First Steps in Quantum Mechanics[1]

by Kent A. Peacock

Department of Philosophy,
University of Lethbridge.

kent.peacock@uleth.ca

"You just have to be able to drill in very hard wood . . . and keep on thinking beyond the point where thinking begins to hurt."

—WERNER HEISENBERG

"Calculemus. . . "'

—G. W. LEIBNIZ

1 Aim and Level of this Handbook

The purpose of this handbook is to summarize the basic elements of non-relativistic QM in convenient form, and to introduce just enough of the mathematics of QM that one can work through an elementary version of Bell's Theorem, which has been called by H. P. Stapp (with some justice I think) "the most profound result of modern science." Many important parts of QM are omitted here; the object of this handbook is simply to serve as a useful entry point to the theory, keeping in mind that quantum mechanics is something that one understands only by practice. An earlier version of this handbook, together with some exercises on basic linear algebra from Chapter 2 of Hughes (1989), served as the basis for a rather successful introductory graduate course on foundations of quantum mechanics at the University of Western Ontario. We needed only about five three-hour classes to work through all the exercises, taking turns doing them at the board. Then we were able to read and conduct seminars on a number of the central papers on foundations of quantum mechanics without being reduced to simply gaping with admiration at the pretty equations. I later tried the same approach with a senior undergraduate class at the University of Lethbridge, in 2002. I am very grateful to my students in these courses for their patience and enthusiasm, and for their detection of several errors in earlier drafts of this handbook.

I learned this way of presenting the material in a very stimulating graduate seminar on philosophy of physics run by Jim Brown and Bill Seager at the University of Toronto, in (I think) 1988. Working together through a set of progressive problems quickly brings everyone in the group to the same level of ignorance, confusion, and humility.

This is a preliminary version, and it contains several omissions which I now (2008) regard as glaring. However, it can still be used for its original and fairly limited purpose, which was to bring graduate students in philosophy (and interested faculty) with some minimal mathematical fluency to the point at which they can grasp a rigorous presentation of Bell's Theorem.

I would be very grateful for any comments, criticism, or corrections, which could be directed to `kent.peacock@uleth.ca`.

I assume that the reader is familiar with certain basic mathematical concepts, especially the most elementary aspects of linear algebra. Chapter 2 of Hughes is an excellent introduction to the linear algebra needed here, and is tailored to the needs of someone learning QM. There are innumerable texts on basic linear algebra that one could consult. At a minimum, the reader should have some idea what a vector space is, what matrices are and how to manipulate them (how to multiply them especially), a bit of very elementary set theory for one optional section at the end, and something about trigonometry and complex numbers. A smattering of the calculus will be useful, although not indispensible. It is enough, for this handbook, to know that the integral sign is analogous to summation. *Expertise* in any of these subjects is certainly not necessary; we will learn as we go along. Beyond these things, all that is needed is what the German mathematicians used to call *Sitzfleisch*—loosely, "sitting muscles", a willingness to apply the seat of the pants to the seat of the chair and calculate.

Apart from occasional asides meant to tantalize the reader, I make no attempt whatsoever here to engage in interpretation or controversy. These are simply the absolute basics necessary to begin thinking productively about the subject. Riches lie beyond.

2 (Some) Basic Elements of QM

I begin by defining a number of terms, and then summarize the postulates of the theory. I reiterate that this is not a complete presentation of QM. It is meant mainly to give the reader enough mathematical machinery to follow the elementary statement of Bell's Theorem given here.

2.1 State Vector

Most versions of QM take the state vector to be the fundamental object of study. It is an abstract object (often called a *ket*) which is an element of a linear vector space called a Hilbert space. For our purposes, the most important aspect of a Hilbert space is that it is complete; i.e., that a vector is an element of a given Hilbert space if and only if it is a linear combination of the basis vectors of the space. For more precise definitions of Hilbert space and related concepts, see Ballantine (1990) or Redhead (1987). The actual physical system being studied is presumed to be be associated with a so-called state space, which is a Hilbert space which has, as a basis, a set of eigenvectors of a complete set of commuting observables of the system. The nature of the "association" between state vectors and the physical realities they are taken to represent is one of the central interpretational problems.

Kets are written conveniently in Dirac notation as $|\,\psi\,\rangle$, and can be represented in matrix form as a column vector:

$$|\,\psi\,\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \tag{1}$$

where the a's are complex numbers. In this equation, n is the dimension of the state space. One can take a ket to "represent" (in a way that is controversial) a macroscopically or classically observable preparation state or perhaps input state of the system. Every ket can be paired with a "bra" $\langle\,\psi\,|$, which can be represented as a row vector:

$$\langle\,\psi\,| = \begin{pmatrix} a_1^* & a_2^* & \ldots & a_n^* \end{pmatrix}, \tag{2}$$

where a^* is the complex conjugate of a. The bra can be thought of as representing a possible macroscopic observation state or output of the system. The correspondence between kets and bras is antilinear, which means that the bra corresponding to $\lambda_1|\,\phi_1\,\rangle + \lambda_2|\,\phi_2\,\rangle$ is $\lambda_1^*\langle\,\phi_1\,| + \lambda_2^*\langle\,\phi_2\,|$.

The inner, or scalar product of a bra with a ket is symbolized by a "bra-ket" $\langle\,\phi\,|\,\psi\,\rangle$; and is spoken of as the amplitude, or probability amplitude, or transition amplitude (or in field theory, the propagator) to go from state ψ to state ϕ. (Read it from right to left.) The amplitude is a complex number, and can be calculated as the matrix product of the representative matrices:

$$\begin{aligned} \langle\,B\,|\,A\,\rangle &= \begin{pmatrix} b_1^* & b_2^* & \ldots & b_n^* \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} && (3) \\ &= b_1^* a_1 + b_2^* a_2 + \ldots + b_n^* a_n. && (4) \end{aligned}$$

Here are the rules for the scalar product:

$$\begin{aligned} \langle\,\phi\,|\,\psi\,\rangle &= \langle\,\psi\,|\,\phi\,\rangle^* && (5) \\ \langle\,\phi\,|\,\lambda_1\psi_1 + \lambda_2\psi_2\,\rangle &= \lambda_1\langle\,\phi\,|\,\psi_1\,\rangle + \lambda_2\langle\,\phi\,|\,\psi_2\,\rangle && (6) \\ \langle\,\lambda_1\phi_1 + \lambda_2\phi_2\,|\,\psi\,\rangle &= \lambda_1^*\langle\,\phi_1\,|\,\psi\,\rangle + \lambda_2^*\langle\,\phi_2\,|\,\psi\,\rangle && (7) \\ \langle\,\psi\,|\,\psi\,\rangle \text{ is real,} \quad \text{and} &\quad \geq 0, = 0 \text{ iff } |\,\psi\,\rangle = 0. && (8) \end{aligned}$$

The complex number $c = \langle\,\phi\,|\,\psi\,\rangle$ can be interpreted geometrically as the projection of $|\,\psi\,\rangle$ along $|\,\phi\,\rangle$.

Two kets $|\,\psi\,\rangle$ and $|\,\phi\,\rangle$ are said to be *orthogonal* if and only if $\langle\,\phi\,|\,\psi\,\rangle = 0$.

The squared modulus $|\langle\,\psi\,|\,\psi\,\rangle|^2$ is always a real number. Usually we consider only kets which have been *normalized;* i.e., multiplied by a factor of proportionality such that

$$\langle\,\psi\,|\,\psi\,\rangle = 1. \tag{9}$$

The normalized basis vectors of a state space obey the orthonormalization conditions

$$\langle\,i\,|\,j\,\rangle = \delta_{ij}, \tag{10}$$

where the *Kronecker Delta Function* is such that

$$\delta_{ij} = 1 \ \text{ when } i = j, \ \text{ and 0 otherwise.} \tag{11}$$

EXERCISE 1: Let $|\,\psi\,\rangle = C(|\,1\,\rangle + |\,2\,\rangle + |\,3\,\rangle)$, with $\langle\,i\,|\,j\,\rangle = \delta_{ij}$. Calculate the normalization coefficient C.

EXERCISE 2: Consider two amplitudes a_1, a_2, such that

$$\begin{aligned} a_1 &= |a_1|\exp(i\theta_1) \\ a_2 &= |a_2|\exp(i\theta_2) \end{aligned}$$

Show that

$$|a_1 + a_2|^2 = |a_1|^2 + |a_2|^2 + 2|a_1 a_2|\cos(\theta_1 - \theta_2).$$

The third term is called the *interference term.* It depends upon $(\theta_1 - \theta_2)$, which is the phase difference, or relative phase, of the two amplitudes.

Not all systems can be represented by a ket; sometimes a so-called *density operator* is needed. (See below.) A state that can be represented by a ket is called a *pure state.* In the Copenhagen interpretation of QM, the state vector of a pure state is assumed to contain all the information available about the system. It is in this sense that Bohr claimed that QM is complete.

2.2 Operators and Observables

A linear operator is a machine that transforms kets to kets:

$$A|\,\psi\,\rangle = |\,\psi'\,\rangle \tag{12}$$

Since they are linear,

$$A(\lambda_1|\,\psi_1\,\rangle + \lambda_2|\,\psi_2\,\rangle) = \lambda_1|\,\psi_1'\,\rangle + \lambda_2|\,\psi_2'\,\rangle, \tag{13}$$

where the λ's are complex constants (scalars). The product of operators is defined by

$$(AB)|\,\psi\,\rangle = A(B|\,\psi\,\rangle). \tag{14}$$

Calculate from the right. *Order is crucial.* We define a special operator known as the *commutator:*

$$[A, B] \equiv AB - BA. \tag{15}$$

It is not always equal to zero.

EXERCISE 3: Show:

1. $[A, BC] = [A, B]C + B[A, C]$
2. $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0.$

Operators can be represented by matrices. The matrix element between vectors $|\,\psi\,\rangle$ and $|\,\phi\,\rangle$ is $\langle\,\phi\,|\,A\,|\,\psi\,\rangle$. In particular, in a space of basis $\{|\,u_i\,\rangle\}$, the elements of the matrix representation of A will be

$$A_{ij} = \langle\,u_i\,|\,A\,|\,u_j\,\rangle. \tag{16}$$

All operators A have an *adjoint* $A^\dagger$. If $A|\,\psi\,\rangle = |\,\psi'\,\rangle$, then $A^\dagger$ is such that $\langle\,\psi\,|A^\dagger = \langle\,\psi'\,|$.

EXERCISE 4: Show:

1. $\langle\,\psi\,|\,A^\dagger\,|\,\phi\,\rangle = \langle\,\phi\,|\,A\,|\,\psi\,\rangle^*$
2. $(A^\dagger)^\dagger = A$.
3. $(AB)^\dagger = B^\dagger A^\dagger$.

An operator A is *Hermitian* if

$$A = A^\dagger \tag{17}$$

EXERCISE 5: Show that if A is Hermitian,

$$\langle\,\psi\,|\,A\,|\,\phi\,\rangle = \langle\,\phi\,|\,A\,|\,\psi\,\rangle^*.$$

Almost all the operators we will consider are Hermitian.

2.3 Projectors

Consider the expression $|\,\psi\,\rangle\langle\,\psi\,|$. If applied to a ket $|\,\phi\,\rangle$ it gives:

$$|\,\psi\,\rangle\langle\,\psi\,|\,\phi\,\rangle = c|\,\psi\,\rangle, \tag{18}$$

where c is the component of $|\,\phi\,\rangle$ along $|\,\psi\,\rangle$. The object $|\,\psi\,\rangle\langle\,\psi\,|$ is therefore an operator which projects another ket along $|\,\psi\,\rangle$. It is called the projector, or orthogonal projector. Suppose we know a column vector representation of $|\,\psi\,\rangle$:

$$|\,\psi\,\rangle = \begin{pmatrix} a \\ b \end{pmatrix}.$$

Then the matrix representation of $|\,\psi\,\rangle\langle\,\psi\,|$ is given by the so-called *outer product* or *tensor product* of $\langle\,\psi\,|$ and $|\,\psi\,\rangle$ as follows:

$$|\,\psi\,\rangle\langle\,\psi\,| = \begin{pmatrix} a \\ b \end{pmatrix}\begin{pmatrix} a & b \end{pmatrix} = \begin{pmatrix} a^2 & ab \\ ab & b^2 \end{pmatrix}. \tag{19}$$

(We multiply the whole first matrix into each element of the second matrix.)

EXERCISE 6: Let $P_\psi = |\,\psi\,\rangle\langle\,\psi\,|$, where ψ is normalized. Show that $P_\psi^2 = P_\psi$.

The condition

$$P^2 = P \tag{20}$$

is known as *idempotency*, and is often taken as the defining condition of a projector (which is then said to be *idempotent*).

EXERCISE 7: Show that $(|\,u\,\rangle\langle\,v\,|)^\dagger = |\,v\,\rangle\langle\,u\,|$.; hence, show that any projector is Hermitian.

2.4 Eigenvalues, Eigenvectors, and Eigenstates

If an operator A acting on a vector $|\,\psi\,\rangle$ simply produces a vector collinear with $|\,\psi\,\rangle$, as

$$A|\,\psi\,\rangle = \lambda|\,\psi\,\rangle, \tag{21}$$

where λ is a complex constant, then we say that $|\,\psi\,\rangle$ is an *eigenvector* ("proper vector") of A with eigenvalue λ. It is often convenient to symbolize eigenvectors by their eigenvalues, as, for example, $|\,\lambda\,\rangle$. Eigenvectors are also called eigenstates.

EXERCISE 8: Show that the eigenvalues of Hermitian operators are *real.*

EXERCISE 9: Show that if two eigenvectors of a Hermitian operator correspond to two different eigenvalues, then they are orthogonal.

Operators will generally have several eigenvalues. If some of an operator's eigenvalues are equal, they are said to be *degenerate*. This introduces niceties which we will ignore here.

EXERCISE 10: From Dirac (1958), p. 34: consider an operator σ such that $\sigma^2 = 1$. Show that σ has two eigenvalues, ± 1. Show that any ket $|\,\psi\,\rangle$ can be expressed as

$$|\,\psi\,\rangle = \frac{1}{2}(1+\sigma)|\,\psi\,\rangle + \frac{1}{2}(1-\sigma)|\,\psi\,\rangle.$$

Show that unless $|\,\psi\,\rangle = 0$, the two terms on the right are eigenkets of σ with eigenvalues 1 and -1 respectively.

2.5 Spectral Representation

Suppose that any vector in a certain Hilbert space $\mathcal{H}_n$ can be expanded as a linear combination of eigenvectors of some operator belonging to that space:

$$|\,\psi\,\rangle = \sum_{i=1}^{n} c_i|\,i\,\rangle. \tag{22}$$

This is called the spectral decomposition of $|\,\psi\,\rangle$. The set of complex numbers $\{c_i\}$ are called the *components* or *expansion coefficients* of $|\,\psi\,\rangle$ in that state space. The set $\{|\,u_i\,\rangle\}$

is said to span the space if and only if every vector in the space can be expanded as above. Such a set is then said to be *complete*, and is spoken of as an *eigenbasis* for that space. If, in addition, the basis set obeys the relation

$$\langle\, u_i \,|\, u_j \,\rangle = \delta_{ij},$$

it is said to be a complete orthonormal basis set.

The elements of an orthonormal basis set obey the so-called *closure relation:*

$$\sum_{i=1}^{n} |\, u_i \,\rangle\langle\, u_i \,| = I, \tag{23}$$

where n is the dimension of the space and I represents the identity operator. This expression is sometimes called the *spectral decomposition of unity.* It expresses the completeness of the basis set, since it states that an operator formed from the sum of all possible projectors formed from basis vectors simply acts like identity; it projects a vector onto itself.

EXERCISE 11: Suppose that two vectors $|\,\psi\,\rangle$ and $|\,\phi\,\rangle$ in a given space have components $\{c_i\}$ and $\{d_i\}$ respectively, relatively to a given basis $\{|\,u_i\,\rangle\}$. Show that

$$\langle\, \psi \,|\, \phi \,\rangle = \sum_i c_i^* d_i$$

and

$$\langle\, \psi \,|\, \psi \,\rangle = \sum_i |c_i|^2.$$

It is, of course, possible for a state space to have many different bases; we will entirely omit the large question of transformation between different bases.

2.6 Observables

An operator which is Hermitian and whose eigenvectors span the state space is said to be an *observable*. If the set $\{|\,u_i\,\rangle\}$ consists of such eigenvectors, it is called an eigenbasis.

EXERCISE 12: Show that the matrix representation of an observable is diagonal in the eigenbasis of that observable. Show that the diagonal elements are the eigenvalues.

EXERCISE 13: Show that if two operators A and B commute, and if $|\,\psi\,\rangle$ is an eigenvector of A, then $B|\,\psi\,\rangle$ is also an eigenvector of A with the same eigenvalue.

EXERCISE 14: Show that if two observables A and B commute, and if $|\,\psi_1\,\rangle$ and $|\,\psi_2\,\rangle$ are two eigenvectors of A with unequal eigenvalues, then $\langle\, \psi_1 \,|\, B \,|\, \psi_2 \,\rangle = 0$.

It is possible to show that if a number of observables commute, then one can construct an orthonormal basis of the state space with eigenvectors common to those observables. (We'll skip the proof, which involves concepts we don't really need here. See Cohen-Tannoudji *et al.,* p. 140.) We note the notion of a *complete set of commuting observables (CSCO):* roughly, a set of mutually commuting observables whose common eigenvectors are sufficient to span the state space.

2.7 Unitary Operators

The inverse of an operator A, if it has one, is an operator A^{-1} such that

$$AA^{-1} = A^{-1}A = I. \tag{24}$$

An operator is said to be *unitary* if its inverse equals its adjoint:

$$A^{\dagger} = A^{-1}. \tag{25}$$

That is,

$$AA^{\dagger} = A^{\dagger}A = I. \tag{26}$$

Unitary operators conserve the scalar product: let

$$|\,\psi_1'\,\rangle = U|\,\psi_1\,\rangle, \qquad |\,\psi_2'\,\rangle = U|\,\psi_2\,\rangle,$$

where U is unitary. Then

$$\langle\,\psi_1'\,|\,\psi_2'\,\rangle = \langle\,\psi_1\,|\,U^{\dagger}U\,|\,\psi_2\,\rangle = \langle\,\psi_1\,|\,\psi_2\,\rangle. \tag{27}$$

Unitary operators are the generalizations to complex spaces of the orthogonal operators of real-valued vector spaces. Important examples: rotation operators, the time evolution operator.

2.8 Postulates of the Theory

We can summarize the basic claims of QM in a number of postulates. The way we do it here is certainly not the only way to define QM; in fact, some of what we say here (especially regarding the projection postulate) could be controversial.

1. **Quantum of Action.** There exists a universal constant of action, Planck's constant, the value of which is determined experimentally to be

   $$h = 6.626 \times 10^{-27}\text{erg.seconds.}$$

   The action associated with any actual physical system always comes in quanta of magnitude h. It is very important to realize that we have absolutely no idea why action should be quantized, or why the quantum of action takes the particular value that it does. The quantization of action implies that any theory that treats action as a continuous, differentiable variable *has to be* an approximation.

2. **Superposition Principle.** Any linear combination of kets of the state space of a system is also a ket of the state space, and hence represents a possible physical outcome. There are some limitations upon superposition, called *superselection* rules, but we shall not concern ourselves immediately with those here.

3. **Measurables are Observables.** A quantity is physically measurable if and only if it can be represented by an observable: a Hermitian operator. Its eigenvalues are the only possible results that could be obtained upon measurement of the observable. The set of possible eigenvalues of an observable are called its *spectrum*. Some spectra are discrete (e.g., spin states, energy of a closed system), while others are continuous (positions, energy of an open system). In the following we shall carelessly identify measurable physical quantities with the observables that represent them.

4. **Conjugate Quantities.** There exist pairs of conjugate quantities which cannot be simultaneously measured with absolute certainty. The conditions which allow a measurement of one to some extent make impossible the measurement of the other. We say that the measurement of one interferes with the measurement of the other, but one must understand this carefully; it is not exactly that the measurement of one somehow jiggles the apparatus so severely that our registration of the *pre-existing* sharp value of the other is blurred; rather, the measurement of one somehow precludes the sharp definability of the other. Exactly what this means is one of the central problems.

   The observables representing conjugate quantities are *non-commuting*. Examples are position and momentum (in the *same* direction; position and momentum in different directions will commute), and the components of angular momentum in different directions. Conjugate observables obey so-called *canonical commutation relations* such as the following for position and momentum:

   $$[X, P_x] = i\hbar I, \tag{28}$$

   where I is, again, the identity operator. Angular momentum obeys the relations

   $$[J_x, J_y] = i\hbar J_z, \tag{29}$$

   and two similar relations formed by cyclically permuting $\{x, y, z\}$. In quantum mechanics many important quantities are defined in terms of their commutation relations.

   It can be shown that any two non-commuting quantities A and B obey an uncertainty relation

   $$\Delta A \Delta B \geq \hbar. \tag{30}$$

   (See Cohen-Tannoudji, or Redhead, for instance, where this is developed from the commutation relations.)

   Energy and time also obey an uncertainty relation:

   $$\Delta E \Delta t \geq \hbar; \tag{31}$$

   but the meaning of this has been the subject of much discussion, since we do not know how to represent time as an observable in quantum mechanics. This seems odd, since one can, after all, read a clock; but there is no clear consensus on whether this represents a defect in the theory, or is just the way things are.[2]

5. **Time Evolution of the State.** The state evolves with time according to the time-dependent Schrödinger Equation:

$$i\hbar\frac{\partial|\,\psi\,\rangle}{\partial t} = H(t)|\,\psi\,\rangle, \tag{32}$$

where $H(t)$ is the Hamiltonian. This time evolution is unitary—meaning reversible, in effect—and can be described as the action of an operator called the evolution operator upon the state. If the potential is independent of the time, the system exists in so-called *stationary states* which are analogous to the normal modes or standing waves that occur in classical vibrating systems. The various atomic orbitals, for instance, are stationary states. A large part of the physicist's work consists in finding ways to calculate stationary states for various systems.

6. **Reduction Postulate.**

If an observable A is measured and found to have eigenvalue a_j, then the system is immediately thereafter in eigenstate $|\,a_j\,\rangle$. (If A is degenerate, a more complicated statement is required, which we will entirely ignore here.) Reduction is therefore a process like this:

$$\sum_i c_i|\,a_i\,\rangle \rightarrow |\,a_j\,\rangle. \tag{33}$$

Suppose, for instance, that a particle was described by a spherically symmetric wave spreading out in space. If it is detected at a certain point with a small uncertainty, we immediately know that it is an eigenstate of position somewhere within that small range; the spherical wave just disappears; it is said to have "collapsed" or "reduced". There are several difficulties with this idea. Note, first, that it is irreversible and therefore nonunitary. QM therefore countenances two sorts of time evolution of the system, a unitary evolution when the system is not being observed, and a nonunitary evolution when it is disturbed by measurement. The problem, in effect, is to understand what the arrow in the above equation represents. One thing that makes this difficult is that meaning of "immediately" is very unclear in the relativistic context. The collapse is apparently a superluminal process. Much effort has been expended to show that it cannot transport energy or information and

[2] Here is an observation for those who have some familiarity with special relativity. The fact that position and time are unified in the position-time 4-vector suggests that if position is an operator, time should be, too—and also strongly suggests that the lack of a time operator is indeed a defect in the usual theory. The difficulty with finding an acceptable time operator is connected with the fact that position, momentum, time and energy are presumed to be capable of taking on a continuous spectrum of values. Many authors in the field of quantum gravity now suspect that in order to construct a fully quantal theory of space and time we may have to give up that assumption. Also note the interesting fact that the corresponding components of the position-time and momentum-energy 4-vectors are conjugate. It seems obvious that the invariant magnitude of these 4-vectors must be conjugate as well; likely every possible 4-vector would have its conjugate.

therefore does not threaten relativity directly. Whether or not these controversial arguments are correct as far as they go, the whole notion seems unsettling in any case. Furthermore, one gets into severe paradoxes unless one supposes that the collapse occurs in an invariant manner; on the other hand, there is no obviously nonarbitrary way to set up a relativistically invariant reduction postulate.[3]

The reduction postulate is one of the most controversial areas of QM. Many authors argue, for various reasons, that it is incorrect or incoherent, and should be dropped from the theory.

7. **Calculation of Probabilities.** When the quantity A is measured on a system in a normalized state $|\,\psi\,\rangle$, the probability $P(a_n)$ of obtaining the eigenvalue a_n as a result is
$$P(a_n) = |\langle\, a_n\,|\,\psi\,\rangle|^2 \tag{34}$$
where $|\,a_n\,\rangle$ is the normalized eigenvector associated with the eigenvalue a_n. In other words, to calculate the probability of a process having a certain outcome, we take the modulus of the amplitude of that process.

EXERCISE 15: Show that $P(a_n) = \langle\,\psi\,|\,a_n\,\rangle\langle\,a_n\,|\,\psi\,\rangle$.

EXERCISE 16: Show that $\sum_i |\langle\, a_i\,|\,\psi\,\rangle|^2 = 1$; i.e., that probabilities sum to unity, as required. Hint: use the Closure Relation.

Suppose that the system can go through more than one intermediate state to get to its final state. (Think of the double-slit experiment, for instance, in which the particles can go through either slit.) Suppose we observe only the initial and final states. If we could tell, *without further measurement or further disturbing the system*, which intermediate states the particle was in, then we get the probability of the final state by directly summing the probabilities of each alternate route, just as we would with classical probabilities. However, if we could not do this—and this would be the case if observing the intermediate states involved the measurement of observables conjugate to the ones required to measure the final state (which would change the final state we would get)—then we get the probability of the final state by first summing the amplitudes for each possible route and *then* taking the modulus.

Again: let α_1, α_2 by amplitudes for two possible routes to a final outcome a. If the routes are distinguishable, then

$$P(a) = |\alpha_1|^2 + |\alpha_2|^2. \tag{35}$$

However, if the routes are indistinguishable without the measurement of noncommuting quantities, then
$$P(a) = |\alpha_1 + \alpha_2|^2 \tag{36}$$

[3] At least, this is the usual view. A case can be made that wave packets collapse in such a way as to satisfy *phase invariance*; this is a manifestly covariant prescription. See Peacock (2006) for further explorations of this possibility.

Recalling Exercise 2, this means that interference terms between the two possible routes will enter into the final probability. We therefore note that interference—and in effect, all the bizarre properties of QM that flow from it—is fundamentally due to the superposition principle. See Feynman (1963), especially chapter 2, for more on this.

2.9 Expectation Values

The *expectation value* $\langle A \rangle$ is the average over a set of identical measurements of A on identically prepared systems. It is not the same thing as a time average of an observable for one system. The expectation values of observables on systems where h can be taken as negligibly small converge to the classically observable values.

Suppose p_i is the probability of obtaining the i-th eigenvalue a_i. Then

$$\langle A \rangle = \sum_i p_i a_i. \tag{37}$$

EXERCISE 17: Show that $\langle A \rangle = \sum_i |\langle a_i \,|\, \psi \rangle|^2 a_i$, where as usual $|\, a_i \rangle$ is the eigenstate corresponding to the eigenvalue a_i.

EXERCISE 18: Show that $\langle A \rangle = \langle \psi \,|\, A \,|\, \psi \rangle$. Hint: use the Closure Relation.

2.10 Mixed versus Pure States

If we know (or are willing to grant) with certainty that a system was prepared in a state $|\, \psi \rangle$, then we say that it is in a pure state. However, suppose we have only probabilistic knowledge of the preparation states of the system, so that there are probabilities p_k that it is in the corresponding states $|\, \psi_k \rangle$, with $\sum p_k = 1$. Then we say that the system is in a *mixed state*. Particles emerging from an aperture in a furnace, for instance, would be in a mixed state, since we could have only statistical information about their energies, polarizations, etc. Any realistic "large" object, such as Schrödinger's cat, for instance, could in practise only be described by a mixed state. Hence the whimsical description that one sometimes sees of the poor cat as being in a pure state,

$$|\, \mathrm{puss} \rangle = \frac{1}{\sqrt{2}}(|\, \mathrm{alive} \rangle \pm |\, \mathrm{dead} \rangle), \tag{38}$$

is somewhat misleading. A statistical mixture must not be confused with a linear superposition. There is no interference between elements of a mixed state—we treat them using classical probabilities.

2.11 Density Operator Formalism

The density operator provides an alternate way of doing QM which is especially useful for mixed states, since one can state equations for the calculation of probabilities and expectation values in terms of it which are true for both kinds of states. From this point of view, the pure state is a special case of the mixed state.

The density or statistical operator of a pure state $|\,\psi\,\rangle$ is

$$\rho_\psi = |\,\psi\,\rangle\langle\,\psi\,|. \tag{39}$$

This of course says no more than the specification of the ket $|\,\psi\,\rangle$, but writing it this way turns out to be very useful. In a mixed state the density operator is

$$\rho = \sum_k p_k|\,\psi_k\,\rangle\langle\,\psi_k\,|. \tag{40}$$

Note, again, that the p_k are to be understood as classical probabilities. In particular, this means that there is no interference between them.

I will pass over the detailed properties of the density operator, since we don't need it here immediately. See Cohen-Tannoudji *et al.,* (1977) p. 295–307, for a concise summary.

Some authors (notably Ballentine 1990) take the density operator as the basic description of the quantum state, rather than the ket. This is useful for some purposes, but I feel that in a foundational study it tends to obscure the basic difference between quantum and classical probability—namely the existence of *amplitudes* in the calculation of probabilities—and makes quantum mechanics more like classical statistical mechanics than it really is. Nevertheless, the density operator formalism is especially useful—in fact, indispensible—in *quantum statistical mechanics*, in which one deals with systems with enormously many degrees of freedom, so that one could never have more than (classically) probabilistic knowledge of the precise preparations of all components of the system.

2.12 Tensor Products and Composite Systems

It is especially important that we grasp some of the basic properties of the QM representation of composite systems, since this is where some of the most surprising manifestations of nonlocality and nonseparability make their appearance.

Consider two systems 1 and 2—particles, say—which have interacted dynamically. Let their state spaces be $\mathcal{E}_1$ and $\mathcal{E}_2$ respectively. Then the possible states of the composite system of 1 and 2 belong to the *tensor product* space $\mathcal{E}_1 \otimes \mathcal{E}_2$, which we define as follows. If $|\,u_1\,\rangle$ and $|\,u_2\,\rangle$ are base states of $\mathcal{E}_1$ and $\mathcal{E}_2$ respectively, then we formally define their tensor product $|\,u_1\,\rangle \otimes |\,u_2\,\rangle \equiv |\,u_1u_2\,\rangle$ as the object with the properties

$$\begin{aligned}
|\,u_1u_2\,\rangle &= |\,u_2u_1\,\rangle && (41)\\
|\,u_1(\lambda u_2)\,\rangle &= \lambda|\,u_1u_2\,\rangle = |\,(\lambda u_1)u_2\,\rangle && (42)\\
|\,u_1(v_1+v_2)\,\rangle &= |\,u_1v_1\,\rangle + |\,u_1v_2\,\rangle. && (43)
\end{aligned}$$

A simple notation is highly desirable here. In the following, we will assume that in a state such as $|\,ab\,\rangle$, $|\,a\,\rangle$ denotes a state of particle 1 and $|\,b\,\rangle$ denotes a state of particle 2. It is sometime also useful to write this as $|\,ab\,\rangle = |\,a\,\rangle|\,b\,\rangle$. These conventions avoid a proliferation of subscripts. The product space $\mathcal{E}_1 \otimes \mathcal{E}_2$ then consists of all possible linear combinations of the base states $\{| u_{1_i} u_{2_j} \rangle\}$. The superposition principle states that any of these are, in general, possible states of the system.

Operators defined on $\mathcal{E}_1$ or $\mathcal{E}_2$ can be extended to $\mathcal{E}_1 \otimes \mathcal{E}_2$ as well. We need to understand how to calculate the expectation value of such an operator.

First, the scalar product in $\mathcal{E}_1 \otimes \mathcal{E}_2$: the scalar product of $| ab \rangle$ with $| cd \rangle$ is given by

$$\langle cd | ab \rangle = \langle c | a \rangle \langle d | b \rangle. \tag{44}$$

Now, let A_1 and B_2 be operators acting in $\mathcal{E}_1$ and $\mathcal{E}_2$ respectively. Then their extension to the product space is an operator $A_1 \otimes B_2$ which acts on the elements of the product space as follows:

$$(A_1 \otimes B_2) | ab \rangle = A_1 | a \rangle \otimes B_2 | b \rangle. \tag{45}$$

The expectation value of $A_1 \otimes B_2$ is given by

$$\langle ab | A_1 \otimes B_2 | ab \rangle = \langle a | A_1 | a \rangle \langle b | B_2 | b \rangle. \tag{46}$$

EXERCISE 19: Show this.

If, instead of a simple product $| ab \rangle$, we had a vector of the form

$$| \psi \rangle = | ab \rangle + | cd \rangle,$$

what would we get? First, for simplicity in notation, write $A_1 \otimes B_2 = C$. Then we get

$$\begin{aligned} \langle \psi | A_1 \otimes B_2 | \psi \rangle &= \langle \psi | C | \psi \rangle & (47) \\ &= (\langle ab | + \langle cd |) C (| ab \rangle + | cd \rangle) & (48) \\ &= \langle ab | C | ab \rangle + \langle cd | C | ab \rangle + \langle ab | C | cd \rangle + \langle cd | C | cd \rangle. & (49) \end{aligned}$$

Then expand the four matrix elements as above (Eq. 46). This will be useful in the derivation of Bell's Theorem below.

2.13 Factorizability

Consider a physical consisting of two subsystems α and β, which may be presumed to have interacted, or still be interacting. A ket belonging to the state space of the combined system is said to be non-factorizable if it cannot be written as a simple product

$$| \Psi \rangle = c | i \rangle \otimes | j \rangle, \tag{50}$$

where $| i \rangle$ and $| j \rangle$ would be states of α and β respectively, and c is some complex constant. A state which can be factored this way is also called a product state. A non-factorizable ket can only be written as

$$| \Psi \rangle = \sum_{i,j} c_{ij} | i \rangle \otimes | j \rangle \tag{51}$$

where the c_{ij} are complex coefficients. The crucial point is that the most general element of the product space $\mathcal{E}_1 \otimes \mathcal{E}_2$ has this form. Such states are said to be "entangled" (Schrödinger's terminology); since the state of one component inevitably becomes enmeshed in the calculation of probabilities for quantities to be measured on the other component. The measurable properties of spatially separate particles in an entangled state can be correlated *even if* the particles are no longer dynamically interacting.

EXERCISE 20: Show that the singlet state,

$$|\,\Psi_S\,\rangle = 1/\sqrt{2}(|\,+-\,\rangle - |\,-+\,\rangle), \tag{52}$$

is non-factorizable. That is, prove we cannot write it in the form $|\,\Psi_S\,\rangle = |\,1\,\rangle|\,2\,\rangle$, where $|\,1\,\rangle$ and $|\,2\,\rangle$ are pure states of the two particles. Hint: write $|\,1\,\rangle$ and $|\,2\,\rangle$ as superpositions of two spin states,

$$\begin{aligned} |\,1\,\rangle &= a|\,1(+)\,\rangle + b|\,1(-)\,\rangle && (53)\\ |\,2\,\rangle &= c|\,2(+)\,\rangle + d|\,2(-)\,\rangle && (54) \end{aligned}$$

and show that there are no possible values of a, b, c, or d which can satisfy the definition of the singlet state.

The violation of Bell's Inequality in the singlet state is not due to interference between putative localized states of the two spatially separate particles—in fact, it can be shown that they do not exist in separate pure states (see Page's paper)—but is due to the interference between the *nonlocal* states $|\,+-\,\rangle$ and $|\,-+\,\rangle$.

2.14 Some Spin Basics

We will describe just enough of the mathematics of spin to enable us to make the essential calculations for Bell's Theorem. There is more, much more, than I indicate here.

What is spin? Good question; Pauli called it a "non-classical two-valuedness" associated with particles. Like many things in QM, it can be given a very clear mathematical specification, but is almost impossible to interpret. Spin is a quantity that has units of angular momentum (or equivalently, action) but which cannot be transformed away by a change of coordinates, the way orbital angular momentum could be (at least in principle). Spin is generally measured in units of $\hbar = h/2\pi$.

EXERCISE 21: Show that angular momentum and action are dimensionally equivalent.

One can loosely imagine spin as an internal rotation state of a particle; it is rather as if each elementary particle were like a tiny spinning magnet. However, remember that the truly elementary particles (leptons, quarks, photons) have no presently measurable diameter; furthermore, we have no idea how to transform to the "internal rest frame" of an elementary particle; so that a picture of a particle with spin as a rotating object is dubious at best.

Spin gives a particle a magnetic moment and hence is manifested through the interaction of particles with an external magnetic fields. A particle with spin will be deflected in certain specific directions in a magnetic field; spin-1/2 particles such as electrons will be deflected either up or down in the direction of an external applied field. Spin exhibits a curious property known as "space quantization". This means that the particle will always be found to have spin only in these specific directions, completely regardless of its initial state of motion. To grasp how odd this is, realize that the external magnetic field can be applied in any direction; nevertheless, the particle will always deflect either up or down with respect to that field. This strongly suggests that the particle did not have an intrinsic angular momentum before the measurement, in the sense that a classical object (say a spinning baseball) would; one could more accurately say that the particle had only a potential or possibility to be in various spin states; the possibility becomes actualized upon the interaction. At least, this is the way many people find themselves speaking of quantum phenomena; it is not clear how meaningful it really is to talk of the quantum state as representing "potentia" or "possibilia".

Particles with integral spin (spin $= \pm n\hbar,\ n = 0, 1, \ldots$) are called *bosons* because they obey so-called Bose-Einstein statistics; particles with half-integral spin (spin $= \pm(2n+1)/2\,\hbar, n = 0, 1, \ldots$) are called *fermions* because they obey so-called Fermi-Dirac statistics. Light quanta (photons) are bosons (with spin 1), electrons are fermions (spin 1/2). Fermi-Dirac particles obey the Exclusion Principle, which means that no two fermions can be in exactly the same quantum state. Bose-Einstein particles, on the other hand, tend to congregate in the same energy state, which can lead to so-called Bose-Einstein condensation and amazing macroscopic quantum behavior such as superfluidity and superconductivity. Roughly speaking, one can think of "solid" matter as constructed of fermions, and mediating fields as constituted of quanta of bosons.

Spin is represented by a vector operator

$$\hat{\mathbf{S}} = \frac{\hbar}{2}\hat{\sigma} = \frac{\hbar}{2}(\sigma_x, \sigma_y, \sigma_z) \tag{55}$$

with

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{56}$$

These are the Pauli matrices.

EXERCISE 22: Calculate the matrix representing $\hat{\mathbf{S}}^2 = \hat{\mathbf{S}} \cdot \hat{\mathbf{S}}$.

The Pauli matrices obey the commutation relations

$$[\sigma_x, \sigma_y] = 2i\sigma_z, \tag{57}$$

and two similar relations obtained by cyclic permutation of $\{x, y, z\}$.

EXERCISE 23: Show this by explicit matrix multiplication.

EXERCISE 24: Show that $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I$

EXERCISE 25: Show that $\sigma_x\sigma_y = -\sigma_y\sigma_x = i\sigma_z$.

EXERCISE 26: Show that $[\hat{\sigma}^2, \sigma_z] = 0$.

Because they commute, $\hat{\sigma}^2$ and σ_z share a common eigenbasis $\{|+\rangle, |-\rangle\}$, which can be represented as column vectors

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{58}$$

$$|-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{59}$$

Matrix multiplication shows that

$$\sigma_z|+\rangle = |+\rangle, \tag{60}$$
$$\sigma_z|-\rangle = -|-\rangle. \tag{61}$$

That is, $|\pm\rangle$ are eigenstates of σ_z with eigenvalues ± 1 units of $\hbar/2$.

EXERCISE 27: Carry out the multiplication to show this.

It can also be shown that

$$\sigma_x|+\rangle = |-\rangle \tag{62}$$
$$\sigma_x|-\rangle = |+\rangle \tag{63}$$

and

$$\sigma_y|+\rangle = i|-\rangle \tag{64}$$
$$\sigma_y|+\rangle = -i|-\rangle \tag{65}$$

EXERCISE 28: Do these too.

These kets obey the orthnormalization relations

$$\langle +|+\rangle = \langle -|-\rangle = 1, \tag{66}$$
$$\langle +|-\rangle = \langle -|+\rangle = 0. \tag{67}$$

EXERCISE 29: And these.

Note that $|\pm\rangle$ are spin states of a single particle. A particle's total state is just a direct tensor product of its spin state with a state giving its position and momentum. Note that it is always assumed that there is no entanglement between spin and spacetime states. This is an example of a superselection rule. That is why we can talk about spin states of particles, without mentioning the other aspects of their state. It is not absolutely clear that this is physically justified; however, entanglement between spin and position would be nonlocality with a vengeance.

If $|+\rangle$ and $|-\rangle$ are the spin states of individual spin-1/2 particles such as electrons, the tensor product spin space is spanned by

$$|+-\rangle = |+\rangle \otimes |-\rangle \tag{68}$$
$$|-+\rangle = |-\rangle \otimes |+\rangle. \tag{69}$$

Physically, $|+-\rangle$ represents a state in which one particle (say on the left) would be found to have spin in the z-direction to be up, and the right particle would be found to be spin down. States such as $|++\rangle$ would be impossible for fermions, by the Exclusion Principle. The general elements of the spin space will be superpositions of these two base states.

2.15 Some Other Terminology

I briefly discuss a few other terms, before proceeding to the main topic.

2.15.1 Wave Functions

The wave function is the form of the state vector $|\,\Psi\,\rangle$ when expressed in a continuous representation, generally either position or momentum. When projected into the $|\,r\,\rangle$ or position representative, the wave function is

$$\langle\, r\,|\,\Psi\,\rangle = \Psi(r) \quad \text{where} \;\; |\,r\,\rangle = \delta(r - r_0). \tag{70}$$

The simplest wave function is the 1-dimensional pure plane wave,

$$\langle\, x\,|\,p\,\rangle = \frac{1}{\sqrt{2\pi\hbar}} \exp{(ipx/\hbar)}. \tag{71}$$

I mention this mainly to make it clear that the state vector and the wave function are not the same thing; think of a wave function as the projection of the state vector either into ordinary spacetime, or the conjugate relativistic position-momentum space.

2.15.2 Hidden Variable Theories

A hidden variable theory is any theory which claims that the probabilistic results of QM are determined by an underlying mechanism or reality which would be described in terms of *real-valued* quantities which are presently hidden from our view.

A *local* hidden variable theory says that the underlying reality exhibits no action at a distance; or equivalently, that the underlying reality is a structure living within relativistic spacetime. Classical electromagnetic theory is a local theory *par excellence*.

A *nonlocal* hidden variable theory says that the underlying reality is not primarily spatio-temporal; the apparent locality of many physical phenomena is (as Leibniz might have said) a “well-founded phenomenon”. Nonlocal h.v. theories allow instantaneous (i.e., superluminal) causation; an example is the Bohm theory in which quantum correlations are mediated by the nonlocal quantum potential. A contextual h.v. theory says that the results of observations are completely determined by variables which pertain both to the measuring apparatus and the system being studied. Bell, Shimony, and others have shown that a nonlocal contextual h.v. theory may is consistent with QM (although there are problems with relativity!)

A *stochastic* h.v. theory says that the hidden variables λ may be subject to random variations or may not be knowable with certainty; hence all functions of them should be weighted by a classical probability density $\rho(\lambda)$.

2.15.3 Dispersion

This term is used variously in the literature. Properly speaking it refers to the tendency for a wave packet to spread with time. It is also used to refer to the measure of uncertainty ΔA of an observable A, defined as

$$\Delta A = \sqrt{\mathrm{var}(A)} = \langle\,(A - \langle\, A\,\rangle)^2\,\rangle^{\frac{1}{2}}. \tag{72}$$

QM states that no system can be observed without uncertainty in some of its attributes. Hidden variables theories postulate that there are "dispersion free" states over which the predictions of QM are some sort of average.

3 A Standard Derivation of Bell's Theorem

In this section we run through a more or less stock derivation of Bell's Theorem, much in the vein of a presentation given by Selleri (1990). There are three parts to this argument, as there are in any version of a Bell's Theorem: derivation of the classical result; derivation of the quantum result; and comparison of the two. Probably the most elementary derivation of Bell's Theorem that is still mathematically rigorous is the ingenious presentation by Tim Maudlin (1994).

We will consider physical systems α and β which have interacted dynamically in the past, and which will accordingly be described by a non-factorizable ket of the tensor product space $\mathcal{H}_{\alpha\beta}$, where

$$\mathcal{H}_{\alpha\beta} = \mathcal{H}_\alpha \otimes \mathcal{H}_\beta, \tag{73}$$

$\mathcal{H}_\alpha$ and $\mathcal{H}_\beta$ being the state spaces of α, β respectively. An example of such an "entangled" system would be a particle ϵ which decays to α and β:

$$\epsilon \rightarrow \alpha + \beta.$$

Let there be N decays, N measurements of an observable $A(\alpha)$ on α, and N corresponding measurements of an observable $B(\beta)$ on β. (In other words, we assume ideal detector efficiency. Some, such as Arthur Fine, have suggested that this may be illegitimate.) We assume that α and β are far enough apart that the time required for each measurement is small in comparison with the time required for light to travel between them. The quantities a and b are parameters such as detector angle ("knob settings") that can be varied locally at α and β respectively.

We design the experiment so that A and B are dichotomic quantities. That is, we assume that the results of measuring them are yes/no answers, which can be represented as having possible values ± 1. Almost any conceivable measurement can be cast in these terms. (For instance, we could measure position by asking whether something was to the left or right of a reference point.)

The observer at α will obtain a set of results $\{A_1, A_2, \ldots, A_N\}$, and the observer at β will obtain results $\{B_1, B_2, \ldots, B_N\}$, with A_i, B_i pertaining to the i-th decay. We want to study the correlation between α and β results, so we define a correlation coefficient

$$P(a,b) \equiv \frac{1}{N}\sum_{i=1}^{N} A_i(a)B_i(b). \tag{74}$$

Clearly, $|P(a,b)| \leq 1$.

EXERCISE 30: Show this.

EXERCISE 31: Let Y be the fraction of results that agree (i.e., the fraction of cases in which $A_i = B_i$) and N the fraction of cases that disagree. Show that

$$P(a,b) = Y - N.$$

Now, we need some way of comparing the QM and local realism predictions about correlations. One way we can do this is to define a quantity Δ:

$$\Delta \equiv |P(a,b) - P(a,b')| + |P(a',b) + P(a',b')| \tag{75}$$

where a, a' and b, b' are knob settings at α and β respectively. We will calculate Δ by local realism, and by QM, and show that for some detector parameters they disagree.

3.1 Correlations According to Local Realism

Local realism postulates that the values of observable quantities are determined by some "hidden" parameters λ such that there is a strict functional dependence of the observables upon λ. Accordingly, we write the A's and B's as $A(a,\lambda)$, $B(b,\lambda)$, to indicate this dependence. Note, however, that A is presumed *not* to be a function of b, nor is B a function of a. This expresses the view that α and β are, indeed, *separate* physical systems.

For generality, we also assume that we have only probabilistic knowledge of λ; in other words, it is a so-called "stochastic" hidden variable. Then our results should be weighted by a probability density $\rho(\lambda)$, obeying the normalization condition

$$\int d\lambda \rho(\lambda) = 1. \tag{76}$$

In these terms, the correlation coefficient will be

$$P(a,b) = \int d\lambda \rho(\lambda) A(a,\lambda) B(b,\lambda). \tag{77}$$

EXERCISE 32: Be sure that you see why this is so.

Then we can show that

$$|P(a,b) - P(a,b')| \leq \int d\lambda \rho(\lambda) |B(b,\lambda) - B(b',\lambda)|. \tag{78}$$

EXERCISE 33: Show this. Hint: use $|\int dx\, f(x)| \leq \int dx |f(x)|$ and the fact that $|A(a,\lambda)| = 1$.

Similarly we get

$$|P(a',b) + P(a',b')| \leq \int d\lambda \rho(\lambda) |B(b,\lambda) + B(b',\lambda)|. \tag{79}$$

Adding these inequalities, we get

$$\Delta_{LR} \leq \int d\lambda\rho(\lambda)|B(b,\lambda) - B(b',\lambda)| + |B(b,\lambda) + B(b',\lambda)|, \tag{80}$$

where Δ_{LR} means Δ according to local realism. Now, we can show that

$$|B(b,\lambda) - B(b',\lambda)| + |B(b,\lambda) + B(b',\lambda)| = 2. \tag{81}$$

EXERCISE 34: Do this. Hint: consider cases.

Thus we finally get

$$\Delta_{LR} \leq 2. \tag{82}$$

EXERCISE 35: Show this. Hint: use the normalization condition.

This is actually the Clauser-Horne-Shimony-Holt version of the Bell Inequality; Bell himself (1964) considered a case in which there were only three detector settings.

Note the essential irrelevance of λ. What makes the above calculation work is the separability—i.e., the factorizability—of the A's and B's.

3.2 The Quantum Prediction

Now we calculate Δ_{QM}. Be sure you have worked through the section above on spin.

Suppose that α and β are spin-1/2 particles such as the electron. We assume that they are emitted from a source with spin 0, and that their spins are measured at spatially remote locations (by a device known as a Stern-Gerlach detector). It can be shown (see Cohen-Tannoudji) that the total, or global, state of the system is given by the so-called singlet state

$$|\,\Psi_S\,\rangle = 1/\sqrt{2}(|\,+-\,\rangle - |\,-+\,\rangle). \tag{83}$$

Measurements of spin in any direction a are given by the projection of the spin operator on a unit vector in that direction:

$$\text{spin in direction } a = \hat{\sigma} \cdot \hat{a}. \tag{84}$$

We describe the product of results $A(a)B(b)$ as a tensor product operator $\hat{\sigma}_\alpha \cdot \hat{a} \otimes \hat{\sigma}_\beta \cdot \hat{b}$, where $\hat{\sigma}_\alpha$, $\hat{\sigma}_\beta$, are spin operators in $\mathcal{H}_\alpha$ and $\mathcal{H}_\beta$ respectively. Then the correlation coefficient we seek is the expectation value of this product:

$$P(a,b)_{QM} = \langle\,\Psi_S\,|\,\hat{\sigma}_\alpha \cdot \hat{a} \otimes \hat{\sigma}_\beta \cdot \hat{b}\,|\,\Psi_S\,\rangle. \tag{85}$$

We set up our apparatus and coordinates so that $\hat{a}$ is in the z direction, and $\hat{b}$ is in the xz plane. Then we get

$$\hat{\sigma}_\alpha \cdot \hat{a} = \sigma_z \tag{86}$$

$$\hat{\sigma}_\beta \cdot \hat{b} = \sigma_z \cos\theta + \sigma_x \sin\theta \tag{87}$$

where θ is the angle between $\hat{a}$ and $\hat{b}$.

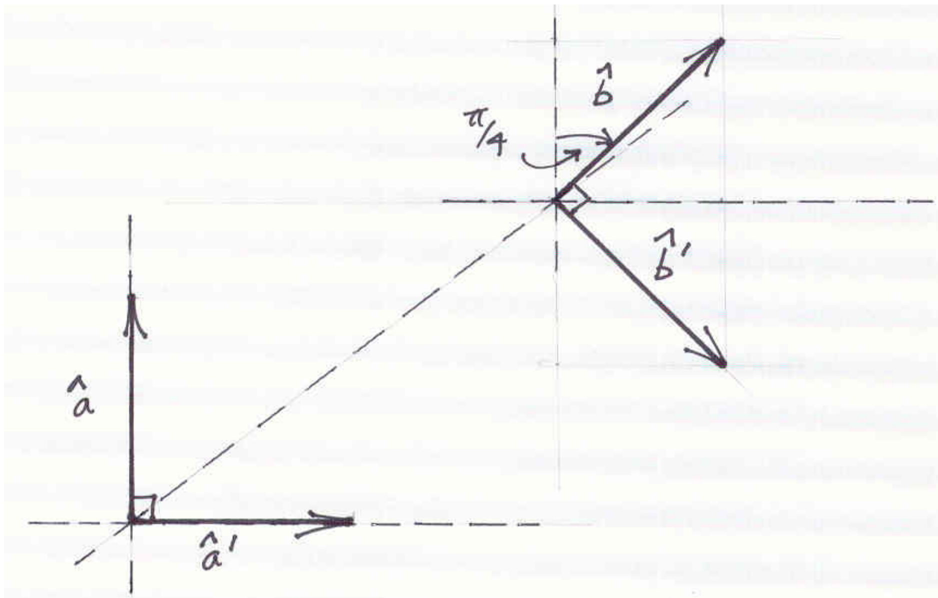

Figure 1: The Setup

EXERCISE 36: Confirm this.

For convenience in calculation, define

$$\sigma_z = A \tag{88}$$
$$\sigma_z \cos\theta + \sigma_x \sin\theta = B. \tag{89}$$

Then we can show that

$$P(a,b)_{QM} = \langle \Psi_S \,|\, A \otimes B \,|\, \Psi_S \rangle \tag{90}$$
$$= -\cos\theta. \tag{91}$$

EXERCISE 37: Do it.

3.3 The Contradiction

To compare Δ_{LR} and Δ_{QM}, we can set up our apparatus so that $\hat{a}$ and $\hat{a}'$, and $\hat{b}$ and $\hat{b}'$ are mutually orthogonal, but $\theta_{\hat{a}\hat{b}} = \pi/4$.

EXERCISE 38: Show that for this arrangement we get $\Delta_{QM} = 2\sqrt{2}$.

Thus, for at least some settings of an EPR-Bohm apparatus, the classical prediction is violated.

The most convincing verification to date of the QM prediction was the experiment by Aspect *et al.* (1981). They did not actually use the singlet state, but a similar entangled state.

4 Could There be a Bell Telephone?

A hypothetical "Bell Telephone" would be some sort of quantum mechanical device which could allow a localized experimenter to superluminally control distant effects by means of quantum mechanical nonlocality. The orthodox view, which has been argued for by many authors, is that a general prohibition against controllable superluminal effects is a consequence of the postulates of quantum mechanics, or at least of local quantum field theory (which arguably is less general than abstract quantum theory as such). (See e.g., Eberhard and Ross 1989.) Abner Shimony in several publications has thus ironically described the relationship between quantum mechanics and relativity theory to be one of "peaceful coexistence," and has described quantum nonlocality not as a form of action at a distance, but "passion" at a distance. At least three authors—the present writer, J. B. Kennedy, Jr., and Peter Mittelstaedt—have criticized the orthodox view on the grounds that it is question-begging: that is, on the grounds that the standard no-signalling arguments rely on specialized locality assumptions of one form or another which undercut the generality of the proofs. (See Peacock 1992, Kennedy 1995, Mittelstaedt 1998, and Peacock and Hepburn 2000; and references therein.) Recent work by Aharonov *et al.* (2004) suggests that it might be possible to bring about controllable superluminal signalling using quantum non-demolition experiments, to which the standard no-signalling arguments are not applicable. This question remains intensely controversial, and a proper treatment of it is decidedly beyond the scope of these notes.

5 A Bell Inequality from Set Theory

To conclude, we will work through a particularly apposite derivation of a Bell's Inequality due to K. C. Hannabuss (1988, p. 107–108). Let A, B, C be three events and A', B', and C' be their complements (i.e., A' is just the event of A not happening, etc.) Let $P(A)$ be the probability of A, and so on. Now define

$$\begin{aligned} P_{AB} &= P(A' \cap B) + P(A \cap B') \qquad (92) \\ &= \text{Prob}(A \text{ or } B \text{ but not both}). \end{aligned}$$

Then:

$$\begin{aligned} & P_{AB} + P_{BC} \\ = \;& P(A' \cap B) + P(A \cap B') + P(B' \cap C) + P(B \cap C') \qquad (93) \\ \geq \;& P(A' \cap B \cap C) + P(A \cap B' \cap C') + P(B' \cap C \cap A') + P(B \cap C' \cap A) \qquad (94) \\ = \;& P(A' \cap C) + P(A \cap C') \qquad (95) \\ = \;& P_{AC}. \qquad (96) \end{aligned}$$

In summary: $P_{AB} + P_{BC} \geq P_{AC}$.

EXERCISE 39: Explain all the above steps.

We can define complementary probabilities $Q_{AB} = 1 - P_{AB}$ and write the inequality as

$$1 + Q_{AC} \geq Q_{AB} + Q_{BC}. \tag{97}$$

EXERCISE 40: Show this.

EXERCISE 41: Show that $Q_{AB} = \text{Prob(results agree)}$.

Eq. (97) is a version of Bell's Inequality. Again, it can be shown to be violated by a variety of entangled quantum systems. One could thus say that Bell's Theorem amounts to the statement that quantum mechanics violates set theory. Cantor (quoted in Kamke (1950), p. 1) defined a set as "a collection into a whole, of definite, well-distinguished objects ... of our perception or of our thoughts." The properties of quantum mechanical systems cannot in general, therefore, be collected into well-defined wholes of well-distinguished objects. More precisely, as Bub (1997) puts it, there is no consistent valuation of all of the possible experimental questions which could be asked of a quantum-mechanical system. Bub and other recent authors thus insist that the basis of quantum mechanics, if there could be such a thing, has to be non-Boolean. (The statement that quantum mechanics is non-Boolean is called the Kochen-Specker Theorem; see Hooker (ed.) 1975, Bub 1997.) Pitowsky (1994) shows that the Bell Inequalities are simply special cases of a class of inequalities between frequencies first defined by George Boole in the 1850s; the Boole inequalities amount to consistency conditions, and were dubbed by Boole 'conditions of possible experience.'

One way of expressing the difference between set theory and quantum mechanics is as follows: classically, if A and B are disjoint events, we have

$$P(A \cup B) = P(A) + P(B) = \text{Prob}(A \text{ or } B \text{ but not both}),$$

while in quantum mechanics we have

$$P(A \cup B) = P(A) + P(B) + \text{interference terms}.$$

This is due to the fact—which we fundamentally do not understand—that in QM there is an *extra step* involved in calculating probabilities, namely the summation of complex amplitudes. (Recall Exercise 2.) Quantum probability reduces to classical probability in the limit in which the interference terms approach zero. Entanglement itself is, as noted earlier, from the mathematical point of view simply the effect of interference between multiparticle states such as $|+-\rangle$ and $|-+\rangle$. Feynman may have been right, therefore, when he hinted (see e.g. Feynman *et al.* 1963) that interference is the fundamental mystery of quantum mechanics.

Bibliography

This bibliography is meant to serve not as an exhaustive reference in foundations of QM—that would run to hundreds of items—but as a guide and entry-point to the subject for philosophers of science wishing to do some fairly serious homework in the subject.

Texts

Ballantine, Leslie E. (1990), *Quantum Mechanics*. Prentice-Hall, Englewood Cliffs, New Jersey.
An extremely useful, thorough and up-to-date mid-level text which is marred by its tendentious presentation of measurement theory entirely in terms of Ballantine's own controversial ensemble interpretation. Very good on properties of correlated systems.

Chester, Marvin (1987), *A Primer of Quantum Mechanics*. Wiley, New York.
Very clear introduction, in spirit of Feynman's interpretation of QM. Demands good background.

Cohen-Tannoudji, C., Diu, B., & Laloë, F., (1977), *Quantum Mechanics* (Two Volumes), John Wiley and Sons, New York.
A really excellent, almost encyclopedic mid- to senior undergraduate text; brings one to the point where one could begin to study quantum electrodynamics. The treatment of entangled states is very thorough.

Dirac, P. A. M. (1958), *The Principles of Quantum Mechanics (Fourth Edition)*, Oxford, the Clarendon Press.
One of the great books.

Feynman, R. P., Leighton, R. L., & Sands, M. (1963), *The Feynman Lectures in Physics, Vol. III.* Addison-Wesley, Reading, Mass.
The finest "introductory" text around. Suffused with Feynman's unique interpretation and physical intuition. Must reading.

Gibbins, Peter (1987), *Particles and Paradoxes: the Limits of Quantum Logic.* Cambridge University Press.
Not really a text, but contains much very useful exposition on basics of quantum mechanics and quantum logic. Very useful complement to this handbook.

Hannabuss, K. C. (1988), *An Introduction to Quantum Mechanics*, University of Oxford, Mathematical Institute.
Excellent, concise, and original exposition of the basics for the serious student.

Kamke, E. (1950). *Theory of Sets*. New York: Dover.

Lawrie, Ian D. (1990), *A Unified Grand Tour of Theoretical Physics.* Adam Hilger, Bristol and New York.
Really excellent, meaty overview (with problems) of the main areas of theoretical physics, including QM and quantum field theory.

Liboff, Richard L., (1992), *Introductory Quantum Mechanics (2nd Edition)*, Addison Wesley Pub. Co., Reading, Mass.
Somewhat more elementary than Cohen-Tannoudji *et al.*, but extremely clear and thorough (except for odd omission of tensor spaces). Excellent collection of problems, many solved; recommended for those who wish to improve their grasp of mathematical technique in particular.

Philosophical, Interpretative, Exploratory

Aharonov, Y., Anandan, J., Maclay, G.. J., and Suzuki, J. (2004) "Model for Entangled States with Spin-Spin Interactions," *Physical Review A* 70, 052114.

Aspect, A., Grangier, P., and Roger, G. (1981) "Experimental tests of realistic local theories via Bell's Theorem." *Physical Review Letters* 47(1), 460–463.

Bell, J. S. (1987), *Speakable and Unspeakable in Quantum Mechanics.* Cambridge University Press.
Basic writings by the master. Indispensible.

Bohm, David, and Hiley, B. J. (1993), *The Undivided Universe: An ontological interpretation of quantum theory.* Routledge, London and New York.
David Bohm is the originator of the most widely-studied hidden variable theory. This book presents the most thoroughly worked-out version of Bohm's alternative to the orthodox view. Very thought-provoking.

Bub, Jeffrey (1974), *The Interpretation of Quantum Mechanics.* Dordrecht, Boston.
Excellent, rigorous introduction to the basic problems, from a quantum logic point of view.

Bub, Jeffrey (1997), *Interpreting the Quantum World.* Cambridge University Press.
Authoritative and detailed review of recent work in foundations of quantum mechanics, from the viewpoint of the modal interpretation; essential reading for the serious student. This book contains detailed proofs and discussion of the Kochen-Specker Theorem and Bell's theorem, and advances the view that quantum mechanics is non-Boolean.

Cini, M., & Lévy-Leblond, J.-M. (eds.) (1990), *Quantum Theory Without Reduction.* Adam Hilger, Bristol and New York.
Some recent attempts to do away with the reduction postulate.

Cushing, J. T., & McMullin, E. (eds.) (1989), *Philosophical Foundations of Quantum Theory: Reflections on Bell's Theorem.* University of Notre Dame Press, Notre Dame .
Recent collection of papers, most not too technical, most very good.

d'Espagnat, B. ([1976] 1989), *Conceptual Foundations of Quantum Mechanics: Second Edition.* Addison-Wesley, Reading, Mass.
Good, although rather conservative, discussion of the implications of nonseparability, at a high level.

Eberhard, P., and Ross, R. R. (1989). "Quantum Field Theory Cannot Provide Faster-than-light Communication," *Foundations of Physics Letters* 2, pp. 127–149.
A thorough presentation of the orthodox "no-signalling" view.

Healey, Richard (1989), *The Philosophy of Quantum Mechanics: An Interactive Interpretation.* Cambridge University Press.

Hodgson, D. H. (1991), *The Mind Matters: Consciousness and Choice in a Quantum World.* Clarendon Press, Oxford.
Explores mental/physical link. An excellent expository chapter on QM.

Hooker, C. A. (ed), (1975), *The Logico-Algebraic Approach to Quantum Mechanics.* D. Reidel Pub. Co., Dordrecht and Boston.
Contains many classic papers on this approach to QM, including Kochen and Specker.

Hughes, R. I. G. (1989), *The Structure and Interpretation of Quantum Mechanics.* Harvard University Press, Cambridge, Mass.
Excellent, serious introduction to the subject; together with Hughes' attempted solution to the mystery.

Kennedy, J. B. Jr., "On the Empirical Foundations of the Quantum No-Signalling Proofs." *Philosophy of Science* 62, 543–560.

Maudlin, Tim (1994), *Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics.* Blackwell, Oxford.
Very clear and challenging overview, insisting that we must accept that Bell's Theorem entails the reality of superluminal causation.

Mittelstaedt, P. (1998). "Can EPR-correlations be used for the transmission of superluminal signals?" *Annalen der Physik* 7, 710–715.

Page, D. N. (1982). "The Einstein-Podolsky-Rosen Physical Reality is completely described by quantum mechanics." *Physics Letters* 91A(2), 57–60.

Peacock, Kent A. (1992). "Comment on 'Tests of Signal Locality and Einstein-Bell Locality for Multiparticle Systems'." *Physical Review Letters* 69(18), 2733-2724.

Peacock, Kent A. (2006), "Temporal Presentness and the Dynamics of Spacetime," in D. Dieks (ed.), *The Ontology of Spacetime*, Elsevier, pp. 247–261.

Peacock, Kent A., and Hepburn, B. S. (2000), "Begging the Signalling Question: Quantum signalling and the dynamics of multiparticle systems." Pre-print at arXiv:quant-ph/9906036v1; published in B. Brown and J. Woods (eds.), *Logical Consequences: Rival approaches and new studies in exact philosophy*. Oxford: Hermes Science Publishers, pp. 279–292.
This is a critique of the orthodox view that quantum mechanics implies a general prohibition against controllable superluminal effects.

Penrose, Roger (1989), *The Emperor's New Mind.* Vintage Edition, Oxford University Press, London.
Among many riches this book contains a very clear, minimally technical exposition of QM.

Pitowsky, Itamar (1994), "George Boole's 'Conditions of Possible Experience' and the Quantum Puzzle," *British Journal for the Philosophy of Science* 45, 95–125.
Pitowsky shows that the Bell Inequalities are special cases of consistency conditions first defined by George Boole.

Prasanna, A. R., Narlikar, J. V., & Vishveshwara, C. V. (eds.) (1980), *Gravitation, Quanta, and the Universe.* Wiley Eastern Limited, New Delhi.
Interesting grab-bag of papers, mostly on interface between QM and general relativity.

Redhead, Michael (1987), *Incompleteness, Nonlocality and Realism: A Prolegomenon to the Philosophy of Quantum Mechanics.* Clarendon Press, Oxford.
High-level, detailed analysis of the implications of nonlocality for metaphysics.

Saunders, S., & Brown, H. (eds.) (1991), *The Philosophy of Vacuum.* Clarendon Press, Oxford.
An introduction to front-line research in quantum field theory.

Selleri, Franco (1990), *Quantum Paradoxes and Physical Reality.* Dordrecht: Kluwer.

van Fraassen, Bas (1991), *Quantum Mechanics: An Empiricist View.* Clarendon Press, Oxford.
Presents a "modal" interpretation of QM; very cogent and thorough.

Zurek, W., & Wheeler, J. A. (eds.) (1983), *Quantum Theory and Measurement.* Princeton University Press.
Essential collection of basic papers, including Aspect *et al.*. Valuable connecting exposition.

Historical

Jammer, Max (1966), *Conceptual Development of Quantum Mechanics.* McGraw-Hill, New York.
Very detailed and complete; strongly influenced by Copenhagen view.

Klein, Martin J. (1970), *Paul Ehrenfest.* American Elsevier, New York.
There are many papers on history of QM by Martin Klein (many published in *Historical Studies in the Physical Sciences.*) All are clear and authoritative; Klein sets the standard in this field.

Kuhn, Thomas (1978), *Black-Body Theory and the Quantum Discontinuity, 1894–1912.* Oxford University Press, New York.
Did Planck really believe in quantization of energy, or was he just cleverly aping a calculational device of Boltzmann's?

Schwinger, J. (1958), *Quantum Electrodynamics.* Dover, New York.
Collection of classic papers in quantum electrodynamics and field theory, up to the early 1950s. Not for the fainthearted.

ter Haar, D. (ed.) *The Old Quantum Theory.* Pergamon Press, Oxford.
Classic papers of pre-1920 era, with excellent introductory exposition by the editor.

van der Waerden, B. L., (1968), *Sources of Quantum Mechanics*, Dover Books, New York.
Reprints of several classical papers in Old Quantum Theory and matrix mechanics; little on wave mechanics. Excellent historical introduction.

Pop

Feynman, R. P. (1985), *QED: The Strange Theory of Light and Matter.* Princeton University Press, Princeton.
Masterful, barely technical exposition.

Herbert, Nick (1985), *Quantum Reality: Beyond the New Physics.* Doubleday, New York.
One of the best non-technical books I have seen in the sense that it comes probably as close as one can without mathematics to giving a sense of what QM is really like.

Hoffman, Banesh (1959), *The Strange Story of the Quantum.* Dover, New York.
Out of date, but a beautiful piece of exposition.

Zukav, Gary (1979), *The Dancing Wu Li Masters.* Bantam, New York.
Covers various interpretations of QM with breathless enthusiasm; concludes that all are true.