Introduction to Quantum Information Processing

Jonathan A. Jones

Chapter 1

The Dirac Notation

The purpose of this chapter is to provide an introduction to Dirac's notation [Dirac, Shankar 1994] for describing quantum information processing. Many areas of quantum mechanics studied in undergraduate degrees can be described without using Dirac notation, and its importance is unclear. In other areas, however, the advantages of Dirac notation are huge, and it is essentially the only notation in use. This is particularly true of quantum information theory [Nielsen 2000, Stolze 2004].

1.1 Hilbert Space

Dirac's notation is closely related to that used to describe abstract vector spaces known as Hilbert spaces, and many formal arguments about the properties of quantum systems are in fact arguments about the properties of Hilbert spaces. Here we aim to steer a careful course between the twin perils of excessive mathematical sophistication and of taking too much on trust. The description given here is closely based on that of [Goldman 1988]; for an alternative view try [Gasiorowicz 2003]. For a detailed introduction to vector spaces see [Halmos 1974].

A Hilbert space is an abstract vector space. As such, it has many properties in common with the ordinary three dimensional vectors which you studied in the first year, but it also differs in several important ways. Firstly, the vector space is not three dimensional, but can have any number of dimensions[1]. Secondly, when the vectors are multiplied by scalar numbers these numbers can be complex. Thirdly, when two vectors are combined by taking their *scalar product* (analogous to the vector dot product, and often called the *inner product*), the result depends on the order in which the vectors are taken, such that

$$\mathbf{v}.\mathbf{u} = (\mathbf{u}.\mathbf{v})^* \tag{1.1}$$

where the asterisk indicates taking the complex conjugate. Clearly the scalar product of any vector with itself is real, as

$$\mathbf{u}.\mathbf{u} = (\mathbf{u}.\mathbf{u})^* \tag{1.2}$$

and the only numbers equal to their complex conjugates are real. It can also be shown that $\mathbf{u}.\mathbf{u}$ is positive, and its positive square root (the *norm* of $\mathbf{u}$) can be thought of as the length of $\mathbf{u}$.

[1]The description below largely assumes that the number of dimensions is finite, but it is also possible to extend these results to infinite dimensional spaces.

As usual it is convenient to describe vectors by taking linear combinations of a set of basis vectors

$$\mathbf{v} = \sum_i \alpha_i \mathbf{u}_i \tag{1.3}$$

where the α_i are complex coefficients, and the $\mathbf{u}_i$ have the property that

$$\mathbf{u}_i.\mathbf{u}_j = \delta_{ij} \tag{1.4}$$

where δ_{ij}, the Kronecker delta, is equal to 1 if $i = j$, and is equal to 0 if $i \neq j$. Such a basis is said to be *orthonormal*. The coefficients α_i can be easily found, as

$$\mathbf{u}_i.\mathbf{v} = \alpha_i \tag{1.5}$$

or

$$\mathbf{v}.\mathbf{u}_i = \alpha_i^* \tag{1.6}$$

where the second version follows from equation 1.1.

1.2 Dirac notation

The essence of Dirac notation is that the state of a quantum system is fully described by a vector in an associated Hilbert space. The notation makes a clear distinction between vectors appearing on the right hand side and on the left hand side of scalar products: vectors of the first kind are called *ket* vectors, or just kets, and are written as $|\psi\rangle$, while vectors of the second kind are called *bra* vectors, or bras, and written as $\langle\psi|$. The scalar product of a bra and a ket (usually called the *inner product*) is represented by the bra(c)ket notation

$$\langle\phi|\psi\rangle \tag{1.7}$$

and equation 1.1 is written as

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*. \tag{1.8}$$

As before, bras and kets are conveniently expanded in an orthonormal basis, that is a set of kets such that

$$\langle i|j\rangle = \delta_{ij}. \tag{1.9}$$

Any ket $|\psi\rangle$ can then be written as

$$|\psi\rangle = \sum_i \alpha_i |i\rangle \tag{1.10}$$

where

$$\langle i|\psi\rangle = \alpha_i. \tag{1.11}$$

The corresponding bra can be written as

$$\langle\psi| = \sum_i \alpha_i^* \langle i| \tag{1.12}$$

with

$$\langle\psi|i\rangle = \alpha_i^*, \tag{1.13}$$

so that the set of bras $\langle i|$ forms an orthonormal basis for the bras. The inner product between $\langle\phi|$ and $|\psi\rangle$ can now be written as

$$\langle\phi|\psi\rangle = \sum_i \sum_j \beta_i^* \langle i|\alpha_j|j\rangle = \sum_{i,j} \beta_i^* \alpha_j \langle i|j\rangle = \sum_{i,j} \beta_i^* \alpha_j \delta_{ij} = \sum_i \beta_i^* \alpha_i \tag{1.14}$$

1.3 Operators

After kets and bras, the most important elements of Dirac notation are operators, which transform kets into other kets according to

$$A|\psi\rangle = |\psi'\rangle. \tag{1.15}$$

The action of an operator on a bra is analogous, but the operator must be written on the right hand side of the bra:

$$\langle\phi|A = \langle\phi'|. \tag{1.16}$$

The relationship between these two actions is defined by the fact that

$$\langle\phi|\psi'\rangle = \langle\phi'|\psi\rangle \tag{1.17}$$

and so the inner product is written as

$$\langle\phi|A|\psi\rangle \tag{1.18}$$

and it is not necessary to specify whether the operator acts on the ket or the bra. These operators are linear, so that

$$A\left(|\psi\rangle + |\phi\rangle\right) = A|\psi\rangle + A|\phi\rangle \tag{1.19}$$

and

$$\left(A + B\right)|\psi\rangle = A|\psi\rangle + B|\psi\rangle. \tag{1.20}$$

The product of two operators acting on a ket is defined by acting first with the rightmost operator, so that

$$AB|\psi\rangle = A\left(B|\psi\rangle\right). \tag{1.21}$$

As discussed above, an operator can be thought to act either on a ket or on a bra, but these operations are not quite identical. In particular the fact that $A|\psi\rangle = |\psi'\rangle$ does not in general imply that $\langle\psi|A = \langle\psi'|$. It is, however, true that

$$\langle\psi|A^\dagger = \langle\psi'| \tag{1.22}$$

where $A^\dagger$ is an operator[2] closely related to A, called the Hermitian conjugate or *adjoint* of A. The form of this operator will be considered below; for the moment it suffices to note that

$$\langle\phi|A|\psi\rangle = \langle\psi|A^\dagger|\phi\rangle^* \tag{1.23}$$

and that this can be used to show that $(A^\dagger)^\dagger = A$.

[2]Readers will probably be familiar with operators written as a and $a^\dagger$ which are used as lowering and raising operators in descriptions of the harmonic oscillator, or (equivalently) as annihilation and creation operators in treatments of light; the relation between these two (apparently distinct) uses of the dagger symbol will eventually become clear.

One important set of operators is the set of *projection operators*. Combining equations 1.10 and 1.11 gives

$$|\psi\rangle = \sum_i \langle i|\psi\rangle |i\rangle \tag{1.24}$$

and since the $\langle i|\psi\rangle$ inner products are just numbers, they can be swapped with the kets $|i\rangle$ to obtain

$$|\psi\rangle = \sum_i |i\rangle\langle i|\psi\rangle = \sum_i P_i|\psi\rangle \tag{1.25}$$

where $P_i = |i\rangle\langle i|$ is an operator which projects $|\psi\rangle$ onto the basis ket $|i\rangle$, that is obtains the component of $|\psi\rangle$ which is parallel to $|i\rangle$. In the same way we can write

$$\langle\psi| = \sum_i \langle\psi|i\rangle\langle i| = \sum_i \langle\psi|P_i. \tag{1.26}$$

As the two equations above are valid for any ket or bra, it follows that

$$\sum_i |i\rangle\langle i| = \sum_i P_i = \mathbb{1} \tag{1.27}$$

where $\mathbb{1}$ is the *identity* operator, which leaves all bras, kets, and operators unchanged, so that

$$\mathbb{1}|\psi\rangle = |\psi\rangle, \qquad \langle\psi|\mathbb{1} = \langle\psi|, \qquad A\mathbb{1} = \mathbb{1}A = A. \tag{1.28}$$

This result is called the *closure theorem*.

Operators can be grouped into various classes according to their properties, and two particularly important groups are *Hermitian* and *unitary* operators. Hermitian operators are simply those which are equal to their adjoint

$$H = H^\dagger \tag{1.29}$$

while unitary operators have their inverse equal to their adjoint, so that

$$UU^\dagger = U^\dagger U = \mathbb{1}. \tag{1.30}$$

Most physical processes are described by Hermitian or unitary operators, and as we shall see below there is a close link between them.

1.4 Vectors and matrices

As shown in equation 1.10, any ket can be thought of as a linear combination of a set of orthonormal basis vectors. Provided there is some agreed basis, it clearly suffices just to list the coefficients: thus for a ket in a three dimensional Hilbert space we can write

$$|\psi\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \tag{1.31}$$

where the coefficients form a column vector. A bra can be written in a similar way

$$\langle\psi| = \begin{pmatrix}\alpha_1^* & \alpha_2^* & \alpha_3^*\end{pmatrix} \tag{1.32}$$

where the coefficients now form a row vector. Reconsidering equation 1.14 shows that when bras and kets are written in this form the inner product is nothing more than a conventional matrix product.

It is also possible to describe operators using a matrix. Clearly

$$A|\psi\rangle = \mathbb{1}A\mathbb{1}|\psi\rangle \tag{1.33}$$

and applying the closure theorem gives

$$A|\psi\rangle = \sum_{i,j} |i\rangle\langle i|A|j\rangle\langle j|\psi\rangle \tag{1.34}$$

$$= \sum_{i,j} \langle i|A|j\rangle\langle j|\psi\rangle|i\rangle \tag{1.35}$$

where we have used the fact that the two inner products in equation 1.34 are just numbers and so can be moved to the front of the formula. Next, note three things. Firstly, using equation 1.11, we know that $\langle j|\psi\rangle = \alpha_j$. Secondly as $\langle i|A|j\rangle$ is just a number we can choose to write it as an element A_{ij} of a matrix A. Finally we can use equations 1.10 and 1.11 to expand $A|\psi\rangle$ in the same way as $|\psi\rangle$,

$$A|\psi\rangle = \sum_i \beta_i|i\rangle. \tag{1.36}$$

Combining all these results gives

$$\beta_i = \sum_j A_{ij}\alpha_j \tag{1.37}$$

and so the coefficients in the new state are obtained from those in the old state by multiplying them by A using conventional matrix multiplication.

Since a matrix can be used to describe an operator, it is instructive to consider how the product of two operators can be described. This can be achieved by considering a single element of the matrix description of the product

$$\langle i|BA|j\rangle = \langle i|B\mathbb{1}A|j\rangle \tag{1.38}$$

$$= \sum_k \langle i|B|k\rangle\langle k|A|j\rangle \tag{1.39}$$

or

$$(BA)_{ij} = \sum_k B_{ik}A_{kj} \tag{1.40}$$

so that the matrix describing the product of two operators is simply the product of their individual matrices.

It is also instructive to consider the relationship between the matrix descriptions of an operator A and its adjoint $A^\dagger$. Applying equation 1.23 to the basis vectors gives

$$\langle i|A^\dagger|j\rangle = \langle j|A|i\rangle^* \tag{1.41}$$

or

$$(A^\dagger)_{ij} = A^*_{ji} \tag{1.42}$$

so that in matrix terms taking the adjoint is equivalent to taking the complex conjugate of the matrix transpose. From this fact it is straightforward to deduce that $(AB)^\dagger = B^\dagger A^\dagger$.

1.5 Eigenvalues and eigenvectors

Consider an operator A and a ket $|\psi\rangle$ such that

$$A|\psi\rangle = \lambda|\psi\rangle \tag{1.43}$$

where λ is just a number. The ket $|\psi\rangle$ is then said to be an eigenket of the operator A, with eigenvalue λ. Alternatively, and equivalently, the vector representation of $|\psi\rangle$ is an eigenvector of the matrix A with eigenvalue λ.

Eigenvalues are most conveniently determined using the matrix formalism. In a Hilbert space with n dimensions, equation 1.43 is equivalent to n simultaneous equations of the form

$$\sum_j A_{ij} a_j = \lambda a_i \tag{1.44}$$

or

$$\sum_j (A_{ij} - \lambda\delta_{ij}) a_j = 0. \tag{1.45}$$

These simultaneous equations only have non-trivial solutions if the determinant of the coefficients on the left hand side is zero, so that

$$\begin{vmatrix} (A_{11}-\lambda) & A_{12} & \dots & A_{1n} \\ A_{21} & (A_{22}-\lambda) & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & (A_{nn}-\lambda) \end{vmatrix} = 0. \tag{1.46}$$

This determinant equation is in fact an n^{th} order polynomial in λ, whose n roots are the n eigenvalues of the matrix A. It should be noted that although the exact form of the determinant equation (1.46) seems to depend on the choice of basis in which A is described, the eigenvalues are fundamental properties of the operator A (equation 1.43) and will be the same in any basis.

Once the eigenvalues have been determined, the eigenvector corresponding to each eigenvalue can be found by solving the set of simultaneous equations (1.45). Unlike eigenvalues, the eigenvectors of an operator obviously *do* depend on the basis used to describe the operator. The eigenkets of the operator, however, are fundamental properties and do not depend on the choice of basis.

The method described above only gives the ratios of the coefficients describing the eigenvector, but this is quite proper as the eigenkets (equation 1.43) are only defined up to a multiplicative factor. It is customary to choose kets of unit norm, but this does not completely define the ket, which can still be multiplied by any complex number of the form $e^{i\phi}$. A more important source of uncertainty, however, may arise when an operator has *degenerate* eigenvalues, arising from repeated

roots in the eigenvalue polynomial. In this case linear combinations of eigenvectors corresponding to the same eigenvalue will also be eigenvectors with the same eigenvalue.

The process of finding eigenvalues and eigenvectors of a matrix is equivalent to *diagonalizing* the matrix: the matrix A can be written in the form

$$A = S\Lambda S^{-1} \tag{1.47}$$

where Λ is a diagonal matrix with the eigenvalues of A along the diagonal and S is formed from the eigenvectors of A.

1.6 Operator trace

The trace of an operator is a particularly important property. As before it is most simply defined by using a matrix description

$$\mathrm{tr}(A) = \sum_i \langle i|A|i\rangle = \sum_i A_{ii} \tag{1.48}$$

but its value does not depend on the basis. This is most easily seen by writing the matrix A in diagonal form and then using the fact that the trace of a product of matrices is invariant under cyclic permutations of the product. Thus

$$\mathrm{tr}(A) = \mathrm{tr}(S\Lambda S^{-1}) = \mathrm{tr}(\Lambda S^{-1}S) = \mathrm{tr}(\Lambda) \tag{1.49}$$

and so the trace of an operator is equal to the sum of its eigenvalues.

1.7 Hermitian operators

As mentioned above, an operator A is Hermitian if it is equal to its adjoint, $A = A^\dagger$. Hermitian operators play a key role in quantum mechanics, and have many useful properties.

Firstly, the eigenvalues of a Hermitian operator are always real. Suppose $|a\rangle$ is an eigenket of A with eigenvalue a, so that

$$A|a\rangle = a|a\rangle, \tag{1.50}$$

or, equivalently,

$$\langle a|A|a\rangle = \langle a|a|a\rangle = a\langle a|a\rangle. \tag{1.51}$$

Using equation 1.23 gives

$$\langle a|A^\dagger|a\rangle = (\langle a|a|a\rangle)^* = a^*\langle a|a\rangle \tag{1.52}$$

and since $A = A^\dagger$ we can immediately deduce that $a = a^*$. Thus a must be real.

Secondly, the eigenkets of a Hermitian operator are mutually orthogonal. Consider two eigenkets such that

$$A|a_1\rangle = a_1|a_1\rangle, \qquad A|a_2\rangle = a_2|a_2\rangle. \tag{1.53}$$

Since A is Hermitian these can be rewritten as

$$\langle a_1|A = \langle a_1|a_1, \qquad \langle a_2|A = \langle a_2|a_2, \tag{1.54}$$

and so the inner product $\langle a_2|A|a_1\rangle$ can be expanded in two different ways:

$$\langle a_2|A|a_1\rangle = a_1\langle a_2|a_1\rangle = a_2\langle a_2|a_1\rangle, \tag{1.55}$$

or

$$(a_1 - a_2)\langle a_2|a_1\rangle = 0. \tag{1.56}$$

The situation is simplest when the two eigenvalues are different, so that $a_1 - a_2 \neq 0$; in this case equation 1.56 immediately requires that $\langle a_2|a_1\rangle = 0$, so that the kets $|a_2\rangle$ and $|a_1\rangle$ are orthogonal. Things are slightly more complex in the presence of degenerate eigenvalues, but in this case it can be shown that it is always possible to take linear combinations of the corresponding eigenkets to obtain orthogonal kets.

Taken together these results imply[3] that for any Hermitian operator in an n dimensional Hilbert space, it is always possible to find n orthonormal eigenkets of the operator. Clearly these orthonormal eigenkets provide a natural basis for describing the operator.

1.8 Commutators

When two operators, A and B are applied in sequence to a ket $|\psi\rangle$ it usually matters which order they are applied in, so that

$$BA|\psi\rangle \neq AB|\psi\rangle \tag{1.57}$$

in general. More fundamentally we note that operator multiplication (like matrix multiplication) is not *commutative*, so that $BA \neq AB$. In some cases, however, the operators do have the property that $BA = AB$, and in this case the operators are said to *commute*[4]. This distinction is usually made by considering the commutator of the two operators

$$[A, B] = AB - BA \tag{1.58}$$

so that two operators commute if their commutator is zero.

Commutators play a key role in quantum mechanics, and it can be useful to consider their properties in the abstract. To give two trivial examples, it is obvious that

$$[B, A] = BA - AB = -[A, B] \tag{1.59}$$

and that

$$\begin{aligned} \mathrm{tr}([A, B]) &= \mathrm{tr}(AB - BA) && (1.60)\\ &= \mathrm{tr}(AB) - \mathrm{tr}(BA) && (1.61)\\ &= 0 && (1.62) \end{aligned}$$

where the last line has used the cyclic invariance of the trace.

[3]This is not a formal proof, as it assumes that the eigenvalues and eigenkets always exist, but a more formal proof is possible and the result is correct.

[4]Note that for two operators to commute it must be true that $BA|\psi\rangle = AB|\psi\rangle$ for *every* ket $|\psi\rangle$, so that we can write $BA = AB$; it is not sufficient if the equality only holds for some particular kets.

1.9 Unitary operators

A unitary operator U was previously defined as an operator whose inverse is equal to its adjoint, but a more fundamental definition is that a unitary operator does not change the norm of a ket. We can now show how these two definitions are related. Consider some arbitrary ket $|\psi\rangle$, such that

$$U|\psi\rangle = |\psi'\rangle \quad \text{and} \quad \langle\psi|U^\dagger = \langle\psi'|. \tag{1.63}$$

It is clear that

$$\begin{aligned} \langle\psi'|\psi'\rangle &= \langle\psi|U^\dagger U|\psi\rangle & (1.64)\\ &= \langle\psi|U^{-1}U|\psi\rangle & (1.65)\\ &= \langle\psi|\psi\rangle & (1.66) \end{aligned}$$

as required. In a similar vein it can also be shown that unitary operators also leaves the scalar product between any two kets unchanged[5].

As with Hermitian operators, unitary operators play a central role in quantum mechanics, and have many important features. For example, we note that the product of two unitary operators U and V is itself unitary, since

$$UV(UV)^\dagger = UVV^\dagger U^\dagger = UU^\dagger = \mathbb{1}. \tag{1.67}$$

More interestingly, it can be shown that the eigenvalues of a unitary operator all have modulus one, and that the eigenvectors of a unitary matrix are orthogonal. Both of these properties can be deduced by considering two eigenkets of U, $|u_1\rangle$ and $|u_2\rangle$, with eigenvalues λ_1 and λ_2. Clearly

$$\begin{aligned} \langle u_2|u_1\rangle &= \langle u_2|U^\dagger U|u_1\rangle & (1.68)\\ &= \lambda_2^*\lambda_1\langle u_2|u_1\rangle & (1.69) \end{aligned}$$

where the first line results from the fact that $U^\dagger U = \mathbb{1}$, and the second line comes from the fundamental properties of operators. Thus

$$(\lambda_2^*\lambda_1 - 1)\langle u_2|u_1\rangle = 0. \tag{1.70}$$

Choosing $|u_2\rangle = |u_1\rangle$ leads immediately to $\lambda_1^*\lambda_1 = 1$, showing that the eigenvalues have modulus one as required. The proof that the eigenvectors are orthogonal is virtually identical to that used for Hermitian operators above.

Finally we consider an important link between unitary and Hermitian operators. Since the eigenvalues of a unitary operator have modulus one, they can all be written in the form

$$\lambda_j = \exp(-ia_j) \tag{1.71}$$

where the numbers a_j are real. These numbers can be though of as the eigenvalues of another operator A which has the same eigenkets as U. Since the eigenvalues of A are real, A must itself be Hermitian. In general we can write

$$U = \exp(-iA) \tag{1.72}$$

[5]This property suggests that unitary operators can be considered as changing between two different bases for describing a system, and this is indeed the case.

connecting any unitary operator with its associated Hermitian operator. The meaning of the exponential of an operator is most simply described by considering the exponential of a matrix, and for a diagonal matrix this process is simple:

$$\exp\left[\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\right] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} + \frac{1}{2!}\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} + \ldots \tag{1.73}$$

$$= \begin{pmatrix} 1 + a + a^2/2 + \ldots & 0 \\ 0 & 1 + b + b^2/2 + \ldots \end{pmatrix} \tag{1.74}$$

$$= \begin{pmatrix} \exp[a] & 0 \\ 0 & \exp[b] \end{pmatrix}. \tag{1.75}$$

The exponential of a general matrix can be calculated in a similar way by first diagonalizing the matrix and then noting that

$$\exp[S\Lambda S^{-1}] = S\exp[\Lambda]S^{-1}. \tag{1.76}$$

This result is easily proved by using a series expansion of the exponential function, as shown above, and canceling matching pairs of S^{-1} and S matrices. More fundamentally, S and S^{-1} are the matrices which transform between the basis we happen to be working in, and the *eigenbasis* of the operator, in which its description is naturally diagonal.

1.10 Physical systems

At last we can proceed to see how Dirac notation can be used to describe a physical system. The most important property of the system is its Hamiltonian operator $\mathcal{H}$ which described the energy of the system. According to the time independent Schrödinger equation the Hamiltonian has an associated set of eigenstates

$$\mathcal{H}|j\rangle = \hbar\omega_j|j\rangle \tag{1.77}$$

which form an orthonormal basis for the system. As the eigenvalues of the Hamiltonian (given by $\hbar\omega_j$) correspond to the energies of the eigenstates they must be real, and so $\mathcal{H}$ must be Hermitian. The most general state of the system is then some *superposition*, or linear combination, of these basis states,

$$|\psi\rangle = \sum_j \alpha_j|j\rangle. \tag{1.78}$$

The evolution of the system is given by the time dependent Schrödinger equation,

$$\frac{\partial}{\partial t}|\psi\rangle = -i\frac{\mathcal{H}}{\hbar}|\psi\rangle \tag{1.79}$$

which has the solution

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle \tag{1.80}$$

with

$$U(t) = \exp(-i(\mathcal{H}/\hbar)t). \tag{1.81}$$

The evolution of quantum states can also be described using the compact notation

$$|\psi\rangle \xrightarrow{\mathcal{H}t} U|\psi\rangle. \tag{1.82}$$

Since $\mathcal{H}$ is Hermitian, the evolution operator U, usually called the *propagator*, must be unitary.

1.11 Time-dependent Hamiltonians

The discussion above assumes that the Hamiltonian is time-independent, that is it does not vary with time. This will not be true in complicated systems, which are controlled by varying the Hamiltonian. In many cases, however, the Hamiltonian is *piecewise constant*, that is it has a constant value for some finite length of time, and is then replaced by a different constant value for another finite time period, and so on. In this case the evolution can be described using a series of propagators

$$|\psi\rangle \xrightarrow{\mathcal{H}_1 t_1} \xrightarrow{\mathcal{H}_2 t_2} \xrightarrow{\mathcal{H}_3 t_3} U_3 U_2 U_1 |\psi\rangle \tag{1.83}$$

with $U_1 = \exp[-i(\mathcal{H}_1/\hbar)t_1]$ and so on. Note that the sequence of Hamiltonians is normally written with time running from left to right (that is the leftmost Hamiltonian is the first to be applied), while the sequence of propagators is always written from right to left, as the rightmost propagator is applied first. It is, of course, possible to combine the sequence of propagators into a single propagator, $U = U_3 U_2 U_1$, by matrix multiplication.

The situation is much more complicated when the Hamiltonian varies continuously with time. It is, of course, possible to write down a formal solution of the form of equation (1.83), but this is not generally a useful approach. For the moment this issue will simply be ignored.

1.12 Global phases

The discussion above has glossed over one important aspect of using kets to represent the state of physical systems. The description of a physical state as a linear combination of basis states (equation 1.78) provides *too much* information, as the kets $|\psi\rangle$ and

$$e^{i\phi}|\psi\rangle = \sum_j e^{i\phi}\alpha_j |j\rangle. \tag{1.84}$$

describe the same physical state. It is safe to use this approach as long as you remember that two kets differing only by an overall phase shift correspond to the same state. Note also that states are only invariant under overall phases (often called *global phase shifts*), and changes in the relative phases of the terms contributing to a superposition *are* important!

Chapter 2

Quantum Information

After all this underlying theory, we will finally turn to quantum information processing. The basic element used in quantum information is the quantum bit, or qubit. This is simply a physical system with two energy levels, which we shall call $|0\rangle$ and $|1\rangle$. Taking the standard approach of quantum information theory, we shall not worry too much about the properties of these states, or even what their energies are; we shall simply assume that they are eigenstates of the Hamiltonian with known eigenvalues (that is, known energies). This approach allows us to concentrate on the fundamental properties of the system, without all the tedious solving of complicated differential equations.

Classical information processing is performed using *bits*, which are just two state systems, with the two states called 0 and 1. By grouping bits together we can represent arbitrary pieces of information, and by manipulating these bits we can perform arbitrary computations. We can in principle perform classical information processing on our quantum system by using the two states $|0\rangle$ and $|1\rangle$ as our logical states 0 and 1 and proceeding in the usual fashion[1], but this misses the point. A qubit[2] is not confined to these two states, but can be found in arbitrary superposition states. Although it is not immediately obvious what a state like

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \tag{2.1}$$

actually means in information processing terms, it is clear that quantum bits are in some sense more powerful than their classical equivalents. Quantum information processing is, of course, the art of exploiting these superposition states to perform information processing tasks which are impossible for classical systems[3].

[1] There are a few technical issues arising in this approach, which is called Reversible Computation; see [Feynman 1996] for details.

[2] There are many possible physical implementations of a qubit, such as spin states of electrons or atomic nuclei, charge states of quantum dots, atomic energy levels, vibrational states of groups of atoms, polarization states of photons, or paths in an interferometer. At this stage the physical implementation is not important: the idea of a qubit is to abstract the discussion away from physical details. Note, however, that the two spin states of a spin-1/2 particle provide a particularly natural implementation of a qubit, and the language of spins is frequently used.

[3] Just as the real power of classical information processing requires groups of bits, the real advantages of quantum information processing only become clear in systems with two or more qubits; for simplicity, however, we are confining ourselves to single isolated qubits at the moment

2.1 The Bloch sphere

The enormous flexibility of a single qubit in comparison with a classical bit can be most clearly seen using the *Bloch sphere* description of a qubit. This also provides a simple but powerful way of visualizing the behavior of a qubit. We begin by looking again at the general state of a single qubit, equation (2.1), and noting that this ket must have unit norm, so that $|\alpha|^2 + |\beta|^2 = 1$. The fact that the state does not change under global phase shifts means that we can always choose α to be *real*, and the normalization constraint is easily imposed by making α and β depend on the cosine and sine of a single parameter. As discussed below, a particularly useful form is to write

$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle \tag{2.2}$$

where $0 \leq \theta \leq \pi$ and $0 \leq \phi < 2\pi$. Note that $\theta = 0$ corresponds to $|\psi\rangle = |0\rangle$, and $\theta = \pi$ corresponds to $|\psi\rangle = |1\rangle$; in these extreme cases the value of ϕ is irrelevant.

There is an obvious analogy between the variables θ and ϕ used above and those used in spherical polar coordinates. Clearly any ket $|\psi\rangle$ can be associated with a single point on the surface of a sphere of radius 1 with co-latitude and azimuth angles θ and ϕ; this sphere is usually called the Bloch sphere. Alternatively (and entirely equivalently) a state can be represented as a unit vector (connecting the origin and a point on the Bloch sphere), called a Bloch vector.

The two basis states $|0\rangle$ and $|1\rangle$, which correspond to the states 0 and 1 of a classical bit, lie at the north and south poles of the Bloch sphere, while a qubit can lie anywhere at all on the surface. One interesting group of states is the set of equally weighted superpositions, with $|\alpha| = |\beta| = 1/\sqrt{2}$, which lie on the equator of the Bloch sphere, with the exact position determined by the relative phase of α and β.

2.2 Density matrices

As described previously, it is frequently convenient to describe the state of a qubit using a vector, written using the basis states $|0\rangle$ and $|1\rangle$ (the computational basis). Thus equation (2.1) can be written as

$$|\psi\rangle = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}, \tag{2.3}$$

while the corresponding bra can be written as

$$\langle\psi| = \begin{pmatrix}\alpha^* & \beta^*\end{pmatrix}. \tag{2.4}$$

The basis states, of course, take the simple forms[4]

$$|0\rangle = \begin{pmatrix}1\\ 0\end{pmatrix} \quad \text{and} \quad |1\rangle = \begin{pmatrix}0\\ 1\end{pmatrix}. \tag{2.5}$$

Bras and kets are normally combined by taking the inner product, such as

$$\langle\psi|\psi\rangle = \begin{pmatrix}\alpha^* & \beta^*\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \alpha^*\alpha + \beta^*\beta = 1 \tag{2.6}$$

[4]There is a potential ambiguity in any description of quantum bits, as to whether $|0\rangle$ and $|1\rangle$ are defined as shown here, or the other way round. Fundamentally, of course, the choice does not matter, as long as one is consistent. Here I follow the most common notation, but both approaches are in use.

but they can also be combined using the *outer product*

$$|\psi\rangle\langle\psi| = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}\begin{pmatrix}\alpha^* & \beta^*\end{pmatrix} = \begin{pmatrix}\alpha\alpha^* & \alpha\beta^*\\ \beta\alpha^* & \beta\beta^*\end{pmatrix}. \tag{2.7}$$

This outer product is called a *density matrix* description of the state.

It is obvious from the form of equation 2.7 that the density matrix describing a qubit is Hermitian, and has trace one; these are in fact general properties which apply to all density matrices. A two by two matrix can always be expanded as a weighted sum of four basic matrices (a matrix basis), and the most useful basis is provided by the Pauli matrices

$$\sigma_0 = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \quad \sigma_x = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} \quad \sigma_y = \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} \quad \sigma_z = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \tag{2.8}$$

where the usual set of three Pauli matrices has been extended to include the *identity matrix* σ_0. As the Pauli matrices are Hermitian, a density matrix can be written as

$$|\psi\rangle\langle\psi| = \tfrac{1}{2}\left(\sigma_0 + s_x\sigma_x + s_y\sigma_y + s_z\sigma_z\right) \tag{2.9}$$

where s_x, s_y and s_z are three *real* coefficients. This might seem excessive, as we know that any pure state can be described using only two numbers (θ and ϕ), but it is easily shown that s_x, s_y and s_z are related; in effect they are the three components of a vector of unit length. Indeed it can be shown that this vector is identical to the Bloch vector, discussed above.

Qubits can also be found in mixed states, which are just weighted sums of pure states of the form

$$\rho = \sum_n P_n|\psi_n\rangle\langle\psi_n| \tag{2.10}$$

where $P_n \geq 0$ is the contribution of the pure state $|\psi_n\rangle\langle\psi_n|$ to the mixture (the probability of the pure state occurring in the mixture). Clearly such mixed states are Hermitian, and as the probabilities of the various contributions must sum to one ($\sum_n P_n = 1$) the density matrix must have trace one. It can be shown that any mixed state of a single qubit corresponds to a point *inside* the Bloch Sphere.

It is useful to be able to calculate the evolution of states described using a density matrix rather than a ket vector. This problem can be addressed directly by solving the Liouville–von Neumann equation (the density matrix equivalent of the time dependent Schrödinger equation), but it is simpler to proceed by analogy. The evolution of a bra vector is clearly closely related to the evolution of the corresponding ket vector, and a little thought shows that

$$(U|\psi\rangle)^\dagger = \langle\psi|U^\dagger \tag{2.11}$$

so that the density matrix description of a pure state evolves according to

$$|\psi\rangle\langle\psi| \xrightarrow{\mathcal{H}t} U|\psi\rangle\langle\psi|U^\dagger \tag{2.12}$$

and the linearity of the operations guarantees that a mixed state will evolve in the same way.

2.3 Propagators and Pauli matrices

We have already noted that the Pauli matrices are Hermitian, and thus provide a natural basis for describing the density matrix corresponding to a qubit. In the same way, the fact that any Hamiltonian is Hermitian means that any Hamiltonian applied to a single qubit can be written as a weighted sum of the four Pauli matrices, equation (2.8), where the weights are *real*. This means that the Pauli matrices provide a natural language for describing single qubit interactions as well as single qubit states.

The fact that any propagator describing the evolution of a quantum system is unitary has several significant consequences. Firstly it means that every propagator has an inverse, and so quantum evolution is *reversible*. (One exception to this general principle is *measurement*, which is discussed in more detail below). Secondly unitary transformations are *length preserving* and can in general be thought of as *rotations* of the vector describing the quantum state. Thirdly we note that the Pauli matrices are unitary, and so correspond to possible propagators. As we shall see later the Pauli matrices viewed as propagators correspond to important quantum logic gates[5].

The fact that the Pauli matrices are *both* unitary and Hermitian has the interesting consequence that

$$\sigma_\alpha^2 = \sigma_0 \tag{2.13}$$

where σ_α are the usual Pauli matrices, with α equal to x, y, or z. This observation can be used to show that

$$\exp(-i\theta\,\sigma_\alpha) = \cos(\theta)\sigma_0 - i\sin(\theta)\sigma_\alpha \tag{2.14}$$

without diagonalizing any matrices, making it easy to calculate many single qubit propagators.

Finally we note that the propagator corresponding to a Hamiltonian which is some multiple of σ_0 is simply a global phase shift, which has no physical significance. In essence this occurs because adding multiples of σ_0 corresponds to moving the zero-point of the energy scale, which has no physical significance.

2.4 Quantum logic gates

The basic idea of quantum information processing is that information is stored in quantum bits and processed by quantum logic gates. Just as classical logic gates take classical bits from one state to another, so quantum logic gates take qubits from one state to another. This can be achieved by modifying the system's Hamiltonian, by applying additional *control fields* to the background Hamiltonian which underlies the system.

Applying Hamiltonians will cause qubits to evolve under unitary transformations, which are reversible. With classical bits there are only two reversible gates which act on a single bit: the NOT gate, which takes a bit in state 0 into state 1 and *vice versa*, and the IDENTITY gate, which just leaves the bit unchanged[6]. There are also two irreversible gates, SET which sets a bit to 1 whatever its initial state, and CLEAR which sets a bit to 0. Clearly these two cannot be achieved with unitary transformations, and so we will neglect them for the moment.

[5]Suspicious minds might surmise that using the Pauli matrices to describe quantum states, Hamiltonians, propagators, and logic gates will inevitably lead to confusion, but in practice such problems rarely occur.

[6]It may seem excessive to consider trivial gates such as IDENTITY, but the formalism works better if they are included.

Returning to the two unitary gates, we must first find propagators that implement them. Clearly σ_0 will perform IDENTITY as

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.15}$$

while σ_x corresponds to NOT as

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{2.16}$$

We now have to find Hamiltonians which can give rise to these propagators. Clearly σ_0 can be achieved simply by doing nothing at all[7], but σ_x is slightly more difficult. For the moment it suffices to note that

$$\exp(-i\pi\sigma_x/2) = -i\sigma_x \tag{2.17}$$

(the reason for dividing the σ_x by 2 will soon become clear), and so a NOT gate can be achieved by evolving the qubit under a Hamiltonian proportional to σ_x for an appropriate time[8]. Note that the factor of $-i$ is just a global phase, and so should be ignored.

The quantum NOT gate behaves exactly like a classical NOT gate when applied to basis states, but it can also be applied to more general states[9]:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \beta \\ \alpha \end{pmatrix}. \tag{2.18}$$

The effect of this gate can be better understood by considering its effect on the Bloch sphere. Rewriting the general state in polar coordinates as before,

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \cos(\theta/2) \\ e^{i\phi}\sin(\theta/2) \end{pmatrix} = \begin{pmatrix} e^{i\phi}\sin(\theta/2) \\ \cos(\theta/2) \end{pmatrix} \tag{2.19}$$

$$= e^{i\phi} \begin{pmatrix} \sin(\theta/2) \\ e^{-i\phi}\cos(\theta/2) \end{pmatrix} \tag{2.20}$$

$$= e^{i\phi} \begin{pmatrix} \cos([\pi-\theta]/2) \\ e^{-i\phi}\sin([\pi-\theta]/2) \end{pmatrix} \tag{2.21}$$

shows that (neglecting the irrelevant global phase) the effect of a NOT gate is to negate both the latitude and longitude coordinates. A little thought shows that this is equivalent to rotating the Bloch sphere by 180° around the x axis. The significance of equation (2.17) should now be clear: the effect of applying some Hamiltonian to a qubit is to rotate the Bloch sphere around an axis

[7] In fact the IDENTITY gate is slightly more subtle than it might seem, as the state of the qubit will evolve under the background Hamiltonian even when no additional control fields are applied. This point will be addressed later.

[8] Once again the situation is subtler than it might seem: the obvious approach is just to apply a control field which generates a Hamiltonian proportional to σ_x, but this is not quite right as the background Hamiltonian will also still be present. The brute force solution is just to make the control field very large in comparison with the background Hamiltonian, but this is rarely practical. A better approach is to apply a weak control field which oscillates at a resonance frequency of the system. This point will be explored in subsequent lectures.

[9] This ability to perform information processing on superposition states lies at the root of the power of quantum computers.

parallel to the Hamiltonian. The angle of rotation depends on both the intrinsic strength of the Hamiltonian, and the time for which it is applied.

Thinking of the NOT gate as a 180° rotation also make sense when considering the effect of applying two NOT gates in sequence. Clearly this should have no overall effect, and it is comforting to note that two successive 180° rotations is equivalent to a 360° rotation, which leaves the Bloch sphere unchanged. Reversing this approach we can also think about rotations through smaller angles, such as a 90° rotation around the x axis. This has the propagator

$$\exp[-i\,\pi/2\,\sigma_x/2] = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \tag{2.22}$$

which acts to convert basis states into superpositions. Applying this propagator twice gives a NOT gate, and so it is called the SQUARE-ROOT-OF-NOT gate. Clearly this gate has no classical equivalent: it is a purely quantum logic gate.

This is not the only purely quantum logic gate: there are an infinite number of such gates! In general any rotation of the Bloch sphere (that is, a rotation by any angle around any axis) can be considered as a quantum logic gate, and can be implemented by applying an appropriate Hamiltonian for an appropriate time. For the moment we will briefly consider two of the more important gates: the Hadamard gate and the phase gate.

The Hadamard gate, usually indicated by the letter H, is similar to the SQUARE-ROOT-OF-NOT gate, but with subtly different effects. It is described by the propagator

$$\mathrm{H} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{2.23}$$

and so acts on the basis states to give

$$|0\rangle \longrightarrow \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) = |+\rangle \quad \text{and} \quad |1\rangle \longrightarrow \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right) = |-\rangle. \tag{2.24}$$

Unlike the SQUARE-ROOT-OF-NOT gate the Hadamard gate is self-inverse, so that applying it twice is equivalent to doing nothing. This means that the Hadamard gate must correspond to a 180° rotation, and it is in fact equivalent to a 180° rotation around an axis tilted at 45° degrees from the x axis towards the z axis.

The phase gate is usually indicated by the letter S, and can be thought of as the SQUARE-ROOT-OF-σ_z gate. It is described by the propagator

$$\mathrm{S} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \tag{2.25}$$

and its effect is simply to change the phase of $|1\rangle$ by 90°, while leaving $|0\rangle$ unaffected, or, equivalently, to rotate the Bloch sphere by 90° around the z axis. Note that the classical states 0 and 1, which lie at the north and south poles of the Bloch sphere, are not affected by this rotation, but the phase of a superposition *will* be changed.

2.5 Gate notation

It is possible to describe quantum gates in many different ways, and this has given rise to a range of notations for discussing them. For example the NOT gate can also be written as X, as σ_x, or

as 180_x. The decision between these forms is usually a matter of context and the background of the person discussing the gate! Researchers with a background in computer science would tend to use the most abstract notation, X, while physicists studying quantum information theory would normally choose the Pauli matrix form, σ_x. By contrast, experimental physicists whoa re interested in actually building quantum computers would usually use the description 180_x, as this corresponds most closely to a physical process. It is usually a good idea to keep an open mind, and be ready to use whatever notation is around. A list of important gates can be found in Appendix A.

There is, however, one important distinction between the theoretician's X and σ_x, on the one hand, and 180_x on the other, and this is the matter of global phases. It is clear from equation 2.17 that 180_x is *not* exactly the same as σ_x, but differs by a global factor of $-i$. In the single qubit case this global phase is completely irrelevant, but in systems of two or more qubits it can be necessary to be a little more careful.

2.6 Quantum networks

Just as a single bit is not much use on its own, very little can be achieved with a single logic gate. Effective information processing requires that gates be joined together to form *networks*, and the same approach cane be used with quantum logic gates. Clearly quantum networks will only be really useful when applied to systems with more than one qubit[10], but even with a single isolated qubit the idea has some use. Gate networks can be used both to explain some classic experiments, such as Ramsey fringes and spin echoes, and also to build single qubit gates out of other gates.

As an example of the second kind, consider the network HSSH, which corresponds to applying first a Hadamard gate, then a phase gate, then another phase gate, and finally a Hadamard gate. The effect of this can be deduced by applying the gates in sequence to a qubit in a general state, but it is more useful to consider the network directly, by simply multiplying out the constituent propagators

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & i\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & i\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} \tag{2.26}$$

to find that this network is equivalent to a NOT gate. Since $\mathrm{SS} = \sigma_z$ this network can also be written as $\mathrm{H}\sigma_z\mathrm{H} = \sigma_x$, or even more simply as HZH = X. Some standard networks are listed in Appendix A.

The network notation does give rise to one serious ambiguity of notation which we have sidestepped above. When describing a process by a sequence of operators, the operators are applied from right to left, so that the first operator applied is the rightmost operator written in the sequence. By contrast, networks are usually written running from left to right, so that the first operator applied is the leftmost operator written in the network. In some cases, therefore, it can be unclear whether to apply the gates from right to left or left to right! In the networks above, of course, this distinction is irrelevant as the networks are symmetric, but in general this ambiguity can be a problem.

Spin echoes occur when a 180° rotation is placed half way through a period of evolution under a background Hamiltonian of the form $\omega\,\sigma_z/2$, and rely on the identity $\phi_z\,180_x\,\phi_z \equiv 180_x$. By

[10]It can be shown that, just as a classical logic network can be built using only one and two bit gates (AND, OR and NOT), any quantum logic network can be built out of one qubit and two qubit quantum logic gates.

this means evolution under the background Hamiltonian can be canceled, making the final state independent of the value of ω. Spin echoes are best known in the context of Nuclear Magnetic Resonance (NMR), but are a universal quantum phenomenon.

An important example of building quantum logic gates out of networks is provided by the Hadamard gate. There are many different ways of implementing this, but the most useful approach is to relate the Hadamard to a 90° rotation. We have already considered a 90°_x rotation, and a 90°_y rotation can be described in much the same way:

$$90^\circ_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \qquad 90^\circ_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}. \tag{2.27}$$

From the form of the 90°_y operator it is obvious that it is closely related to a Hadamard, and calculations show that the Hadamard is equivalent to a 180°_z operation followed by a 90°_y operation, or to a 90°_{-y} followed by a 180°_z.

2.7 Initialization and measurement

So far we have only considered unitary gates (gates that can be described by unitary matrices), but some important gates are obviously not unitary. For example consider the CLEAR gate, which sets a qubit to the state $|0\rangle$ whatever its initial state is; clearly this process cannot be described by matrix multiplication. This might seem problematic, as evolution of a quantum system under a Hamiltonian is *always* unitary, and it is not clear how a quantum system can evolve other than in response to a Hamiltonian.

The solution to this quandary is that while a single isolated qubit can only undergo unitary evolution, there isn't really any such thing as an isolated qubit. The fact that we can use control fields to alter the state of the qubit means that the qubit *must* have some interaction with the rest of the world. It can be shown that non-unitary evolutions of a qubit can be achieved by performing a unitary evolution on a composite system, comprising the qubit and some environment, and then ignoring the state of the environment. A detailed analysis of this process clearly requires an understanding of two qubit systems, and so is beyond the scope of this chapter; as usual it suffices to note that non-unitary operations can be performed.

Another important non-unitary gate is the READOUT gate, which simply performs a classical measurement of the state of a single qubit. A full discussion of what measuring the state of a quantum system really means would very complicated, and we don't yet have a complete understanding, but fortunately it is easy to give an accurate mathematical description of what the measurement process *does* to the quantum state. As usual we start by considering a single qubit in a general state

$$|\psi\rangle\langle\psi| = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} = \begin{pmatrix} \alpha\alpha^* & \alpha\beta^* \\ \beta\alpha^* & \beta\beta^* \end{pmatrix} \tag{2.28}$$

and then consider the state of the qubit *after* the measurement. Assuming we measure in the computational basis[11] we know that the result of the measurement will be either that the qubit

[11] We can of course choose to measure the qubit using some other basis, but this would simply make the process appear more complicated without changing any fundamentals. Furthermore a measurement of a single qubit in any basis can always be achieved by using a measurement in the computational basis preceded and followed by appropriate unitary transformations.

is in state $|0\rangle$, or that it is in state $|1\rangle$, and that after the measurement the qubit will be found in the appropriate state. We also know that the probability of getting the result $|0\rangle$ is given by $|\alpha|^2 = \alpha\alpha^*$, and the probability of getting the result $|1\rangle$ is given by $|\beta|^2 = \beta\beta^*$.

We could choose to stop the discussion here, but it would be useful to be able to describe the state of the system *after* the measurement in the language we have used before. We don't know what the state of the system is after the measurement, because we don't know what the result of the measurement! We can, however, make probabilistic statements about it, and this is the way to proceed. Clearly the final state is a mixed state, with the form of equation (2.10), and can be written as

$$\rho = \alpha\alpha^*|0\rangle\langle 0| + \beta\beta^*|1\rangle\langle 1| \tag{2.29}$$

$$= \alpha\alpha^* \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \beta\beta^* \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.30}$$

$$= \begin{pmatrix} \alpha\alpha^* & 0 \\ 0 & \beta\beta^* \end{pmatrix} \tag{2.31}$$

showing that from a mathematical point of view the effect of a measurement is simply to zero the off-diagonal elements of the density matrix.

The dephasing process is sometimes called *decoherence* as the elements of the density matrix which are lost are those which correspond to the system being in a *coherent superposition* state. Decoherence is almost always the enemy of quantum information processing, and vast effort is put into controlling it. The topic is, however, terribly complicated and here we simply note that the random interactions between a quantum system and its environment have the same form as measurements. Thus decoherence results from the environment measuring the state of the system, and so the system must be well insulated from the surroundings if it is to exhibit interesting quantum behavior.

Chapter 3

Atom in a Laser Field

In this chapter we will use a succession of different methods[1] to calculate the interaction between an atom and the light field from a laser. We will see that the effect of the light is to cause *transitions* between different energy levels in the atom, but that these transitions only occur if the frequency of the light is tuned to match the energy gap between the levels

$$h\nu = \hbar\omega = E_f - E_i \tag{3.1}$$

so that the light is *resonant* with the transitions.

Atoms have an infinite number of energy levels, and might seem to be rather complex systems, but the resonance condition means that our treatment of them can be greatly simplified. In most cases it will be sufficient to consider a *two level atom*, which is assumed to have a ground state $|g\rangle$ and a single excited state $|e\rangle$, and a laser field which is close to resonance with this transition. Other transitions are far from resonance and so can be ignored.

3.1 Time-dependent systems

Consider a quantum mechanical system with a Hamiltonian $\mathcal{H}^0$, which is subjected to a time-varying perturbation[2] $\delta(t)$. The total Hamiltonian of the system is then

$$\mathcal{H} = \mathcal{H}^0 + \mathcal{H}^1(t). \tag{3.2}$$

As usual the eigenstates of $\mathcal{H}^0$ form a complete set, and so we can write the wavefunction of the system in this basis,

$$|\psi(t)\rangle = \sum_j c_j(t)|j\rangle \tag{3.3}$$

with the time dependence of $|\psi(t)\rangle$ arising from the time dependence of the coefficients. If there was no perturbation present then these coefficients would still oscillate at their natural frequencies,

$$c_j(t) = c_j(0)e^{-iE_jt/\hbar}, \tag{3.4}$$

[1] All the methods used in this chapter are *semi-classical* treatments, in which we treat the light field as a classical system; a brief introduction to fully quantum approaches can be found in Appendix B.

[2] The treatment here is closely based on that given by Prof. Steane in his A3 lectures notes "Further Quantum Physics," which is itself largely taken from [Shankar 1994].

and so it is useful to separate the time variation into that which would occur without the perturbation, and any additional variation which can be ascribed to the perturbation. Thus we write

$$|\psi(t)\rangle = \sum_j d_j(t) e^{-iE_j t/\hbar} |j\rangle \tag{3.5}$$

with all the interesting behavior now found in the values of $d_j(t)$. Now we know from the time-dependent Schrödinger equation that $[i\hbar\,\partial/\partial t - \mathcal{H}^0 - \mathcal{H}^1(t)] = 0$, and applying this operator to both sides of equation 3.5 gives

$$0 = \sum_j \left(i\hbar\, \dot{d}_j(t) - d_j\, \mathcal{H}^1(t) \right) e^{-iE_j t/\hbar} |j\rangle \tag{3.6}$$

or

$$\sum_j i\hbar\, \dot{d}_j(t) e^{-iE_j t/\hbar} |j\rangle = \sum_j d_j\, e^{-iE_j t/\hbar} \mathcal{H}^1(t) |j\rangle. \tag{3.7}$$

We can pick out the time-dependence of one of the coefficients, say d_k, by taking the inner product of $\langle k|$ with equation 3.7 giving

$$i\hbar\, \dot{d}_k e^{-iE_k t/\hbar} = \sum_j d_j e^{-iE_j t/\hbar} \langle k|\mathcal{H}^1(t)|j\rangle \tag{3.8}$$

which can be written as

$$\dot{d}_k = -i \sum_j d_j e^{i\omega_{kj} t} \mathcal{H}^1_{kj}(t)/\hbar \tag{3.9}$$

where $\omega_{kj} = (E_k - E_j)/\hbar$ and $\mathcal{H}^1_{kj}(t) = \langle k|\mathcal{H}^1(t)|j\rangle$ are called the *matrix elements* of $\mathcal{H}^1$. Note that this equation is *exact*, and is really just the time-dependent Schrödinger equation in disguise.

3.2 Sudden jumps

As a first attempt at solving this equation, consider a really simple (indeed stupidly simple) model system[3], namely a two level atom with a single electron which experiences an electric field $\boldsymbol{E}$ for a time τ. The perturbation Hamiltonian is then

$$\mathcal{H}^1 = -\boldsymbol{\mu} \cdot \boldsymbol{E} = \begin{cases} ezE & 0 \le t \le \tau \\ 0 & \text{otherwise} \end{cases} \tag{3.10}$$

where $\boldsymbol{\mu} = -e\mathbf{r}$ is the dipole moment of the atom arising from the separation of the electron and the nucleus, and the electric field direction has been taken as defining the z-axis. From symmetry grounds it is obvious that

$$\langle g|\mathcal{H}^1|g\rangle = \langle e|\mathcal{H}^1|e\rangle = 0 \tag{3.11}$$

[3]As we shall see below, the two level atom model is wildly inappropriate in this case; however some of the ideas we come across here will carry over into more realistic systems. The treatment here largely follows that in Chapter 6 of [Atkins 1997].

and we can choose to write

$$\langle g|\mathcal{H}^1|e\rangle = \hbar V \qquad \langle e|\mathcal{H}^1|g\rangle = \hbar V^* \tag{3.12}$$

where the second result is deduced from the first by the fact that the Hamiltonian is Hermitian, although in this case $V^* = V$. Thus the time dependence of the coefficients is given by

$$\dot{d}_e = -i\, d_g\, e^{i\omega_0 t} V \tag{3.13}$$

$$\dot{d}_g = -i\, d_e\, e^{-i\omega_0 t} V \tag{3.14}$$

where $\omega_0 = \omega_{eg} = -\omega_{ge}$ corresponds to the energy gap between the ground and excited states. These coupled differential equations can be solved by differentiating one equation with respect to time and substituting the other equation into the result, to give a single second order ordinary differential equation. The procedure is fairly straightforward but messy[4]. It is useful to start by considering the simplest case where the field is very strong, or the two energy levels are almost degenerate, so that $V \gg \omega_0$ and the exponential terms can simply be ignored. The equations are now easy to solve; assuming the atom starts in the ground state (so that $d_g = 1$ and $d_e = 0$) the result is

$$d_g = \cos(Vt), \qquad d_e = -i\sin(Vt). \tag{3.15}$$

The effect of the sudden strong perturbation is to cause the system to make transitions from the ground state to the excited state and back again: the amplitude of the excited state is modulated sinusoidally at a rate given by V. The exact result has the same broad form: assuming that the atom starts in the ground state then

$$d_e = -i\sqrt{\frac{4V^2}{4V^2+\omega_0^2}}\sin\left(\frac{t\sqrt{4V^2+\omega_0^2}}{2}\right)e^{i\omega_0 t/2} \tag{3.16}$$

which reduces to equation 3.15 when $\omega_0 \to 0$.

This sinusoidal modulation is called *Rabi flopping* and is also found in more realistic treatments of transitions. Note that flopping will only occur at all if the perturbation *connects* the two transitions, that is

$$V = \langle e|\mathcal{H}^1|g\rangle/\hbar \neq 0, \tag{3.17}$$

and is only efficient if $V > \omega_0$, where $\hbar\omega_0$ corresponds to the gap between the energy levels. Thus a static field can be very effective at inducing transitions between degenerate energy levels, but will have little effect on non-degenerate levels unless it is very strong. In this latter case the field will cause transitions between many different pairs of levels[5], and the two level atom assumption will not be valid. Fortunately there are more subtle ways of inducing transitions.

3.3 Oscillating fields

A much more practical approach is to note that transitions can be induced by a small oscillating field, as long as the field is close to resonance with the desired transition. In many texts this

[4]The gory details are in [Atkins 1997].

[5]Indeed a sufficiently strong field will cause transitions to unbound states, effectively tearing the atom apart!

result is derived using time-dependent perturbation theory, but it is more insightful to begin with an analytic result. Consider a cosinusoidal oscillating electric field, with an angular frequency $\omega = 2\pi\nu$ and intensity $\mathcal{E}$; this can be rewritten as the sum of two complex fields

$$\mathcal{E}(t) = \mathcal{E}\cos\omega t = \tfrac{1}{2}\mathcal{E}\left(e^{i\omega t} + e^{-i\omega t}\right) \tag{3.18}$$

and for the moment we will only consider the first term in this sum and will ignore the *counter-rotating* component; justifications of this approach, which is called the *rotating wave approximation* will be given below. The matrix elements of the perturbation Hamiltonian are now given by

$$\langle g|\mathcal{H}^1|e\rangle = \tfrac{1}{2}\hbar V e^{i\omega t} \tag{3.19}$$

$$\langle e|\mathcal{H}^1|g\rangle = \left(\tfrac{1}{2}\hbar V e^{i\omega t}\right)^* = \tfrac{1}{2}\hbar V e^{-i\omega t}. \tag{3.20}$$

Inserting these into equation 3.9 gives for the time-dependence of the coefficients

$$\dot{d}_e = -\tfrac{1}{2}i\, d_g\, e^{i(\omega_0-\omega)t}V \tag{3.21}$$

$$\dot{d}_g = -\tfrac{1}{2}i\, d_e\, e^{-i(\omega_0-\omega)t}V \tag{3.22}$$

which are exactly our previous results, except that ω_0 has now been replaced by $\omega_0 - \omega$, that is the difference between the frequency of the light and the resonance frequency of the system, and the strength of the perturbation has been halved[6]. In particular if the light is exactly resonant with the transition, so that $\omega_0 - \omega = 0$, then the simple results

$$d_g = \cos(Vt/2), \qquad d_e = -i\sin(Vt/2) \tag{3.23}$$

are recovered. Thus Rabi flopping can be induced by a weak field oscillating in resonance with a transition. The populations of the ground and excited states are given by

$$P_g = \cos^2(Vt/2) = \tfrac{1}{2}[1 + \cos(Vt)] \tag{3.24}$$

$$P_e = \sin^2(Vt/2) = \tfrac{1}{2}[1 - \cos(Vt)] \tag{3.25}$$

and are sinusoidally modulated at a frequency Vt, called the *Rabi frequency*. Note that the Rabi frequency refers to the rate of modulation of the *populations*, not the probability amplitudes, which are modulated at half this frequency[7].

This method can, of course, also be used to calculate the effects of off-resonance excitation, and the key results are implied above. However more insight into this problem can be gained by using the rotating frame transformation and the vector model, which will be discussed in the next chapter. Although the discussion there is formally concerned with spins in magnetic fields, the results can, of course, be applied to any other two level quantum system.

[6]There is considerable variation among (and even within) textbooks as to whether V is taken as the strength of the *oscillating* field (as used here), or as the strength of the *rotating* field; this leads to minor variations in equations, and in particular in the formula for the Rabi frequency. Similarly some authors incorporate a factor of $\hbar$ into V rather than separating it out as done here.

[7]This can be seen as an example of *spinor* behavior: when a spin is coherently rotated from its ground state through an excited state and back to the ground state again its wavefunction picks up a sign of -1.

3.4 Time-dependent perturbation theory

Two approximations were made in deriving the previous result: firstly that we can treat the system as a two level atom, and secondly that the counter-rotating component can be ignored. This seems reasonable in the light of the final result: as the counter-rotating component is far from resonance it will not be effective at inducing Rabi flopping unless the field is very strong. Furthermore, since the light will only induce transitions at frequencies close to ω it seems reasonable to ignore all other excited states. It might, however, be argued that this proof is circular, since the final result is assumed at the start!

To make a more rigorous argument it is necessary to return to equation 3.9, which is exact. We could solve this fairly easily for a two level atom, but with an n level system we will end up having to solve an nth order differential equation. Furthermore, this equation will become extremely complex unless the form of the time-varying perturbation is extremely simple. To make further progress we will have to make approximations from the start, and if the perturbation is *small* then it makes sense to use a power series in $\mathcal{H}^1(t)$.

Consider a multi-level atom, and suppose that the system begins in some initial state $|i\rangle$ and we wish to obtain the amplitude of the system making a transition to some final state $|f\rangle$. The zero order result is obtained by ignoring the perturbation completely (effectively setting $\mathcal{H}^1 = 0$), and substituting this into equation 3.9 gives the trivial result that the coefficients do not evolve. The first order result is then obtained by using the zeroth order wavefunction (that is, the unperturbed coefficients) with the first order Hamiltonian, giving

$$\dot{d}_f = -ie^{i\omega_{fi}t}\langle f|\mathcal{H}^1(t)|i\rangle/\hbar. \tag{3.26}$$

(Note that at small times after the perturbation is first applied all the coefficients will be close to zero, except for d_i which will remain close to one). The solution is

$$d_f(t) = -\frac{i}{\hbar}\int_0^t e^{i\omega_{fi}t'}\langle f|\mathcal{H}^1(t')|i\rangle\,\mathrm{d}t'. \tag{3.27}$$

where t' is just a dummy variable for the integration and we have assumed that $\langle i|\mathcal{H}^1|i\rangle = 0$ as before.

This integral is, of course, a Fourier transform, suggesting that the process will be sensitive to components of $\mathcal{H}^1(t)$ oscillating near the frequency ω_{fi}. Furthermore, because the Fourier transform is *linear* the total effect of applying several different perturbations is simply the sum of the effect of the individual perturbations. In particular it is possible to treat an oscillating perturbation as the sum of two counter-rotating perturbations, equation 3.18, and it is possible to treat *any* perturbation as a sum of oscillating terms. For a single oscillating term with angular frequency ω the solution is

$$d_f(t) = -\frac{i}{\hbar}\int_0^t e^{i(\omega_{fi}-\omega)t'}\hbar V(\omega)\,\mathrm{d}t' = -iV(\omega)\times\left[\frac{e^{i(\omega_{fi}-\omega)t'}}{i(\omega_{fi}-\omega)}\right]_0^t \tag{3.28}$$

$$= -iV(\omega)e^{i(\omega_{fi}-\omega)t/2}t\,\mathrm{sinc}[(\omega_{fi}-\omega)t/2] \tag{3.29}$$

where $\mathrm{sinc}(x) = \sin(x)/x$. The sinc function arises naturally whenever a Fourier transform is taken of an oscillation with a finite extent, and can be considered as measuring the uncertainty in the frequency of the oscillation.

First, consider the case when the oscillation is exactly resonant with the transition, so that $\omega_{fi} - \omega = 0$. Since $\mathrm{sinc}(0) = 1$ equation 3.29 reduces to

$$d_f(t) = -iVt. \tag{3.30}$$

For short times

$$\sin(Vt) \approx (Vt) \tag{3.31}$$

and this result is identical to the previous result for a two level atom, equation 3.15. At longer times the treatment breaks down, as it is no longer reasonable to assume that d_i is always one. This can be overcome by using higher orders of perturbation theory: the conceptually simplest method is to feed the first order wavefunctions back into the algorithm to obtain second order wavefunctions, and so on, giving series expansions of the underlying sine and cosine modulations.

Equation 3.29 can also be used to look at the effects of excitation away from resonance. This will be identical to excitation on-resonance, except that the strength of the interaction is scaled down by $\mathrm{sinc}[(\omega_{fi} - \omega)t/2]$. Clearly the effect will be very small unless ω is close to resonance, justifying our previous decision to use the two-level atom model and to ignore the counter-rotating component of the excitation field. More interestingly, this result shows that excitation becomes more "choosy" as time goes on; this is not particularly surprising, however, as it simply reflects the fact that the frequency of an oscillation becomes better defined as it is observed over a long period.

3.5 Fermi's Golden Rule

The treatment above looks good, but unfortunately clashes with both common sense and common experience. The first clash shouldn't worry you at all (quantum mechanical systems are famous for their strange behavior!), but the second clash is more worrying.

Consider the effect of on-resonance excitation, equation 3.30, and calculate how the population of the final state varies with time. Since $P_f = |d_f|^2$ this is given by

$$P_f(t) = V^2t^2 \tag{3.32}$$

so the degree of excitation varies *quadratically* with time, or, equivalently, the rate at which transitions occur increases linearly with time. In fact, however, the excited state population is often observed to grow *linearly* with time, so that the transition rate is constant. Fortunately this apparent discrepancy is easily explained. So far we have assumed that the energy levels of an atom are perfectly sharp, so that any transition has a single exact frequency, but this is quite untrue. Every excited state of an atom has a finite lifetime (limited by the spontaneous emission lifetime), and so has a corresponding uncertainty in its energy; thus the frequency of a transition is not in fact well defined! It is, therefore, usually necessary to integrate the transition probability over the whole range of transition frequencies, and when this is done[8] it is found that

$$P_f(t) \propto V^2t \tag{3.33}$$

in agreement with naive expectations.

[8]For the detailed derivation see any standard text.

3.6 Rabi or Fermi?

Who then is right? Does light cause an atom to undergo Rabi flopping, or does excitation follow Fermi's Golden Rule? This question can be considered from the point of view of theory and from that of experiment.

The essential reason underlying the linear behavior in equation 3.33 is easy to understand. As previously noted, the system becomes increasingly choosy about whether or not to make a transition as time goes on, and the increasing fussiness counteracts the intrinsic tendency of the transition rate to grow, resulting in a constant transition rate overall. This effect is only important, however, for times which are long in comparison with the inverse of the width of the transition; in effect this means times which are long in comparison with the lifetime of the excited state. The time for which the light is applied will obviously depend on the time it takes to have a significant effect, which is conveniently parameterized by the oscillation frequency in equation 3.15. We can thus distinguish two extreme regimes of behavior, depending on V and the state lifetime τ:

1. Strongly driven transitions: $V\tau \gg 1$
   In this case the system undergoes Rabi oscillations between the ground and excited state. This case is sometimes called *coherent control*, and is suitable for quantum information processing experiments.

2. Weakly driven transitions: $V\tau \ll 1$
   In this case the system obeys Fermi's Golden Rule and a constant transition probability is observed. The long term behavior of the system is described by *rate equations*.

For transitions at optical frequencies, the lifetimes of the excited state are usually fairly short[9], and transitions are usually weakly driven, although it is possible to observe Rabi oscillation behavior, either by using very high power lasers, or by artificially suppressing spontaneous decay[10]. Thus direct optical transitions are not normally suitable candidates for coherent quantum control. The obvious solution is to use transitions to a low lying quantum state, so that the transitions occur at much lower frequencies. As we will see later, transitions between nuclear spin states, which occur at frequencies below 1 Ghz, are easy to drive coherently. There are, however, two problems with this approach.

Firstly, most transitions from the ground state to low lying excited states are *forbidden*, that is the matrix element for the transition $\mathcal{H}^1_{eg} = 0$. Although working out the exact matrix element connecting two states for a given perturbation is quite complicated, it is relatively simple to list *selection rules* which determine whether it is zero (a forbidden transition) or non-zero (an allowed transition). So far we have been considering transitions induced by the interaction between the electric field of light and the electric dipole moment of an atom, and so it is more accurate to state that most transitions from the ground state to low lying excited states are *electric dipole forbidden*. It is, of course, possible to find many low frequency transitions between pairs of excited states which are electric dipole allowed, but in this case *both* states will be broadened by spontaneous emission, and the possibility of coherent control is further suppressed.

[9] From the Einstein A and B coefficients we know that the relative importance of spontaneous decay and driven transitions goes as the third power of the frequency.

[10] This can be achieved by placing the atom is a high finesse optical cavity, so that the system can only emit photons into the resonant modes of the cavity.

A second problem is that low frequency light has a long wavelength, and this makes it difficult to focus. For quantum information processing it is usually necessary to excite one atom without exciting another similar atom which is physically close by. Such selective excitation can only be achieved if the light can be focused down to a spot which is small compared with the separation of the atoms, and this spot size is limited by the wavelength of the light. For visible light this resolution limit will be around 1 μm, but for 1 GHz radiation the limiting separation will be around 1 m.

3.7 Raman transitions

The solution to both these problems is to use *Raman* transitions to connect two low lying energy levels. Many textbooks do not discuss Raman transitions at all, and most of the rest only discuss the *Raman effect* used in spectroscopy, rather than coherent Raman transitions. Fortunately the basic idea is easy to understand.[11]

Consider a system with three energy levels, $|g\rangle$ and $|e\rangle$, which form the basis of our qubit, and an additional level $|a\rangle$. We will assume that transitions between $|g\rangle$ and $|e\rangle$ are forbidden, but that both of these can make transitions to $|a\rangle$. Suppose the system is illuminated by two lasers, one in resonance with each of the two allowed transitions. The result of this process will be a complicated evolution of the system between the three states, with transfers from $|g\rangle$ to $|e\rangle$ occurring *via* the additional state $|a\rangle$. This (in principle) solves the problem of making forbidden transitions, but is not an effective solution for two reasons. Firstly, the system cannot go from $|g\rangle$ to $|e\rangle$ without passing through $|a\rangle$, and thus we no longer have a proper two level system, and secondly it remains hard to drive the system strongly enough that Rabi behavior occurs.

The solution is to tune both lasers *away* from the frequencies of the two allowed transition by the same amount, so that the energy difference between photons in the two beams still matches the energy gap between $|g\rangle$ and $|e\rangle$. The remarkable result is that although the two allowed transitions no longer occur, Rabi flopping occurs for the *forbidden* transition between $|g\rangle$ and $|e\rangle$. This is an example of a *two photon* process: in effect a photon is absorbed from one laser beam, while the other beam stimulates the emission of a second photon. The transition is sometimes described as occurring via a *virtual state*, but in fact occurs via off-resonance interactions with the (real) state $|a\rangle$. Because these transitions are off-resonance the Rabi frequency is scaled down from its naive value Ω by a factor Ω/Δ, where Δ is the frequency offset from resonance, but this is not a major problem. An important advantage is that the system can be driven strongly, as the relevant state lifetimes (τ, see the previous section) are those of $|g\rangle$ and $|e\rangle$; the lifetime of the additional state $|a\rangle$ is irrelevant[12] as this state is never populated! Raman transitions provide an almost ideal solution to the problem of inducing Rabi transitions between atomic energy levels, and very commonly used[13].

[11] For a slightly more detailed treatment see the notes by Dieter Jaksch.

[12] This is only true in an ideal world; in real life there is always some population of the additional state and its lifetime cannot be completely ignored.

[13] The most popular alternative so far in experimental implementations of quantum information processing has been to use so called quadrupole transitions, which are electric dipole forbidden but can be induced by strong laser fields. More recently some researchers have considered using very intense lasers to perform direct Rabi flopping on electric dipole transitions.

Chapter 4

Spins in magnetic fields

Spins in magnetic fields provide one of the simplest and most natural physical systems for implementing quantum bits; indeed the relationship between a spin and a qubit is so close that the terms are sometimes used interchangeably. Experimental spin physics is rarely studied in physics courses in the UK, which is a pity, as it provides one of the simplest examples of coherent quantum control available. The treatment is essentially identical to that of two level atoms in laser fields, except that transitions can almost always be treated as strongly driven. For details see [Goldman 1988].

4.1 The nuclear spin Hamiltonian

Just like electrons, atomic nuclei possess an intrinsic angular momentum, called spin[1]. This arises from the coupling between the intrinsic spins of the protons and neutrons making up the nucleus. A nucleus with spin quantum number I has spin angular momentum $\hbar\boldsymbol{I}$ and an associated magnetic moment $\boldsymbol{\mu}$, given by

$$\boldsymbol{\mu} = \gamma\hbar\boldsymbol{I} \tag{4.1}$$

where γ is called the *gyromagnetic ratio* of the nucleus, and is in some sense analogous to the Landé g-value of an electron in an atom. Although it is possible to calculate these properties from first principles, for most purposes it is best to treat the details of nuclear spins as experimentally measured quantities.

If the spin is placed in a magnetic field $\boldsymbol{B}$ then the interaction between the magnetic moment and the field is described by the Zeeman Hamiltonian

$$\mathcal{H} = -\boldsymbol{\mu}\cdot\boldsymbol{B} \tag{4.2}$$

and the standard convention is to orient the z-axis along the magnetic field, so that

$$\mathcal{H} = -\mu_z B = -\hbar\gamma B I_z = -\hbar\omega_L I_z \tag{4.3}$$

where I_z is the projection of $\boldsymbol{I}$ onto the z-axis and ω_L is called the *Larmor frequency*. As this is a quantum mechanical system, I_z cannot taken any value, but can only take values between $-I$ and I in integer steps; the simplest situation is when $I = \frac{1}{2}$ (a spin-$\frac{1}{2}$ nucleus), in which case there

[1]It is also possible to use electron spins as qubits; the basic techniques are very similar, but electron transitions usually occur at higher frequencies than nuclear transitions, as electrons have a much larger magnetic moment

are only two possible values, $I_z = \pm\frac{1}{2}$. The most important example of a spin-$\frac{1}{2}$ nucleus is the hydrogen (^{1}H) nucleus, but several others exist, most notably ^{13}C, ^{15}N, ^{19}F and ^{31}P.

The effect of a magnetic field on a spin-$\frac{1}{2}$ nucleus is to split apart the two spin states[2], with a splitting $\hbar\omega_L$, and these two energy levels provide an obvious implementation of a qubit. Note that in this case the system really does have only two levels, and so we do not need to make a two-level approximation. The transitions between these two spin state are *electric dipole forbidden*, as they violate the electric dipole selection rule $\Delta S = 0$, but they can be induced by magnetic fields. Another way of looking at this is that the electric field matrix element $\langle 1|\boldsymbol{E}|0\rangle = 0$, while the magnetic field matrix element $\langle 1|\boldsymbol{B}|0\rangle$ will be non-zero as long as the magnetic field is not parallel to the z-axis. Thus if a strong magnetic field is suddenly applied at right angles to the main magnetic field then transitions between the two spin states will occur. More realistically, the same effect can be achieved by applying an weak oscillating magnetic field as long as it oscillates in resonance with the transition, that is at the Larmor frequency.

4.2 The rotating frame

These transitions can be treated using exactly the same techniques as we used previously to study transitions in a two level atom, but it is more common to use a subtly different (though ultimately equivalent) approach, based on transforming the problem into a *rotating frame*. Consider a general wavefunction $|\psi\rangle$, which we choose to write as

$$|\psi\rangle = U|\tilde{\psi}\rangle \tag{4.4}$$

where U simply describes the transformation between two different bases which can be used to describe the wavefunction Note that

$$|\tilde{\psi}\rangle = U^{-1}|\psi\rangle = U^\dagger|\psi\rangle \tag{4.5}$$

where we have used the fact that basis-state transformations are unitary. If we transform the wavefunction into a new basis then we must also transform the Hamiltonian, and this transformation can be worked out using the time-dependent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}|\psi\rangle = \mathcal{H}|\psi\rangle. \tag{4.6}$$

In the transformed basis

$$i\hbar\frac{\partial}{\partial t}|\tilde{\psi}\rangle = i\hbar\frac{\partial}{\partial t}\left(U^\dagger|\psi\rangle\right) \tag{4.7}$$

$$= i\hbar\left[U^\dagger\frac{\partial}{\partial t}|\psi\rangle + \left(\frac{\partial U^\dagger}{\partial t}\right)|\psi\rangle\right] \tag{4.8}$$

$$= \left[U^\dagger\mathcal{H} + i\hbar\left(\frac{\partial U^\dagger}{\partial t}\right)\right]|\psi\rangle. \tag{4.9}$$

[2]There is enormous variation in the notation used to describe these two spin states in the literature. Two relatively common notations are to call the spins states α and β, or to call them spin-up ($\uparrow$) and spin-down ($\downarrow$). As usual we will avoid these arguments by calling the two states $|0\rangle$ and $|1\rangle$ or $|g\rangle$ and $|e\rangle$.

Using equations 4.4 and 4.6 gives

$$i\hbar\frac{\partial}{\partial t}|\tilde{\psi}\rangle = \left[U^\dagger \mathcal{H} U + i\hbar\left(\frac{\partial U^\dagger}{\partial t}\right)U\right]|\tilde{\psi}\rangle = \tilde{\mathcal{H}}|\tilde{\psi}\rangle. \tag{4.10}$$

and so the transformed Hamiltonian has the form

$$\tilde{\mathcal{H}} = \left[U^\dagger \mathcal{H} U + i\hbar\left(\frac{\partial U^\dagger}{\partial t}\right)U\right]. \tag{4.11}$$

The first term in the transformed Hamiltonian is simply the obvious transformation of $\mathcal{H}$ into the new basis, but the second term is more subtle. This term is zero for fixed transformations, and corresponds to a *fictitious energy*, which is analogous to the fictitious forces which arise in classical mechanics when working in accelerating frames.

To take a concrete example, we consider a spin-$\frac{1}{2}$ particle in a static magnetic field along the z-axis and experiencing an oscillating magnetic field at right angles. This oscillating field can be achieved by using the magnetic component of an appropriate oscillating electromagnetic field, that is light at the resonance frequency[3]. Thus the Hamiltonian can be written in matrix form as

$$\mathcal{H} = \begin{pmatrix} -\frac{1}{2}\hbar\omega_0 & \hbar V\cos\omega t \\ \hbar V\cos\omega t & \frac{1}{2}\hbar\omega_0 \end{pmatrix} \tag{4.12}$$

where we have used $|g\rangle$ and $|e\rangle$ as our basis states and have chosen to place the energy zero half way between our two states. We then choose the transformation

$$U = \begin{pmatrix} e^{i\omega t/2} & 0 \\ 0 & e^{-i\omega t/2} \end{pmatrix} \tag{4.13}$$

which corresponds to using basis states which rotate in synchrony with one component of the oscillating field. Applying equation 4.11 the Hamiltonian in this new frame is

$$\tilde{\mathcal{H}} = \begin{pmatrix} \frac{1}{2}\hbar(\omega-\omega_0) & \hbar V\cos(\omega t)e^{-i\omega t} \\ \hbar V\cos(\omega t)e^{i\omega t} & -\frac{1}{2}\hbar(\omega-\omega_0) \end{pmatrix}. \tag{4.14}$$

Next we define the *detuning* as $\delta = \omega - \omega_0$ and separate the oscillating term into two counter-rotating terms. Finally we apply the rotating wave approximation[4] as before, and simply ignore the rapidly varying terms. Thus to a good approximation

$$\tilde{\mathcal{H}} = \begin{pmatrix} \frac{1}{2}\hbar\delta & \frac{1}{2}\hbar V \\ \frac{1}{2}\hbar V & -\frac{1}{2}\hbar\delta \end{pmatrix}. \tag{4.15}$$

It is important to remember that although this result has been derived for the case of a spin in a magnetic field, the method is entirely general, and an identical result could have been derived for an atom in a laser field: all two level quantum systems (qubits) are basically the same!

[3] As we shall see this turns out to correspond to radio-frequency (RF) radiation.

[4] The rotating wave approximation is not quite so good in this case as it was for transitions between atomic energy levels as the frequencies involved are much lower. Careful calculations indicate that the counter-rotating component gives rise to a small shift in the transition frequencies known as a Bloch–Siegert shift. This is an example of a more general phenomenon called the AC Stark shift which is discussed in [Budker 2004].

4.3 On-resonance excitation

As usual the simplest case occurs when the excitation is on-resonance, so that $\delta = 0$ and the Hamiltonian in the rotating frame is

$$\tilde{\mathcal{H}} = \begin{pmatrix} 0 & \frac{1}{2}\hbar V \\ \frac{1}{2}\hbar V & 0 \end{pmatrix} \tag{4.16}$$

which is clearly related to one of the Pauli matrices, that is

$$\tilde{\mathcal{H}} = \tfrac{1}{2}\hbar V \sigma_x. \tag{4.17}$$

We can calculate the evolution under this Hamiltonian by using the method of propagators, described in section 1.10, to get

$$\tilde{U} = \exp(-i\tilde{\mathcal{H}}t/\hbar) = \exp(-i\theta\sigma_x) \tag{4.18}$$

where $\theta = Vt/2$. Now we have previously derived a formula for the matrix exponential of σ_x (equation 2.14), and so know that

$$\tilde{U} = \begin{pmatrix} \cos(Vt/2) & -i\sin(Vt/2) \\ -i\sin(Vt/2) & \cos(Vt/2) \end{pmatrix}. \tag{4.19}$$

If the system starts off in the ground state $|g\rangle$ then the state at later times is given by

$$\tilde{\psi} = \tilde{U}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(Vt/2) \\ -i\sin(Vt/2) \end{pmatrix} \tag{4.20}$$

in complete agreement with equation 3.23.

None of this should be surprising: it is all exactly as expected from the discussion in section 2.4 where we considered how to implement quantum logic gates, and showed that a NOT gate could be implemented by applying a Hamiltonian proportional to σ_x for an appropriate time. A NOT gate interconverts $|0\rangle$ and $|1\rangle$, which is exactly what occurs in Rabi flopping.

4.4 Excitation phases

We have seen that resonant radiation can be used to produce a Hamiltonian proportional to σ_x, but in order to implement general single qubit gates it is useful to be able to implement Hamiltonians proportional to σ_y. This can be achieved by simply altering the *phase* of the radiation, so that the perturbation takes the form $\hbar V\cos(\omega t + \phi)$. The overall Hamiltonian can then be written as

$$\tilde{\mathcal{H}} = \begin{pmatrix} 0 & \frac{1}{2}\hbar V[e^{i(\omega t+\phi)} + e^{-i(\omega t+\phi)}]e^{-i\omega t} \\ \frac{1}{2}\hbar V[e^{i(\omega t+\phi)} + e^{-i(\omega t+\phi)}]e^{i\omega t} & 0 \end{pmatrix} \tag{4.21}$$

(where we have assumed the radiation is applied on-resonance) and making the rotating wave approximation as usual gives

$$\tilde{\mathcal{H}} = \begin{pmatrix} 0 & \frac{1}{2}\hbar V e^{i\phi} \\ \frac{1}{2}\hbar V e^{-i\phi} & 0 \end{pmatrix} = \tfrac{1}{2}\hbar V(\sigma_x\cos\phi + \sigma_y\sin\phi). \tag{4.22}$$

Thus by appropriate choice of ϕ we can generate Hamiltonians proportional to σ_x, or σ_y, or at any angle between them. If we take the case $\phi = \pi/2$, so that $\tilde{\mathcal{H}} \propto \sigma_y$, then the evolution propagator is

$$\tilde{U} = \begin{pmatrix} \cos(Vt/2) & -\sin(Vt/2) \\ \sin(Vt/2) & \cos(Vt/2) \end{pmatrix} \tag{4.23}$$

and the system undergoes Rabi oscillations at the same frequency as before. If only populations are considered then the phase of the radiation has no effect, but if the amplitudes of the two states are considered then the phase is important.

Sceptical readers might point out that the *absolute phase* of an oscillation is largely meaningless, with only the relative phase of two oscillations being well defined. This is correct, and leads to a corresponding result that it is not possible to define an absolute phase for a single Rabi pulse, but it is possible to define a relative phase for two or more pulses. This will be explored in the discussion of Ramsey fringes below.

4.5 Off-resonance excitation

Next we consider the case when the radiation is not quite in resonance with the transition frequency, so that the Hamiltonian takes the general form, equation 4.15. The propagator is then

$$\tilde{U} = \exp\left[-i \times \begin{pmatrix} \delta/2 & V/2 \\ V/2 & -\delta/2 \end{pmatrix} \times t\right] \tag{4.24}$$

and brute force calculation[5] gives the result

$$\tilde{U} = \begin{pmatrix} \cos(\Omega t/2) - i(\delta/\Omega)\sin(\Omega t/2) & -i(V/\Omega)\sin(\Omega t/2) \\ -i(V/\Omega)\sin(\Omega t/2) & \cos(\Omega t/2) + i(\delta/\Omega)\sin(\Omega t/2) \end{pmatrix} \tag{4.25}$$

where $\Omega = \sqrt{V^2 + \delta^2}$. Note that on-resonance $\delta = 0$ and $\Omega = V$, and our previous results are recovered. Off-resonance, we see that the frequency of the Rabi oscillations is increased ($\Omega > V$), but the efficiency is reduced ($|0\rangle$ cannot be completely converted to $|1\rangle$).

Seen from the conventional point of view, off-resonance excitation is a bad thing, but from the viewpoint of quantum control it provides a direct route to other quantum logic gates. The most important example occurs in the case when $V = \delta$, so that $\Omega = \sqrt{2}V$. Choosing t such that $\Omega t/2 = \pi/2$ gives

$$\tilde{U} = -i \times \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix} \tag{4.26}$$

which (neglecting an irrelevant global phase) is the Hadamard gate, one of the most important single qubit logic gates. However it is usually simpler in practice to consider only on-resonance excitation, and to construct gates such as the Hadamard using gate networks as described in section 2.6.

[5]Essentially this requires finding the eigenvalues and eigenvectors as described previously; a symbolic mathematics program such as Mathematica or Maple is a great help.

4.6 Practicalities

I have already hinted that spins in magnetic fields are, in some sense, a more quantum mechanical system than atoms in a laser field. The reason for this is not any fundamental property of the two systems, but simply a matter of practicalities. The most important difference is that the transitions between nuclear spin energy levels occur at very much lower frequencies.

It is obvious from section 4.1 that nuclear spin transition frequencies depend both on the magnetic field strength used and on intrinsic properties of the nuclei. The largest static magnetic fields available to us[6] are around 20 T, and the most sensitive of the stable nuclei is ^{1}H (hydrogen), for which transitions frequencies in the range up to 1 GHz are found, corresponding to the radio-frequency (RF) portion of the spectrum. Most studies of ^{1}H take place at frequencies in the range 400–800 MHz, and studies of other nuclei (except for ^{19}F and the dangerously radioactive nucleus ^{3}H) take place at significantly lower frequencies.

There are two principal advantages of working with RF frequencies. The first is that spontaneous emission rates at these low frequencies are completely negligible, and so it should be easy to reach the coherent control region. The second advantage is that RF radiation is extremely easy to generate and control. While experimentalists working with lasers have to work hard to control the frequency, amplitude and phase of laser light, any desired RF pattern can be obtained simply by asking a computer to generate it! For this reason coherent control of nuclear spins has flourished for decades under the name of *nuclear magnetic resonance* or *NMR*.

There are, however, two major disadvantages of working with RF. The first is that the wavelength of RF radiation is so large that spatially selective excitation is essentially impossible, as discussed previously. The second is that the energy of RF photons is so small that it is virtually impossible to detect single photons. Both of these problems have major consequences for the use of NMR as an implementation of quantum information processing.

4.7 The vector model

There is another way of looking at spins in magnetic fields, usually called the vector model, which was developed by Bloch.[7] This is an entirely classical method for thinking about the situation, but it turns out to give a pretty accurate description of a single isolated spin;[8] by the obvious extension it can also be used to describe any other single qubit system.

We have already seen that the state of a spin can be represented as a Bloch vector, pointing from the origin to an appropriate point on the Bloch sphere. The vector model represents a spin by a classical magnetic moment pointing along this vector. If the spin is placed in a magnetic field along the z-axis then it will *precess* around the field, at the Larmor frequency, which depends on the strength of the magnetic field and the size of the magnetic moment. This process corresponds perfectly with the way in which the two basis states $|0\rangle$ and $|1\rangle$ pick up a relative phase shift at the Larmor frequency.

[6] These fields are achieved using superconducting electromagnets, and are limited by the *critical field* and *critical current* of the superconducting wires; larger fields are available for *short* periods of time by using pulsed electromagnets, and extremely large fields are available for very short times using destructive techniques.

[7] The vector model is very widely used in the field of NMR; see, for example, [Freeman 1998].

[8] This is a consequence of Ehrenfest's theorem.

The effect of resonant RF fields can be treated in much the same way. The oscillating magnetic field component is divided into two counter-rotating components in the xy-plane, one of which rotates around the field in the same direction and at the same rate as the spin. If we transform into a rotating frame which also goes round in the same way then both the magnetization and the RF field component will appear to be static in the xy-plane (the exact position depending on the phase of the RF). The situation now looks just like a magnetic moment in a normal magnetic field, and the spin will precess around the RF field component (which is along, say, the x-axis) at a rate which depends on its strength. It is also clear why the counter-rotating component can be ignored: this is moving so fast that the spin sees it as a rapidly fluctuating field which basically cancels out.

One might ask what has happened to the main magnetic field in this picture. The long answer is that the rotating-frame transformation is an example of a gauge transformation, which results in a gauge field, in this case a fictitious magnetic field which exactly cancels the main field. The short answer is that since the spin does not precess around the field direction in the rotating frame, then the field cannot be there!

The vector model can also be used to model off-resonance excitation. In this case the frame rotates at the RF frequency, not the Larmor frequency, and so the spin is not quite static. This means that the fictitious field does not quite cancel the main field, and a small residual magnetic field remains. The total field experienced by the spin is then the vector sum of the residual field along z and the excitation field along x, and the spin precesses around this vector sum. This sum is longer than the excitation field, and so the precession frequency (Rabi frequency) is increased, but it is tilted away from the xy-plane, so that precession will no longer drive the spin from the $+z$ to the $-z$ axis (from $|0\rangle$ to $|1\rangle$).

Finally, as always, it is important to remember that the vector model is not peculiar to the description of nuclear spins, although that is where it is most frequently used. The underlying nature of any two level quantum system interacting with a radiation field is basically the same, and so all these ideas can equally well be applied to atoms in laser fields. This approach ultimately leads to the *optical Bloch equations*, which are analogous to the Bloch equations used to describe the vector model in NMR systems.

4.8 Single qubit experiments

We have now seen two different possible implementations of a qubit. These are clearly very closely related to one another: although each implementation has its own traditional language, each can be described using the other's language, or by the common language of qubits. We shall now consider some simple single qubit experiments, and show how they can be considered as networks of quantum gates.

The simplest experiment in coherent quantum control is Rabi flopping, which we have already considered in some detail. The basic idea is that a quantum system begins in some ground state $|g\rangle$, and radiation is applied which is resonant with an allowed transition to some excited state $|e\rangle$. The populations of the ground and excited states are measured as a function of the time for which the radiation is applied.

As we have seen the Hamiltonian for this system[9] can be written as

$$\mathcal{H} = \hbar \times \begin{pmatrix} 0 & \Omega/2 \\ \Omega/2 & 0 \end{pmatrix} = \hbar\Omega\,\sigma_x/2 \tag{4.27}$$

where Ω is the Rabi frequency. Neglecting global phases, the corresponding propagator is

$$U = \begin{pmatrix} \cos(\Omega t/2) & -i\sin(\Omega t/2) \\ -i\sin(\Omega t/2) & \cos(\Omega t/2) \end{pmatrix}. \tag{4.28}$$

This propagator has already been examined in some detail, but it can be viewed in a quite different way as the nth-POWER-OF-NOT quantum logic gate[10], with $n = \Omega t/\pi$. When $\Omega t = \pi$ (a π-pulse, or 180° pulse) we get a NOT gate, which interconverts $|0\rangle$ and $|1\rangle$, while a 90° pulse produces equally weighted superpositions of $|0\rangle$ and $|1\rangle$.

4.9 Ramsey fringes

Ramsey fringes occur when two 90° pulses are applied to a two level quantum system, separated by a time period during which the system is allowed to undergo free evolution. The overall result is a oscillation with a frequency depending on the energy gap between the two levels of the system. Here we will show how this situation can be analyzed using a gate network.

The details of the analysis depends on the exact form of the 90° pulses, and in particular what phase they have. The simplest situation occurs when they can be treated as Hadamard gates; as shown in section 2.6 these are closely related to 90° pulses. Next we must consider how to represent the period of free precession, during which the system evolves under the background Hamiltonian. Working in the rotating frame this takes the form

$$\tilde{\mathcal{H}} = \begin{pmatrix} \frac{1}{2}\hbar\delta & 0 \\ 0 & -\frac{1}{2}\hbar\delta \end{pmatrix} = \hbar\delta\sigma_z/2 \tag{4.29}$$

and so the propagator describing the evolution is

$$U = \exp(-i\mathcal{H}t/\hbar) = \exp(-i\delta t\,\sigma_z/2) = \cos(\delta t/2)\sigma_0 - i\sin(\delta t/2)\sigma_z \tag{4.30}$$

where the last step is based on equation 2.14.

The experiment can now be described by the gate network HUH, and using the linearity of matrix operations

$$HUH = \cos(\delta t/2)H\sigma_0 H - i\sin(\delta t/2)H\sigma_z H. \tag{4.31}$$

This expression can be simplified using $H\sigma_0 H = H^2 = \sigma_0$ and $H\sigma_z H = \sigma_x$ to obtain

$$HUH = \cos(\delta t/2)\sigma_0 - i\sin(\delta t/2)\sigma_x = \exp(-i\delta t\,\sigma_x/2) \tag{4.32}$$

showing that the overall effect of the sequence is to perform Rabi flopping at a frequency which depends on the *internal frequency*, $\delta = (E_1 - E_0)/\hbar$, of the system.

[9] Strictly speaking this is the Hamiltonian in the rotating frame after making the rotating wave approximation; however we are free to work in any convenient frame and the approximation is generally good.

[10] As usual we are ignoring global phases.

So why is this experiment known as Ramsey fringes? Suppose that the system starts off in the ground state $|0\rangle$, and after a time t we measure the state in the computational basis. The probability that the system is still in the ground state is then $\cos^2(\delta t/2)$. In otherwise the signal is *time-modulated*, and the modulation takes exactly the same form as the spatial modulation seen in a two-slit experiment.

Now consider the more general case, when the two pulses are arbitrary 90° pulses. A detailed analysis shows that the results are just the same, but that the phase of the fringes is shifted, with the phase angle depending on the relative phase of the two pulses; the case described previously is recovered when the phases of the two pulses are 180° apart.

4.10 Interferometry

We have already noted the similarity between Ramsey fringes and the two slit experiment, but there is an even closer link with another experiment, the Mach–Zender interferometer. The key component here is a *beam splitter*, which for a conventional interferometer (using visible light) is a half silvered mirror, which has the property that if a photon is incident on the mirror it has a 50% chance of being transmitted and a 50% chance of being reflected. If the reflected and transmitted beams are then recombined using a second beam splitter, then conventional interference effects are seen. If the interferometer is correctly set up, then light will only emerge on one side of the second beam splitter, showing that the beam splitters split the light wave coherently, rather than in a naive probabilistic fashion. Furthermore, the output result can be controlled by applying a relative phase shift to the two beams in the interferometer.

A detailed analysis shows that Mach–Zender interferometry is essentially identical to the Ramsey fringes experiment, with the two beam splitters corresponding to the two 90° pulses, and the relative phase shift corresponding to the period of free evolution. For this reason beam splitters (which can be built for many different particles, not just photons) are sometimes described as Hadamard gates.

4.11 Spin echoes

We have already analyzed spin echoes as quantum networks in section 2.6, but it is interesting to look at them using the vector model. Suppose we have a spin which starts along the x-axis, and we allow it to precess at its own Larmor frequency ω for a time t; this will cause it to rotate around the xy-plane through an angle $\phi = \omega t$. Next we apply a NOT gate, which is a 180° rotation around the x-axis. This leaves the spin within the xy-plane, but moves it to a position with angle $-\phi$. After precessing at the same rate ω for another period t the spin is once more found along the x-axis. The second NOT gate has no effect,[11] and so the overall effect is that the spin ends the sequence precisely where it started.

[11] The second NOT gate is necessary to deal with spins that start out away from the x-axis.

Chapter 5

Two qubits and beyond

As we have seen, even a single qubit is a surprisingly interesting object. However the real power of quantum information processing begins with systems of two ore more qubits. Before studying these in detail we need to expand our notation a little.

Consider a system of two qubits, labeled a and b, each of which has two basis states, $|0\rangle$ and $|1\rangle$. The whole system then has four basis states, which can be written as $|0_a 0_b\rangle$, $|0_a 1_b\rangle$, $|1_a 0_b\rangle$, and $|1_a 1_b\rangle$, and can be found in any general superposition of these states, so that it occupies a four-dimensional Hilbert space. In the same way, a system of three qubits inhabits an eight-dimensional Hilbert space, and so on. This exponential increase in the size of the Hilbert space with a linear increase in the number of qubits underlies the power of quantum computers.

5.1 Direct products

The size of the Hilbert spaces involved can also be a huge problem, however, making it difficult to describe states of systems with many qubits. A partial solution is to note that some states can be described in a simpler way, using the concept of *direct products*. These states, in which the individual qubits can in principle be discussed separately, make up a tiny minority of the states accessible to a multi-qubit system, but include many important states, most notably the basis states. States of this kind are said to be *separable*, and states which are not separable are said to be *entangled*. Entangled states are much more interesting than separable ones, but it is wise to begin with the simpler case.

By the basis state $|1_a 0_b\rangle$ we mean a state where qubit a is in state $|1_a\rangle$ and qubit b is in state $|0_b\rangle$, and we can write this as $|1_a\rangle \otimes |0_b\rangle$, where the symbol $\otimes$ indicates a direct product. For the moment we shall not worry too much about what a direct product really is, and just think of it as a way of combining two subsystems; a more mathematical discussion can be found in the next section. There is considerable variation in the way these states are described: $|1_a 0_b\rangle$ can also be written as $|1\rangle \otimes |0\rangle$, as $|10\rangle$, or most simply of all as $|2\rangle$, where this last version is obtained by interpreting the 1 and the 0 as the two bits making up the binary number 10, or decimal 2. Most authors move back and forth between these different notations, using whatever is most convenient at the time. Of course, if the more compact forms are used, it is essential to use a consistent ordering of the qubits to avoid ambiguous notation.

The direct product approach can also be used to describe more complex states. Suppose, for

example, as Hadamard gate is applied to the second qubit of a system starting in the state $|00\rangle$. This can be written as

$$|00\rangle = |0\rangle \otimes |0\rangle \xrightarrow{\mathrm{H}_b} |0\rangle \otimes (|0\rangle + |1\rangle)/\sqrt{2} = (|00\rangle + |01\rangle)/\sqrt{2}. \tag{5.1}$$

Similarly, direct products can be used to write down single-qubit operators in a multi-qubit system without the need for explicit labels: thus, for example, $\mathrm{H}_b = \mathbb{1} \otimes \mathrm{H}$ (that is, do nothing to the first qubit and apply a Hadamard to the second qubit), while $\mathrm{H}_a = \mathrm{H} \otimes \mathbb{1}$. Simultaneous Hadamard gates can also be applied to both qubits, using $\mathrm{H}^{(2)} = \mathrm{H} \otimes \mathrm{H}$.

5.2 Matrix forms

Much of the point of the direct product approach is to avoid writing out explicit matrix descriptions of states of multi-qubit systems, but sometimes it is useful to do so. The basic idea behind a direct product is to multiply a copy of the second matrix by each element of the first matrix in turn: thus

$$\begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} a\alpha \\ a\beta \\ b\alpha \\ b\beta \end{pmatrix}. \tag{5.2}$$

Note that, for example, the matrix representation of $|10\rangle$ is

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \tag{5.3}$$

exactly what would be naively expected. An equivalent approach can be used for operators, so that the matrix representation of H_b is

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}. \tag{5.4}$$

When an operator separately affects two different qubits it my be useful to use the fact that the operator can be considered as two sequential operators, one affecting each qubit; thus applying $\mathrm{H}^{(2)}$ is the same as applying H_a followed by H_b, or *vice versa*. Similarly direct products and conventional matrix products can be carried out in either order,

$$(a \otimes b) \cdot (c \otimes d) = (a \cdot c) \otimes (b \cdot d). \tag{5.5}$$

These methods frequently allows calculations to be simplified.

5.3 Two qubit gates

We have already seen some two qubit gates: for example H_b implements a single qubit Hadamard in a two qubit system, while $H^{(2)}$ represents simultaneous Hadamard gates in a two qubit system. However these gates, which can all be written using direct products, are in some sense a trivial extension of the corresponding gates in a single qubit system, and are usually described as single qubit gates. A much more interesting two-qubit gate is the controlled-NOT gate which has the explicit matrix form

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{5.6}$$

and a little thought shows that this matrix cannot be written as a direct product. It seems that this operator might be more interesting than those discussed above, and this is indeed the case. In particular the controlled-NOT gate is a key gate in the generation of entangled states. Furthermore, it can be shown that the combination of the controlled-NOT gate and a small set of single qubit gates is *universal* for quantum information processing, meaning that any desired operation can be built from a network of these gates. The proof of this statement is quite complex, but an outline can be found in [Stolze 2004].

The reason why this gate is called a controlled-NOT gate can be easily seen by applying it to the four basis states in turn, effectively evaluating its *truth table*. Clearly $|00\rangle$ and $|01\rangle$ are unaffected, while $|10\rangle$ and $|11\rangle$ are interchanged. Thus, the effect of the controlled-NOT gate is to apply a NOT gate to the second qubit *if and only if* the first qubit is in state $|1\rangle$. This is an example of controlled evolution, in which the state of one qubit is used to influence the state of another, a process at the heart of computation.

Yet another way of looking at the action of the controlled-NOT gate is to use the concept of *bitwise addition modulo 2*, which simply means adding two bits, throwing away any carries that are generated. Thus $0 \oplus 0 = 0$, and $0 \oplus 1 = 1 \oplus 0 = 1$ as normal, but $1 \oplus 1 = 0$. Note that $a \oplus b$ is equal to zero if a and b are the same, and is equal to one if a and b are different. Alternatively, $a \oplus b$ is equal to the XOR (exclusive-OR) of a and b.

A final description of the controlled-NOT gate is provided by noting that NOT is equivalent to the X gate, and then seeking an expansion of the gate in terms of direct products. The action of controlled-NOT is to apply $\mathbb{1}$ to qubit 2 if qubit 1 is in state $|0\rangle$, and to apply X to qubit 2 if qubit 1 is in state $|1\rangle$. This can be written as

$$\text{controlled-NOT} = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes \sigma_x \tag{5.7}$$

which can be confirmed by direct matrix calculations. Note that in cases such as this, where an operator is written as a sum of direct products, it is essential to be careful about global phases: it is not possible to simply replace the σ_x in equation 5.7 by 180_x as these differ by a factor of i.

The controlled-NOT gate is commonly used in theoretical discussions of quantum information processing, but in many experimental implementations it is easier to use a closely related gate, the controlled-Z gate, which performs the transformation

$$|11\rangle \xrightarrow{c-Z} -|11\rangle \tag{5.8}$$

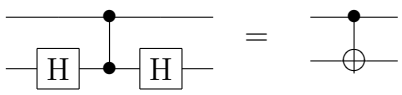
while leaving the other three basis states unaffected. This can be converted to a controlled-NOT gate using a pair of Hadamard gates; the equivalence can be proved by brute force multiplication

$$\frac{1}{2}\begin{pmatrix}1 & 1 & 0 & 0\\ 1 & -1 & 0 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & 1 & -1\end{pmatrix}\cdot\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}\cdot\begin{pmatrix}1 & 1 & 0 & 0\\ 1 & -1 & 0 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & 1 & -1\end{pmatrix}=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{pmatrix} \tag{5.9}$$

or by more cunning methods.

5.4 Networks and circuits

Unlike networks of single qubit gates, networks of two qubit gates are usually difficult to describe in simple words, and it is much better to draw a *quantum circuit*. For example, equation 5.9 can be redrawn as a network as shown below.

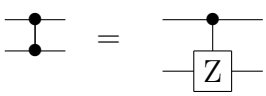


(5.10)

In circuits like this each line corresponds to one qubit, and time runs from left to right, so that the leftmost gate is the first gate applied. Single qubit gates are drawn on the relevant line, while two qubit gates connect two lines.

The gate on the far right is a controlled-NOT gate, and the symbol is made up of three parts. On the top line is a small filled circle, indicating that this qubit *controls* the gate. This control mark is connected by a vertical line to the symbol $\oplus$ on the second qubit (the *target* qubit), and this symbol indicates a NOT gate (this is a slight abuse of notation, which can be partly justified by noting that $a \oplus 1 = \text{NOT}(a)$).

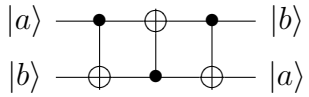
On the left hand side we have three gates, including two single-qubit Hadamard gates, applied to the second qubit, and a peculiar two-qubit gate, comprising two control dots connected by a control line. Once again this is an example of abuse of notation, and is used to indicate a controlled-Z gate,



(5.11)

The justification for this abuse is the fact that the controlled-Z gate, unlike the controlled-NOT gate, is *symmetric* between the control and target qubits: it is not meaningful to say which is which, as the nominal roles could be interchanged with no effect! This symmetry is a characteristic of physical interactions, and explains the importance of the controlled-Z gate in physical implementations.

An interesting circuit which can be built entirely out of controlled-NOT gates is the SWAP circuit shown below

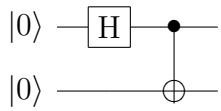


(5.12)

which acts to interchange the states of two different qubits. Exploring this circuit is a good first exercise in playing with gates.

5.5 Entangled states

We have already hinted at the existence of entangled states, which are states of a system of two or more qubits which cannot be written as a direct product of single qubit states. Here we will confine ourselves to two-qubit systems where the phenomenon of entanglement is relatively simple and well understood. A simple way to generate an entangled state from a basis state is to use the network



(5.13)

and we can follow through this network a step at a time

$$|00\rangle \xrightarrow{\mathrm{H}_a} (|00\rangle + |10\rangle)/\sqrt{2} \tag{5.14}$$

$$\xrightarrow{c-X} (|00\rangle + |11\rangle)/\sqrt{2} \tag{5.15}$$

where the first line follows from the properties of the Hadamard gate and the second line is obtained by using the fact that the controlled-NOT gate is a unitary operation, and thus a linear operation, and so its effect on a superposition can be obtained by applying the gate to each of the terms in turn. The final state looks simple enough, but has some very peculiar properties! This state is *inseparable*, which means that it cannot be written as a direct product of states of the two individual qubits. In turn this means that the properties of the state cannot be fully described by listing the properties of the two qubits involved: rather they are properties of the two qubit state taken as a whole.

To take a simple example, consider measuring the state of the first qubit. The system is in an equally weighted superposition of two states, in one of which it the first qubit is in $|0\rangle$, and in the other the first qubit is in $|1\rangle$. Thus any measurement of the first qubit will return either $|0\rangle$ or $|1\rangle$, at random and with equal probability. This is not particularly odd; what is odd is the effect that measuring the first qubit has on the *second* qubit. Our entangled state is a superposition of two states, in both of which the two individual qubits have the same state; thus if we measure qubit one and find it is in $|0\rangle$ we know immediately that qubit two must also be in state $|0\rangle$! Similarly, if we find qubit one in $|1\rangle$ then qubit two will also be in $|1\rangle$. The behavior of the two qubits is completely intertwined, or *entangled*.

The state discussed above is only one example of an infinite number of possible entangled states. Particularly important among these are the four Bell states, which are *maximally entangled* states, meaning that their behavior is as unlike a direct product state as possible. These states are defined by

$$\phi^{\pm} = (|00\rangle \pm |11\rangle)/\sqrt{2} \qquad \psi^{\pm} = (|01\rangle \pm |10\rangle)/\sqrt{2} \tag{5.16}$$

and so the state discussed previously is a ϕ^+ Bell state. Note that the four Bell states form an orthonormal basis for the maximally entangled states.

Chapter 6

A peek into the future

We finally have all the basic tools we will need to describe systems of one and two qubits. In this last chapter we look at some of the consequences of the peculiar properties of quantum systems, and see how they might prove useful.

6.1 Measuring a single qubit

A key result in quantum information theory is that it is impossible to accurately characterize a single qubit. In other words, there is no experiment, or sequence of experiments, which allow us to find out the state of a single quantum bit.

The reason for this problem is two-fold. Firstly, we have to make some sort of decision about the basis we will use for our measurement. For example, when measuring a single qubit the most popular choice is to make a measurement in the computational basis. This is equivalent to asking the qubit whether it is in $|0\rangle$ or $|1\rangle$. If the qubit is indeed in $|0\rangle$ or $|1\rangle$ the measurement process is simple, and we will get the obvious answer. If, however, the qubit is in a superposition, such as $\alpha|0\rangle + \beta|1\rangle$, then the situation is more difficult. Characterizing the state now means determining the values of α and β, but the measurement can only return the answer 0 or 1, and for a superposition state one of these two answers will be returned at random, with probabilities $|\alpha|^2$ and $|\beta|^2$ respectively.

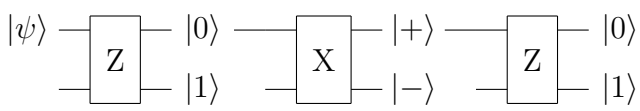
Clearly we cannot characterize a superposition state in a single measurement, but why not just make repeated measurements, and so gain statistical information about α and β? The problem is that the first measurement does not leave the state unaffected: if the first measurement returned 0 then the state is *changed* to $|0\rangle$, and if the first measurement returned 1 the state is changed to $|1\rangle$. Any subsequent measurement of the state will therefore return the same answer as before, and no more information can be obtained. This point was addressed briefly in section 2.7.

This measurement behavior is so strange that it is good to go back to a few simple examples. The classic example of a quantum measurement is a Stern–Gerlach apparatus, which measures the projection of a spin onto some axis. This is, of course, entirely equivalent to measuring a qubit in some basis: a Stern–Gerlach apparatus aligned along the z-axis is equivalent to a measurement in the computational basis, while one aligned along the x-axis is equivalent to a measurement in the $|\pm\rangle$ basis. These measurements always produce one of two results: the spin is either parallel or antiparallel to the measurement axis. If the spin is neither parallel or antiparallel then one of the two permitted results is returned at random, with probabilities depending on the projection of the

spin onto the axis.

One might ask what is so special about the computational basis, and the short answer is that measuring in any other basis is ultimately equivalent. The key thing about the computational basis is simply that it is an orthonormal basis, as the only quantum measurements that work reliably are those where the possible answers correspond to orthonormal states[1]. However any orthonormal basis is (in principle) as good as any other, and we can choose some other basis without altering the underlying ideas.

Returning to our Stern–Gerlach apparatus, the situation becomes much more interesting when considering a cascaded network of measurements such as that shown below.



(6.1)

Suppose we measure the z-component of a spin, and find it to be $|0\rangle$, and the measure the x-component of a spin and find it to be $|+\rangle$. Finally we measure the z-component again, and get the remarkable effect that the result is either $|0\rangle$ or $|1\rangle$, each with 50% probability. This makes easy to explain using quantum mechanics (the effect of the x-measurement is to *change* the state to either $|+\rangle$ or $|-\rangle$), but hard to make sense of in any other way.

The same result can be experimentally demonstrated even more simply using polarized light. In this case horizontally and vertically polarized light corresponds to the computational basis states, $|0\rangle$ and $|1\rangle$, while light polarized at $\pm 45^\circ$ corresponds to the $|\pm\rangle$ states. Suppose a photon is passed through a pair of crossed polarizers: the only photons that can get through are those which were in $|0\rangle$ at the first polarizer and in $|1\rangle$ at the second one, and so no light is transmitted. But if a polarizer at $\pm 45^\circ$ is placed *between* the two crossed polarizers, then the intervening measurement changes the states of the photons, allowing some of them to pass!

If we do know something about the state of the qubit, the situation is radically changed. For example, if we know the qubit is in either $|0\rangle$ or $|1\rangle$ then a measurement in the computational basis will immediately tell us which one it is. Next consider a qubit which is known to be in one of the two superposition states $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. In each case measurement in the computational basis will return either $|0\rangle$ or $|1\rangle$ with 50% probability, and so this measurement basis tells us nothing useful about the state. The two states are, however, orthonormal, and so should be completely distinguishable! This can be achieved by measuring in a suitable basis usually called the X-basis, which always returns either $|+\rangle$ or $|-\rangle$.

Now suppose we have a qubit in either $|0\rangle$ or $|1\rangle$ and choose to measure its state in the X-basis. In this case the measurement will return either $|+\rangle$ or $|-\rangle$, each with 50% probability, and we learn nothing at all about the state. While we can optimize our measurement process for any particular pair of states we cannot simultaneously optimize it for all possible states. Clearly if the state of the qubit is unknown it is impossible to optimize the measurement.

More insight into the measurement process can be obtained by using a Bloch sphere picture. The two states $|0\rangle$ and $|1\rangle$ lie at the north and south poles of the Bloch sphere, and a measurement in the computational basis applied to a general state is essentially a method of estimating its

[1]There is a substantial subfield of quantum information theory which considers how to make measurements that seek to distinguish states that are not orthonormal. The detailed answers are quite complicated, but the key result is that such measurements cannot be made to work with perfect reliability, and they do not fundamentally change the discussion above.

projection onto the z-axis. The two basis states $|0\rangle$ and $|1\rangle$ lie at opposite ends of the z-axis, and are easily distinguished. By contrast the states $|\pm\rangle$ lie at opposite ends of the x-axis, and the projection onto the z-axis is zero for both states, showing that they cannot be distinguished. These states are best distinguished by their projections onto the x-axis, that is by measurements in the X-basis, but this is completely useless for the states $|0\rangle$ and $|1\rangle$, whose projection onto the x-axis is zero. A measurement will only be perfect if the measurement axis is parallel to the state, and will be completely useless if the measurement axis is perpendicular to the state. For a completely unknown state there is no sensible way to choose the axis, and any measurement is as good as any other.

6.2 Ensembles

So far we have assumed that we have only one copy of our unknown quantum state. Suppose, however, that we have a large number of identical copies of the state, a situation usually called an ensemble. If the state is $|\psi\rangle$ then the ensemble can be written as

$$(|\psi\rangle)^n = \bigotimes^n |\psi\rangle = \overbrace{|\psi\rangle \otimes |\psi\rangle \otimes \cdots \otimes |\psi\rangle}^{n \text{ terms}} \tag{6.2}$$

which is a direct product of copies of the unknown state. The significance of the direct product form is that the individual copies of the state are *independent*, in the sense that manipulating one qubit does not affect any others. By performing several different measurements on many copies of the state we can get a very good idea of what the state is.

It might seem that this offers a solution to the problem of characterizing an unknown state: all that is necessary is to make an ensemble of copies of the state and then measure these! As we shall see, however, this process is impossible.

6.3 The no-cloning theorem

The no-cloning theorem is arguably one of the most important results in the whole of quantum information theory, but it is also one of the simplest. It is possible to copy classical information without limits[2], but it is almost trivial to prove that an unknown quantum state cannot be copied (cloned). A brief proof is sketched below; for more details see [Stolze 2004] or [Nielsen 2000].

The proof proceeds by contradiction. Suppose a quantum cloning device, capable of accurately copying a completely unknown state, did in fact exist. Clearly such a device must be capable of copying the two basis states $|0\rangle$ and $|1\rangle$. Copying the two basis states is easy, and can in fact be achieved by a controlled-NOT gate, which performs

$$|0\rangle|0\rangle \longrightarrow |0\rangle|0\rangle \qquad |1\rangle|0\rangle \longrightarrow |1\rangle|1\rangle \tag{6.3}$$

as desired. However this approach *cannot* be used to clone a general state, such as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle. \tag{6.4}$$

[2]This fact has caused considerable annoyance to companies trying to sell recorded music!

If a controlled-NOT gate is used to "copy" this state, the result will be

$$\alpha|0\rangle|0\rangle + \beta|1\rangle|1\rangle, \tag{6.5}$$

a result which follows immediately from the linearity of the controlled-NOT gate. This state should be compared with the desired state, which has the form

$$|\psi\rangle \otimes |\psi\rangle = \alpha^2|0\rangle|0\rangle + \alpha\beta|0\rangle|1\rangle + \beta\alpha|1\rangle|0\rangle + \beta^2|1\rangle|1\rangle. \tag{6.6}$$

Clearly these states are only the same in the extreme limits of $\alpha = 1$ (implying $\beta = 0$) or $\beta = 1$, that is when the state being copied is a basis state. This result suggests that quantum cloning is indeed impossible, but is not completely convincing as one particular operation for cloning the basis states has been assumed. However a little thought shows that any other putative cloning method is fundamentally equivalent to a controlled-NOT gate, and the argument is basically sound. Just as was the case for characterizing a quantum state, it is possible to optimize the cloning process to work for a particular pair of states, but no solution exists for an unknown state.

Accepting that an unknown quantum state cannot be copied accurately, one might still ask whether the output of a quantum cloner could in any way assist an attempt to characterize an unknown state. In fact the form of the state in equation 6.5 immediately rules this out, as this is an entangled state, in which the properties of the two qubits are completely correlated. Once the first qubit has been measured we know that the second qubit will have the same state; actually measuring this state tells us nothing new about the system.

There is in fact an important link between the problem of measurement and the no-cloning theorem. The fact that an unknown quantum state cannot be copied means that the measurement problem cannot be overcome by copying. Similarly, the inability to accurately characterize an unknown state rules out an obvious cloning strategy: measuring the state precisely and crafting identical copies. These two interlinked phenomena lie at the heart of one of the most important applications of quantum information processing, called quantum cryptography.

6.4 Fidelity

The discussions above prove that it is impossible to perform certain operations on quantum bits without the possibility of error. The obvious question is then how accurately these operations can be performed. Critical to these questions is the concept of *fidelity*, which measures how close two states (or, by extension, two operations) are to one another.

For assessing state fidelity it seems obvious that the measure should be built around the inner product. Since the two basis states are orthonormal, we know that

$$\langle 0|0\rangle = \langle 1|1\rangle = 1 \qquad \langle 0|1\rangle = \langle 1|0\rangle = 0 \tag{6.7}$$

which makes sense. However the inner product of two general states will be a complex number, while fidelity should be a real number between 0 and 1. A better definition of the fidelity of one ket $|\phi\rangle$ with respect to another ket $|\psi\rangle$ is[3]

$$F(|\phi\rangle, |\psi\rangle) = |\langle\phi|\psi\rangle|^2 = \langle\psi|\phi\rangle\langle\phi|\psi\rangle \tag{6.8}$$

[3]Some authors, notably [Nielsen 2000], use the square root of this definition. Clearly these two definitions are very closely related.

This definition can be extended to measure the fidelity between a pure state $|\psi\rangle$ and a mixed state[4] described by a density matrix ρ

$$F(\rho, |\psi\rangle) = \langle\psi|\rho|\psi\rangle \tag{6.9}$$

which clearly reverts to the original form for two pure states, when $\rho = |\phi\rangle\langle\phi|$. To take a simple example, consider the fidelity between a general state and itself:

$$\begin{aligned} F &= \begin{pmatrix}\alpha^* & \beta^*\end{pmatrix}\begin{pmatrix}\alpha\alpha^* & \alpha\beta^* \\ \beta\alpha^* & \beta\beta^*\end{pmatrix}\begin{pmatrix}\alpha \\ \beta\end{pmatrix} && (6.10)\\ &= (|\alpha|^2 + |\beta|^2)^2 && (6.11)\\ &= 1. && (6.12)\end{aligned}$$

A more interesting case is the fidelity between an arbitrary pure state and the same state after a measurement in the computational basis:

$$\begin{aligned} F &= \begin{pmatrix}\alpha^* & \beta^*\end{pmatrix}\begin{pmatrix}\alpha\alpha^* & 0 \\ 0 & \beta\beta^*\end{pmatrix}\begin{pmatrix}\alpha \\ \beta\end{pmatrix} && (6.13)\\ &= |\alpha|^4 + |\beta|^4 && (6.14)\\ &= 1 - 2|\alpha|^2|\beta|^2. && (6.15)\end{aligned}$$

Clearly the process of measurement damages a state unless the state is a basis state of the measurement (so that either α or β is equal to zero). The worst case occurs for states like $|\pm\rangle$ for which $|\alpha| = |\beta| = 1/\sqrt{2}$, resulting in a fidelity of $F = 1/2$. For the general case it can be shown that the average fidelity of a state after a measurement is $2/3$.

This result can be interpreted in two different ways. Firstly, as noted above, measuring an unknown state will damage it, and so it is always possible to check whether someone has been looking at your "secret" state.[5] Secondly, the state after the measurement also describes the state of knowledge of the person who performed the measurement, and this shows that the knowledge of the state can never be perfect unless the correct measurement basis was used.

6.5 Local operations and classical communication

So far we have considered two qubit states where both qubits are accessible to us. The situation becomes much more interesting when different qubits are controlled by different people.

Consider two people, traditionally called Alice and Bob, each of whom have one qubit[6] of a two qubit system. We assume that they can both manipulate their *own* qubit in any way they desire: they can apply single qubit logic gates, make measurement, etc., but they have no direct access to the other person's qubit. Speaking technically, we say that Alice and Bob have access to the complete set of *local* operations. We also assume that Alice and Bob can communicate by sending classical messages, reporting the results of measurements on their own qubits, or asking

[4]It can also be extended to measure the fidelity between two mixed states, but the definition becomes complicated; see [Nielsen 2000] for details.

[5]This fact underlies the idea of *quantum money*, which was invented by Steven Wiesner long before quantum information theory was thought of, and ultimately underlies *quantum cryptography*.

[6]That is, they have possession and control of the physical system used to implement the qubit.

that certain gates be applied to the other person's qubit. This set of abilities is described as *local operations and classical communication*, usually abbreviated to *LOCC*.[7]

If the two qubit system is in a separable state then nothing mysterious occurs. Recall that it is the nature of a separable state that the two qubits have individual properties, and it makes sense to treat them as individual objects. If the two qubits are entangled, however, then the situation is entirely different! It is no longer really possible to talk about the two qubits as separate objects. As a simple example, suppose Alice and Bob share a pair of qubits in the entangled state

$$\phi^+ = (|00\rangle + |11\rangle)/\sqrt{2} \tag{6.16}$$

and that Alice applies a NOT gate to her qubit (assumed to be the qubit listed first in our notation). The result is

$$(|10\rangle + |01\rangle)/\sqrt{2} = \psi^+. \tag{6.17}$$

In a similar way, Alice can convert the state into any one of the four Bell states, and Bob can do the same thing. It is no longer possible to divide up the state into portions controlled by Alice and portions controlled by Bob: they both have equal control over the entire state. This behavior lies at the heart of quantum communication protocols such as *quantum dense coding* which will be explored later; less positively it also makes certain elementary cryptographic operations impossible in the quantum world, most notably the impossibility of *quantum bit commitment*.

Given this key distinction between separable and entangled states, it is reasonable to ask whether Alice (with or without help from Bob) can turn an initially separable state into an entangled state using only local operations and classical communications. It is a key result in quantum information theory that this is impossible, and more generally the amount of entanglement in a quantum system cannot be increased by LOCC. If Alice and Bob wish to use an entangled state they must either create one by applying two qubit gates (which requires the two qubits to be brought into direct contact), or use a state prepared by some third party. For simplicity we can often assume that Alice prepares the entangled state and gives one qubit to Bob. It is a curious fact about many quantum communication protocols that it does not matter where the entangled state comes from: if a malicious person seeks to cheat by providing the wrong state then this fact can be easily detected.

6.6 The EPR problem

One of the most vivid illustrations of the sheer weirdness of entangled states is provided by the Einstein–Podolsky–Rosen (EPR) thought experiment, which was put forward in 1935 as an argument that quantum mechanics could not possibly be correct: the predictions quantum mechanics made about the results of this experiment were so ludicrous[8] that they could not possibly be true. While thought experiments can be interesting, real experiments are more convincing, and in 1951 David Bohm tightened up the EPR argument and described a thought experiment which could

[7] By local, here, we mean local in the obvious every day sense, rather than in the relativistic sense; however the extension to include classical communications, which we assume to be limited by the speed of light, means that LOCC is equivalent to relativistically local.

[8] In particular when interpreted naively they appeared to be inconsistent with relativity; whether they actually are inconsistent is a much more interesting question.

potentially be done. This experiment has subsequently been performed, and quantum mechanics appears to have won.[9]

The modern form of the EPR argument assumes that Alice and Bob each have one qubit from an entangled two qubit state which is in the Bell state $|\psi^-\rangle$. This state is often called the *singlet* state, because if the two qubits are spins then this is the state with total spin zero; the other three Bell states have total spin one and correspond to triplet states. If Alice and Bob measure their respective qubits in the computational basis then they will each get $|0\rangle$ and $|1\rangle$ at random; if, however, they compare their results they will notice that whenever Alice got $|0\rangle$ then Bob got $|1\rangle$, and *vice versa*. We have seen this behavior before, but the choice of the singlet state (rather than one of the triplet Bell states) makes the situation even more interesting, as Alice and Bob will get corresponding results whatever basis they choose for their measurements.[10] Suppose, for example, they choose to measure in the $|\pm\rangle$ basis: in this case they will both observe $|+\rangle$ and $|-\rangle$ at random, but whenever Alice observes $|+\rangle$ then Bob will observe $|-\rangle$, and so on.

Now let us look at the situation purely from Alice's point of view: suppose she measures her qubit in the $|\pm\rangle$ basis and observes $|+\rangle$. From this fact she immediately knows that if Bob measures his qubit in the $|\pm\rangle$ basis then he will observe $|-\rangle$. It seems as if Alice's local actions can affect the state of Bob's qubit, even though the two qubits are not interacting with one another. Furthermore, this effect appears to occur *instantaneously*, rather than propagating at or below the speed of light. Einstein referred to this as "spooky action at a distance" (spukhafte Fernwirkung).

6.7 The Bell inequalities

In an attempt to exorcize the spooks, various physicists tried to explain the EPR effects by *hidden variable theories*. The key idea behind these theories is that, in addition to the obvious properties which we can measure directly, quantum particles can possess additional properties which are hidden from direct observation but which determine the behavior of the particles in situations such as the EPR experiment. This whole enterprise was crucially undermined in 1964 by John Bell, who showed that any local hidden variable theory must necessarily make predictions which are inconsistent with those made by traditional quantum mechanics. In essence the procedure is to compare the results of a large number of measurements where Alice and Bob choose their measurements at random; the results of these experiments can be boiled down into a single number, which for any local hidden variable model must always be less than on equal to 2. By contrast quantum mechanics predicts that this number can rise as high as $2\sqrt{2}$.

The Bell inequalities will be studied in considerable detail next year;[11] for the moment we simply note three consequences. Firstly, the arguments used by Bell are of a simple and fundamental kind,

[9] Purists would point out that experimental imperfections in the current systems mean that the results can in principle be explained away, and it is true that an absolutely convincing *loophole free* test has yet to be carried out. Indeed coming up with new loopholes forms a minor industry among quantum information theorists. However as experiments improve these theories are becoming increasingly bizarre and contrived, and they are usually called *conspiracy theories* in recognition of this fact. A more hard line approach is taken by a small band of philosophers, who claim that the reasoning underlying Bell's argument is fundamentally flawed. For a friendly introduction to these ideas see [Mermin 1990].

[10] This fact is most simply proved by showing that $|\psi^-\rangle$ is left unchanged by any bilateral unitary transformation, that is any case where the *same* unitary transformation is applied to both qubits.

[11] For a friendly introduction see [Mermin 1990]; for a bit more detail try [Bell 2004].

and this is even more true of later refinements: a system that breaks a Bell inequality *cannot* be described in any straightforward way as a system made up of two independent subsystems. Secondly, many experiments appear to break the Bell inequalities, suggesting that the real world cannot be described in any straightforward way. Thirdly, the impossibility of "mocking up" behavior which breaks the Bell inequalities means that it is essentially impossible for a dishonest third party to deceive Alice and Bob by providing fake entangled states.

6.8 Faster than light?

We have already noted that one of the most intriguing properties of the correlations observed in singlet states is that the effect appears to be instantaneous, and this immediately suggests the possibility of using entanglement to build a faster than light communicator. Sadly,[12] however, this turns out to be impossible. It is straightforward to show that a faster than light communicator could be built if it were possible to completely characterize an unknown quantum state. But, as we have already seen, this cannot be done. Combing our previous results, it seems that the measurement problem, the no-cloning theorem, and the impossibility of faster than light communication are all inextricably interlinked. It can also be shown that these limits are linked to limits on the power of computers: a world which permits quantum cloning (or, equivalently, accurate quantum measurement) would also permit any mathematical problem to be solved arbitrarily fast. Faster than light communication is also closely linked to the possibility of time travel. Rather than being a limitation, the no-cloning theorem seems to be essential to prevent a quantum universe from descending into insanity!

[12] Or happily, depending on your point of view.

Appendix A

Single qubit gates

Here I list some of the most important single qubit gates, and some relationships between them. The notation is partly based on that of [Nielsen 2000], which also covers multi-qubit gates.

$$\mathbb{1} = \sigma_0 = \mathrm{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{A.1}$$

$$i180_x = \sigma_x = \mathrm{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{A.2}$$

$$i180_y = \sigma_y = \mathrm{Y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{A.3}$$

$$i180_z = \sigma_z = \mathrm{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{A.4}$$

$$90_x = \sqrt{\mathrm{X}/i} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \tag{A.5}$$

$$90_y = \sqrt{\mathrm{Y}/i} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \tag{A.6}$$

$$\mathrm{S} = \sqrt{\mathrm{Z}} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \tag{A.7}$$

$$\mathrm{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{A.8}$$

Note that $\mathrm{X}^2 = \mathrm{Y}^2 = \mathrm{Z}^2 = \mathrm{H}^2 = \mathrm{I}$ and that $\mathrm{S}^2 = \mathrm{Z}$. The Hadamard gate H can be implemented in may different ways, such as $\mathrm{H} = \mathrm{S}\, 90_x\, \mathrm{S} = 90_y\, \mathrm{Z}$. Other identities include $\mathrm{HZH} = \mathrm{X}$ and $\mathrm{HXH} = \mathrm{Z}$.

Appendix B

Quantum optics

In the main text I have adopted a traditional *semi-classical* approach to describing transitions between energy levels of a quantum system, in which light is treated as a classical oscillating electric or magnetic field. This approach works well, but clearly it cannot be quite right, as light is itself a quantum system. Here I briefly discuss how a full quantum mechanical treatment of light-matter interactions can be developed, and why it is rarely necessary.

The key step is to see that, from a quantum mechanical point of a view, a light field is nothing more than a thinly disguised harmonic oscillator, with the number of photons in the light field corresponding to the quantum number in the harmonic oscillator. Note that the energy of a light field containing n photons is normally written as $E_n = n\,h\nu$ (the photons do not interact with one another and so the energy of n photons is simply n times the energy of a single photon), while the energy of the nth level of a harmonic oscillator is $E_n = (n + \frac{1}{2})\,\hbar\omega$. Clearly these formulae are essentially equivalent.

The standard way to treat a quantum harmonic oscillator is in terms of the raising operator $A^\dagger$ and the lowering operator A, which perform

$$A^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle \qquad A|n\rangle = \sqrt{n}\,|n-1\rangle \tag{B.1}$$

increasing or decreasing the quantum number by 1. Entirely equivalent operators can be used to describe a light field: $a^\dagger$, the creation operator, creates an additional photon in the light field (increases n by 1), while a, the annihilation operator, destroys a photon (reduces n by 1).

A similar approach can be used to define two related operators for a qubit, σ and $\sigma^\dagger$, defined by

$$\sigma = \tfrac{1}{2}\,(\sigma_x + i\sigma_y) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{B.2}$$

and

$$\sigma^\dagger = \tfrac{1}{2}\,(\sigma_x - i\sigma_y) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{B.3}$$

which convert $|1\rangle$ to $|0\rangle$ and *vice versa*. Note that $\sigma^\dagger$ is the Hermitian conjugate of σ as expected; the same thing is, of course, true for a and $a^\dagger$, although in this case a matrix representation must be infinite dimensional! If, however, we assume that the system can only contain zero or one photons, then a and $a^\dagger$ take the same forms as σ and $\sigma^\dagger$.

We are now in a position to write down the interaction between the light field and the atom. Two basic events can occur: the absorption of a photon, with consequent excitation of the atom, and emission of a photon, with de-excitation, and both of these processes happen at the same rate, which we can call $\Omega/2$. Thus the Hamiltonian is

$$\mathcal{H} = \tfrac{1}{2}\hbar\Omega\left(a\sigma^\dagger + a^\dagger\sigma\right) \tag{B.4}$$

where we are assuming that the light is exactly resonant with the field, and we have chosen one particular phase for the light. This can be written out explicitly by noting that the light field can be treated as another qubit (remember that we are assuming the light field contains zero or one photons) and using the usual direct product approach gives

$$\mathcal{H} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \hbar\Omega/2 & 0 \\ 0 & \hbar\Omega/2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{B.5}$$

showing that the Hamiltonian is block-diagonal, dividing into a central two by two block and two outer elements; the corresponding propagator

$$U = \exp(-i\mathcal{H}t/\hbar) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\Omega t/2) & -i\sin(\Omega t/2) & 0 \\ 0 & -i\sin(\Omega t/2) & \cos(\Omega t/2) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{B.6}$$

has the same structure. Clearly the state $|00\rangle$ does not evolve (the ground state of the atom cannot emit a photon, and there is no photon present for it to absorb), and a similar (but more subtle) argument applies to $|11\rangle$. The interesting behavior occurs in the central block, which is identical to the Hamiltonian describing the interaction of a qubit with a classical field. Thus the state of the atom undergoes Rabi oscillations, while the state of the light field also oscillates in the opposite direction.

Another way of looking at this situation is to consider the propagator in the case that $\Omega t = \pi$ (a 180° pulse), for which

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{B.7}$$

Neglecting a few phases this is essentially a SWAP quantum logic gate: the interaction between the light and the atom acts to swap quantum information between the two qubits. More interesting behavior occurs when $\Omega t = \pi/2$, as this creates an *entangled* state of the atom and the light field.

Quantum optics is an interesting and important topic in its own right, but the above hints at a potential problem with implementing quantum computers with atomic states in the way described previously: it appears that applying single qubit gates to the atom will inevitably entangle it with the light field. Any process that effectively measures the light field, causing it to decohere, will also cause the atomic state to decohere at the same time.

The solution to this problem is that a *classical light field* does not correspond to a state with a well defined number of photons (a *Fock state*), but to a much more interesting state known as a *coherent state*. Coherent states take the form

$$|\alpha\rangle = \exp(-|\alpha|^2/2) \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{B.8}$$

where α is some complex number, and have the useful property that they are eigenstates (technically, right-eigenstates) of the annihilation operator:

$$a|\alpha\rangle = \alpha|\alpha\rangle \tag{B.9}$$

(for the details see [Barnett 1997]). Thus it is possible to remove a photon for a coherent state without changing it in any way! If a coherent state is used to excite an atom then the light and the atom do not become entangled.

Bibliography

[Atkins 1997] *Molecular Quantum Mechanics*, P. W. Atkins and R. S. Friedman (1997).

[Barnett 1997] *Methods in Theoretical Quantum Optics*, S. M. Barnett and P. M. Radmore (1997).

[Bell 2004] *Speakable and Unspeakable in Quantum Mechanics*, J. S. Bell, 2nd edition (2004).

[Budker 2004] *Atomic Physics*, D. Budker, D. F. Kimball and D. P. DeMille (2004)

[Dirac] *The Principles of Quantum Mechanics*, P. A. M. Dirac, 4th edition (1958).

[Feynman 1996] *Feynman Lectures on Computation*, R. Feynman, edited by A. J. G. Hey and R. W. Allen (1996).

[Freeman 1998] *Spin Choreography*, R. Freeman (1998).

[Gasiorowicz 2003] *Quantum Physics*, S. Gasiorowicz (2003).

[Goldman 1988] *Quantum Description of High-Resolution NMR in Liquids*, M. Goldman (1988).

[Halmos 1974] *Finite-Dimensional Vector Spaces*, P. R. Halmos (1974).

[Mermin 1990] *Boojums All The Way Through*, N. D. Mermin (1990).

[Nielsen 2000] *Quantum Computation and Quantum Information*, M. A. Nielsen and I. L. Chuang (2000).

[Shankar 1994] *The Principles of Quantum Mechanics*, R. Shankar (1994).

[Stolze 2004] *Quantum Computing*, J. Stolze and D. Suter (2004).