001)) have proposed a way to circumvent this restriction and implement efficient quantum computation using only passive linear optics, photodetectors, and single photon sources. This efficient linear optical approach to quantum computing is distinct from all other linear optical schemes which are not efficiently scalable. The first experimental research targets are to produce a prototype two qubit gate for photons using linear optics and then to develop a blue-print for a multiple qubit device.