

# 1 Summary of things you should already know

*“I think I can safely say that nobody understands quantum mechanics”* Richard Feynman

## 1.1 Operators and Observables

It is a premise of QM that any measurable quantity is associated with a Hermitian operator. So far we are used to analytic expressions for wavefunctions and operators. But if a ket is full description of the state of a system, it must also contain some implicit information. The abstract bra-ket notation includes this.

Consider the electric charge. Obviously this is measurable, so it should be associated with an operator  $\hat{Q}$ , such that e.g.

$$\hat{Q}|\Phi\rangle = -e|\Phi\rangle$$

where  $\Phi$  is the wavefunction of an electron.  $-e$  meets all the criteria for a quantum number, and the above equation is obviously a true representation of reality, but it cannot be proved algebraically. Thus the meaning of the ket  $|\Phi\rangle$  is broader than a simple spatial function, and operators can also be non-algebraic. This is especially important in particle physics where all manner of quantum numbers appear (isospin, strangeness, baryon number etc. etc.)

## 1.2 Hamiltonians and eigenstates

Schroedinger's equation  $\hat{H}\phi = i\hbar\partial\phi/\partial t$  shows us that the Hamiltonian (energy operator) is related to the change in wavefunction in time. A system prepared in an eigenstate of the Hamiltonian has time invariant probability density. A system prepared in an eigenstate of a non-commuting operator has a probability density which varies in time. It is this time independence (conservation law) which makes eigenstates of the energy operator so useful.

## 1.3 Matrix and operator mechanics

There are two equivalent mathematical ways of calculating physical properties, Schroedinger's wave mechanics and Heisenberg's matrix mechanics. In each systems are represented in terms of eigenstates and measurables as eigenvalues. In matrix mechanics the operator is represented by a Hermitian matrix of elements  $\langle m|\hat{Q}|n\rangle$  which depends on the choice of basis set  $|m\rangle$ . Any state can be represented by a normalised *vector*, which also depends on the basis set. The eigenvalues and eigenvectors of the matrix, however, do not depend on the choice of basis - the eigenvectors are, in fact, the eigenbasis of the operator.

For a set of basis 'vectors' of size  $N$ , there are  $N \times N$  possible matrix elements. For an infinite basis set, of course, there are infinite elements.

## 1.4 Wavefunction collapse and measurement

When we measure some property of a system, the act of making the measurement collapses the system into an eigenstate of the appropriate operator. The system then evolves according to Schroedinger's equation. All memory of the previous state of the system is lost in this collapse, except in the special case when the state is degenerate, as we'll see later.

## 1.5 Formal definition of a complete, orthonormal basis set

Consider a basis set  $|i_n\rangle$ . It is *orthonormal* if  $\langle i_n|i_m\rangle = \delta_{mn}$ . It is *complete* if any wavefunction can be written as  $|\phi\rangle = \sum_n c_n|i_n\rangle$  and the  $c_n$  are uniquely defined. If the wavefunction cannot be so written, the basis set is *incomplete*, if there exists more than one possible set of  $c_n$ , the basis set is *overcomplete*. Choosing a basis set in a Hilbert space (see 1.7) is analogous to choosing a set of coordinates in a vector space. Note that completeness and orthonormality are well defined concepts for both vector spaces and function spaces.

## 1.6 Example of matrix representation method and choice of basis

In practical quantum problems, we almost always describe the state of the system in terms of some basis set. Consider a simple spin 1/2 system, choosing as basis states  $S_z = \pm\frac{1}{2}$ . Consider this system in a magnetic field pointing in the  $x$  direction, the operator corresponding to this is  $\mu B \hat{S}_x$ . We wish to find the eigenstates and eigenenergies.

Evaluating the required matrix elements such as  $\langle S_z = \frac{1}{2}|\mu B \hat{S}_x|S_z = \frac{1}{2}\rangle$  (see QP3) gives a matrix:

$$\begin{pmatrix} 0 & \mu B/2 \\ \mu B/2 & 0 \end{pmatrix}$$

The normalised eigenvectors of this are  $(\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$  and  $(\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$  with eigenvalues  $(\mu B/2)$  and  $(-\mu B/2)$ . Of course these represent the eigenstates  $|S_x = \pm\frac{1}{2}\rangle$  in the basis of  $|S_z = \pm\frac{1}{2}\rangle$ :

$$|S_x = \pm\frac{1}{2}\rangle = \left[ |S_z = \frac{1}{2}\rangle \pm |S_z = -\frac{1}{2}\rangle \right] / \sqrt{2}$$

Had we chosen  $|S_y = \pm\frac{1}{2}\rangle$  as our basis set, then the matrix would have been:

$$\begin{pmatrix} 0 & -i\mu B/2 \\ i\mu B/2 & 0 \end{pmatrix}$$

Once again, the eigenvalues of this matrix are  $(\mu B/2)$  and  $(-\mu B/2)$ , as they must be since these are the measurable quantities. Coincidentally, the eigenvectors in this basis set are also  $(\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$  and  $(\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$ .

Had we chosen  $|S_x = \pm\frac{1}{2}\rangle$  as our basis set in the first place, the problem would have been much simplified. The matrix would then be:

$$\begin{pmatrix} \mu B/2 & 0 \\ 0 & -\mu B/2 \end{pmatrix}$$

Once again, the eigenvalues of this matrix are  $(\mu B/2)$  and  $(-\mu B/2)$ , and now the eigenvectors are (1,0) and (0,1): i.e. the eigenstates are simply the basis states.

## 1.7 Dirac Notation - Analogies with vectors and matrices

You probably remember Dirac notation as a shorthand for integrals, for example the overlap between two wavefunctions can be written as:

$$\langle \chi|\phi\rangle \quad \text{instead of} \quad \int \int \int \chi^*(\mathbf{r})\phi(\mathbf{r})d^3\mathbf{r}.$$

(Where  $d^3\mathbf{r}$  is the scalar volume element, sometimes called  $r^2 \sin \theta d\theta d\phi dr$ ,  $dx dy dz$ ,  $dV$  or  $d\tau$ )

But also if we have a complete set of orthonormal basis states  $i$ , the overlap is also the sum of the overlaps between each  $i$  and  $\chi$  and  $\phi$

$$\langle \chi | \phi \rangle = \sum_i \langle \chi | i \rangle \langle i | \phi \rangle$$

**Warning:** A summation convention is also sometimes used, such that when a state symbol appears twice, first as a ket, then as a bra, it is assumed to be summed over a complete set of orthonormal basis states. The expression above is then further abbreviated to  $\langle \chi | i \rangle \langle i | \phi \rangle$ . This convention can be confusing and will not be used in these notes.

Compare this with the vector dot product formula

$$\mathbf{b} \cdot \mathbf{a} = b_x a_x + b_y a_y + b_z a_z = \sum_i (\mathbf{b} \cdot \mathbf{e}_i) (\mathbf{e}_i \cdot \mathbf{a})$$

where  $\mathbf{e}_i$  are the unit vectors in x, y and z directions. Just as any vector can be expressed as a linear combination of  $\mathbf{e}_i$ , so any quantum state can be expressed as a linear combination of basis states  $i$ . There are certain conditions on the basis states, e.g. they must be ‘orthonormal’  $\langle j | i \rangle = \delta_{ij}$  just as  $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ . Just as the three Cartesian vectors span a three dimensional space, so the many basis states span a many-dimensional space. In some cases (e.g. Fourier expansions, hydrogen wavefunctions) there are an infinite number of basis states which are therefore related to spanning an infinite-dimensional space. Mathematicians call these ‘Hilbert spaces’. Any state  $\phi$  can thus be viewed as a vector in a multi-dimensional space, where each dimension corresponds to one of the basis functions. It is thus common to use the words eigenstate and eigenvector interchangeably to refer to  $|\phi\rangle$  Even before the discovery of quantum mechanics, mathematicians had solved many of the problems in this area.

In Dirac notation we have two quantities, the bra and the ket, whereas in vector algebra we have only one, this is because there is not an exact analogy to commutation for Dirac brackets:  $\langle \chi | \phi \rangle = \langle \phi | \chi \rangle^*$  includes taking a complex conjugate. Consider manipulating the bras and kets. We can write a vector in terms of its components thus

$$\mathbf{A} = \sum_i \mathbf{e}_i (\mathbf{e}_i \cdot \mathbf{A})$$

where  $(\mathbf{e}_i \cdot \mathbf{A})$  is the amount of  $\mathbf{A}$  along the  $\mathbf{e}_i$  axis; the components. The quantities on either side of the equation are not numbers but *vectors*. We can generate a whole algebra based on vectors.

Likewise we can write a state thus:  $|\phi\rangle = \sum_i |i\rangle \langle i | \phi \rangle$

where  $\langle i | \phi \rangle$  is the amount of  $\phi$  along the  $i$  basis state; the components or *expansion coefficients*. The quantities on each side of this equation are not numbers but *functions*.  $\phi$  is a normalised wavefunction iff  $\sum_i |\langle i | \phi \rangle|^2 = 1$ . We can then generate a whole algebra based on bras and kets.

For any different complete sets of basis states  $i$  and  $j$ , we can write:  $|\phi\rangle = \sum_j |j\rangle \langle j | \phi \rangle$ , and  $\langle \phi | = \sum_i \langle i | \langle i | \phi \rangle$ . Expansions in  $i$  and  $j$  are called different *representations* of  $\phi$ . This is very similar to using different coordinate systems: the bases  $i$  and  $j$  are analogous to two sets of axes rotated with respect to one another. We might choose complete set of wavefunctions including  $\phi$  as a representation, just as we sometimes choose axes such that some special vector points along the  $z$ -axis.

Going even further, the expansion in a basis can be done for any  $|\phi\rangle$ , so perhaps we can dispense with  $|\phi\rangle$  and write:

$$1 = \sum_i |i\rangle\langle i|, \text{ the unit operator}$$

All this means is that in any equation you can always proceed by breaking the states down into a complete, orthonormal set of basis functions. This may be useful when dealing with a Hamiltonian for which the eigenstates  $i$  with eigenenergies  $E_i$  are already known. A general mixed state  $|\phi\rangle$  has energy:

$$\langle\phi|H_0|\phi\rangle = \sum_i \sum_j \langle\phi|i\rangle\langle i|H_0|j\rangle\langle j|\phi\rangle = \sum_i |\langle\phi|i\rangle|^2 E_i \quad \text{since} \quad \langle j|E_i|i\rangle = 0$$

So we could use the solution to an easier problem (the eigenvalue problem, which we need solve only once per Hamiltonian) so that we never need to apply the complicated Hamiltonian to the complicated mixed state! This is a very useful trick - reformulating a problem so that we can make use of some work that has already been done. In this case the single, hard, problem of finding the energy of a mixed state is changed to the many, easier, problems of finding the energy of the eigenstates and the amount of each eigenstate in the mixed state.

## 1.8 Using Bras to pick Kets

One of the most useful algebraic tricks in quantum mechanics is to multiply a sum of terms by a complex conjugate wavefunction, and integrate the product over all space. Orthogonality often means that this procedure can be used to ‘pick’ a single term from the sum. In Dirac notation this procedure simply becomes applying  $\langle i_m|$ .

For example, if we have an expansion of a mixed state  $\Phi$  in eigenstates  $i_n$ :  $\hat{H}|\Phi\rangle = \sum_n \hat{H}|i_n\rangle\langle i_n|\Phi\rangle$ , we can remove the sum by  $\langle i_m|$ :

$$\langle i_m|\hat{H}|\Phi\rangle = \langle i_m|\sum_n \hat{H}|i_n\rangle\langle i_n|\Phi\rangle = E_m\langle i_m|\Phi\rangle$$

This works because  $\langle i_m|\hat{H}|i_n\rangle = E_m\delta_{nm}$ . This procedure is analogous to taking components of a vector.

## 1.9 Good quantum numbers

It is normal to think of the eigenvalues as labelling states. In that case they are just called quantum numbers. A set of eigenvalues from a complete commuting set of operators are called good quantum numbers. The eigenvalues from a non-commuting operator are a bad quantum numbers, because their values cannot be known simultaneously.

This is not quite as simple as it seems. In real systems the Hamiltonian may contain many small terms (perturbations) which may not commute with the operators which commute with the unperturbed Hamiltonian. Although in principle the quantum numbers are no longer good, in practice they are often used.

An example of this is in spin-orbit coupling of angular momenta in many-electron atoms. Here  $L_z = \sum_i l_{iz}$  is a good quantum number in the absence of spin-orbit coupling, but  $\hat{l}_{iz}$  does not commute with the spin orbit coupling operator  $\sum_i \hat{l}_i \cdot \hat{s}_i$ . Thus for light atoms, where spin-orbit coupling is weak,  $L_z$  is often used although it is not strictly a good quantum number.

## 2 Review: Time-Independent Non-degenerate Perturbation Theory

### 2.1 Small changes to the Hamiltonian

There are very few problems in quantum mechanics which can be solved exactly. However, we are often interested in the effect of a small change to a system, and in such cases we can proceed by assuming that this causes only a small change in the eigenstates. *Perturbation theory* provides a method for finding approximate energy eigenvalues and eigenfunctions for a system whose Hamiltonian is of the form

$$\hat{H} = \hat{H}_0 + \lambda \hat{V}$$

where  $\hat{H}_0$  is the ‘main bit’ of the Hamiltonian of an exactly solvable system, for which we know the eigenvalues,  $E_n^{(0)}$ , and eigenfunctions,  $|n^{(0)}\rangle$ , and  $\lambda \hat{V}$  is a *small*, time-independent perturbation.  $\lambda$  is a small expansion parameter,  $\hat{H}$ ,  $\hat{H}_0$  and  $\hat{V}$  are Hermitean operators. Using perturbation theory, we can get approximate solutions for  $\hat{H}$  using as basis functions eigenstates of the similar, exactly solvable system  $\hat{H}_0$ .

Assuming that  $\hat{H}$  and  $\hat{H}_0$  possess discrete, non-degenerate eigenvalues only, we write

$$\hat{H}_0 |n^{(0)}\rangle = E_n^{(0)} |n^{(0)}\rangle$$

in Dirac notation. The states  $|n^{(0)}\rangle$  are orthonormal. The effect of the perturbation will be modify each state and its corresponding energy slightly; The eigenstate  $|n^{(0)}\rangle$  will become  $|n\rangle$  and  $E_n^{(0)}$  will shift to  $E_n$ , where

$$\hat{H} |n\rangle = E_n |n\rangle$$

We solve the full eigenvalue problem by assuming that we can expand  $E_n$  and  $|n\rangle$  in a power series in  $\lambda$ :

$$\begin{aligned} E_n &= E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \\ |n\rangle &= |n^{(0)}\rangle + \lambda |n^{(1)}\rangle + \lambda^2 |n^{(2)}\rangle + \dots \end{aligned}$$

where the correction terms,  $E_n^{(1)}$ ,  $E_n^{(2)}$ ,  $\dots$  and  $|n^{(1)}\rangle$ ,  $|n^{(2)}\rangle$ ,  $\dots$  are of successively higher order in the power  $\lambda$  (i.e. they get smaller)

Substituting these expansions into the eigenvalue equation for the full Hamiltonian:

$$\begin{aligned} (\hat{H}_0 + \lambda \hat{V}) (|n^{(0)}\rangle + \lambda |n^{(1)}\rangle + \lambda^2 |n^{(2)}\rangle + \dots) = \\ (E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots) (|n^{(0)}\rangle + \lambda |n^{(1)}\rangle + \lambda^2 |n^{(2)}\rangle + \dots) \end{aligned}$$

and equating terms of the same power in  $\lambda$ , gives

$$\begin{aligned} \lambda^0 : \quad \hat{H}_0 |n^{(0)}\rangle &= E_n^{(0)} |n^{(0)}\rangle \\ \lambda^1 : \quad (\hat{H}_0 - E_n^{(0)}) |n^{(1)}\rangle &= (E_n^{(1)} - \hat{V}) |n^{(0)}\rangle \\ \lambda^2 : \quad (\hat{H}_0 - E_n^{(0)}) |n^{(2)}\rangle &= (E_n^{(1)} - \hat{V}) |n^{(1)}\rangle + E_n^{(2)} |n^{(0)}\rangle \end{aligned}$$

## 2.2 First order energy shifts

The first equation simply restates that  $|n^{(0)}\rangle$  is the eigenvector of  $\hat{H}_0$  with eigenvalue  $E_n^{(0)}$ .

The second equation is more interesting; premultiplying with, say, the  $k$ th unperturbed eigenstate yields

$$\langle k^{(0)} | (\hat{H}_0 - E_n^{(0)}) | n^{(1)} \rangle = \langle k^{(0)} | (E_n^{(1)} - \hat{V}) | n^{(0)} \rangle$$

The first term on the left-hand side can be rewritten

$$\begin{aligned} \langle k^{(0)} | \hat{H}_0 | n^{(1)} \rangle &= \langle n^{(1)} | \hat{H}_0^\dagger | k^{(0)} \rangle^* \\ &= \langle n^{(1)} | \hat{H}_0 | k^{(0)} \rangle^* && \text{since } \hat{H}_0 \text{ is Hermitean} \\ &= E_k^{(0)} \langle n^{(1)} | k^{(0)} \rangle^* && \text{since } E_k^{(0)} \text{ is real} \\ &= E_k^{(0)} \langle k^{(0)} | n^{(1)} \rangle \end{aligned}$$

Thus the equation becomes

$$(E_k^{(0)} - E_n^{(0)}) \langle k^{(0)} | n^{(1)} \rangle = E_n^{(1)} \delta_{k,n} - \langle k^{(0)} | \hat{V} | n^{(0)} \rangle$$

Since we have chosen  $k$  arbitrarily, let's see what happens when we make the particular choice  $k = n$ . The left-hand side vanishes, and the Kronecker delta on the right-hand side is equal to one, yielding

$$E_n^{(1)} = \langle n^{(0)} | \hat{V} | n^{(0)} \rangle \equiv V_{nn}$$

$$E_n = \langle n^{(0)} | \hat{H}_0 + \hat{V} | n^{(0)} \rangle$$

which is one of *the most useful results in quantum mechanics*. It tells us how to calculate the change in the  $n$ th energy eigenvalue, to first order:

**The shift in energy induced by a perturbation is given to first order by the expectation value of the perturbation with respect to the unperturbed state.**

Thus first order time independent perturbation is equivalent to making the approximation that *the wavefunction does not change*. Since  $E_n^{(1)}$  represents the *shift* in the energy to first order, it is often written  $\Delta E_n$ .

Note that the  $\lambda$ s from 2.1 have disappeared: we used them only to remind ourselves how small each term was. From now on when we use perturbation theory we will just have to remember that  $\hat{V}$  is small.

### 2.3 Mixing of the eigenstates of $\hat{H}_0$

If, on the other hand, we choose  $k \neq n$  as the premultiplying bra, we can find the first-order correction to the eigenvectors. We have

$$(E_k^{(0)} - E_n^{(0)})\langle k^{(0)}|n^{(1)}\rangle = -\langle k^{(0)}|\hat{V}|n^{(0)}\rangle \quad k \neq n$$

We can expand the first-order correction term  $|n^{(1)}\rangle$  in terms of the complete set of unperturbed energy eigenstates;

$$|n^{(1)}\rangle = \sum_m c_{nm}^{(1)} |m^{(0)}\rangle$$

Taking the scalar product with  $\langle k^{(0)}|$  yields

$$\begin{aligned} \langle k^{(0)}|n^{(1)}\rangle &= \sum_m c_{nm}^{(1)} \langle k^{(0)}|m^{(0)}\rangle \\ &= \sum_m c_{nm}^{(1)} \delta_{k,m} \\ &= c_{nk}^{(1)} \end{aligned}$$

so that the coefficients are given by

$$c_{nk}^{(1)} = \frac{\langle k^{(0)}|\hat{V}|n^{(0)}\rangle}{(E_n^{(0)} - E_k^{(0)})} \quad k \neq n$$

Unfortunately we see that we have not determined the coefficient for the case  $n = k$ . However, this may be set to zero without loss of generality, so that we end up with the following expression for the first-order correction to the  $n$ th energy eigenvector:

$$|n^{(1)}\rangle = \sum_{k \neq n} \frac{\langle k^{(0)}|\hat{V}|n^{(0)}\rangle}{(E_n^{(0)} - E_k^{(0)})} |k^{(0)}\rangle \equiv \sum_{k \neq n} \frac{V_{kn}}{(E_n^{(0)} - E_k^{(0)})} |k^{(0)}\rangle$$

We speak of the perturbation *mixing the unperturbed eigenfunctions* since the effect is to add to the unperturbed eigenfunction,  $|n^{(0)}\rangle$ , a small amount of each of the other unperturbed eigenfunctions. The denominator suggests that states with similar energies are more strongly mixed.

Unlike the formula for the energy shift, we are faced in general with evaluating an infinite sum to find the correction to the eigenfunctions.

### 2.4 Notes

- For the first-order changes to the eigenfunction to be small we must have:

$$\langle k^{(0)}|\hat{V}|n^{(0)}\rangle \equiv V_{kn} \ll |(E_n^{(0)} - E_k^{(0)})| \quad \text{for all } k \neq n$$

- Similarly, we require that the level shift be small compared to the level spacing in the unperturbed system:

$$|E_n^{(1)}| \ll \min |(E_n^{(0)} - E_k^{(0)})|$$

- These conditions may break down *if there are degeneracies in the unperturbed system*. However, we need only assume that the *particular energy level whose shift we are calculating is non-degenerate* for the preceding analysis to be correct.

## 2.5 Higher Orders

It may turn out that the first-order calculation yields zero shift, typically due to symmetry. In this case we must consider what happens at second order. Going back to our results for the  $\lambda^2$  terms in the expansion we saw that:

$$(\hat{H}_0 - E_n^{(0)}) |n^{(2)}\rangle = (E_n^{(1)} - \hat{V}) |n^{(1)}\rangle + E_n^{(2)} |n^{(0)}\rangle$$

Proceeding as before and taking the scalar product with an unperturbed eigenstate gives:

$$\langle k^{(0)} | (\hat{H}_0 - E_n^{(0)}) |n^{(2)}\rangle = \langle k^{(0)} | (E_n^{(1)} - \hat{V}) |n^{(1)}\rangle + E_n^{(2)} \delta_{k,n}$$

The left-hand side may be simplified as before, replacing  $\hat{H}_0$  by the eigenvalue  $E_k^{(0)}$  so that

$$(E_k^{(0)} - E_n^{(0)}) \langle k^{(0)} |n^{(2)}\rangle = \langle k^{(0)} | (E_n^{(1)} - \hat{V}) |n^{(1)}\rangle + E_n^{(2)} \delta_{k,n}$$

When  $k = n$  this simplifies to give

$$\begin{aligned} \Delta E_n = E_n^{(2)} &= \langle k^{(0)} | \hat{V} |n^{(1)}\rangle - E_n^{(1)} \langle k^{(0)} |n^{(1)}\rangle \\ &= \langle k^{(0)} | \hat{V} |n^{(1)}\rangle \quad \text{using the orthogonality of } |n^{(1)}\rangle \text{ and } |n^{(0)}\rangle \\ &= \sum_{m \neq n} c_{nm}^{(1)} V_{nm} \quad \text{using the expansion of } |n^{(1)}\rangle \\ &= \sum_{m \neq n} \frac{V_{mn} V_{nm}}{(E_n^{(0)} - E_m^{(0)})} \\ &= \sum_{m \neq n} \frac{|V_{mn}|^2}{(E_n^{(0)} - E_m^{(0)})} \end{aligned}$$

Again we are in the position in general of having to evaluate an infinite sum of terms to get the desired result, although sometimes there are reasons, such as symmetry, why only a finite number of terms are non-zero.

## 2.6 Example

Consider a simple harmonic oscillator in its ground state, to which we apply a perturbation  $\hat{V} = \lambda x^2$ . We know the wavefunction  $|n_0^{(0)}\rangle = [m\omega_0/\pi\hbar]^{1/4} \exp\{-m\omega_0 x^2/2\hbar\}$ , so we can evaluate the first order shift in energy according to the perturbation theory:

$$\Delta E_0 = \langle n_0^{(0)} | \lambda x^2 |n_0^{(0)}\rangle = \lambda \sqrt{m\omega_0/\pi\hbar} \int x^2 \exp\{-m\omega_0 x^2/\hbar\} dx = \frac{\lambda}{2} \frac{\hbar}{m\omega_0}$$

In this case we know the exact shift, since the perturbation is simply an additional harmonic potential, giving a total  $k = m\omega_0^2 + 2\lambda$  and an exact ground state energy of  $\frac{1}{2}\hbar\sqrt{\omega_0^2 + 2\lambda/m}$ . It is easy to verify that to first order in  $\lambda$  these expressions are identical.

To determine the amount of mixing of states, we need to evaluate matrix elements like  $\langle n_0^{(0)} | \lambda x^2 |n_i^{(0)}\rangle$ . We won't evaluate these here, but we will note that for odd  $i$  the integral is zero - the symmetric perturbation only mixes in symmetric excited states.

### 3 Dealing with Degeneracy

#### 3.1 Time-Independent Degenerate Perturbation Theory

We have seen how we can find approximate solutions for a system whose Hamiltonian is of the form

$$\hat{H} = \hat{H}_0 + \lambda \hat{V}$$

When we assumed that  $\hat{H}$  and  $\hat{H}_0$  possess discrete, non-degenerate eigenvalues only. This led to a mixing of states  $|n\rangle = |n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \dots$ , where

$$|n^{(1)}\rangle = \sum_{k \neq n} \frac{\langle k^{(0)} | \hat{V} | n^{(0)} \rangle}{(E_n^{(0)} - E_k^{(0)})} |k^{(0)}\rangle \equiv \sum_{k \neq n} \frac{V_{kn}}{(E_n^{(0)} - E_k^{(0)})} |k^{(0)}\rangle$$

Clearly, if  $E_n^{(0)} = E_k^{(0)}$  this diverges. As do the higher order energy shifts (see 2.5). Thus for the degenerate case we cannot associate a perturbed state  $|n\rangle$  with a particular unperturbed state  $|n^{(0)}\rangle$ : we need to take a different approach. In fact, the approximation we make is completely different: we assume that the small perturbation *only mixes those states which are degenerate*. We then solve the problem exactly for that subset of states.

Assume that  $\hat{H}_0$  possesses  $N$  degenerate eigenstates with eigenvalue  $E_d^{(0)}$ . It may also possess non-degenerate eigenstates, which can be treated separately by non-degenerate perturbation theory. We write the perturbed eigenstate as an linear expansion in the unperturbed degenerate eigenstates only:

$$|n\rangle = \sum_m |m^{(0)}\rangle \langle m^{(0)} | n \rangle = \sum_m c_{nm} |m^{(0)}\rangle$$

Where  $m$  here runs over degenerate states only. The TISE now becomes:

$$[\hat{H}_0 + \hat{V}] |n\rangle = [\hat{H}_0 + \hat{V}] \sum_m c_{nm} |m^{(0)}\rangle = E_n \sum_m c_{nm} |m^{(0)}\rangle$$

but we know that for all degenerate eigenstates  $m$ .  $\hat{H}_0 |m^{(0)}\rangle = E_d^{(0)} |m^{(0)}\rangle$ . So we obtain:

$$\sum_m c_{nm} \hat{V} |m^{(0)}\rangle = (E_n - E_d^{(0)}) \sum_m c_{nm} |m^{(0)}\rangle$$

premultiplying by some unperturbed state  $\langle k^{(0)} |$  gives

$$\sum_m c_{nm} [\langle k^{(0)} | \hat{V} | m^{(0)} \rangle - \delta_{mk} (E_n - E_d^{(0)})] = 0$$

There is a similar equation from each unperturbed state  $|k^{(0)}\rangle$ . We thus have an eigenvalue problem: the eigenvector has elements  $c_{nm}$  and the eigenvalues are  $\Delta E_n = E_n - E_d^{(0)}$ . Writing the matrix elements between the  $k^{th}$  and  $m^{th}$  unperturbed degenerate states as  $V_{km} \equiv \langle k | \hat{V} | m \rangle$  we recover the determinantal equation:

$$\begin{vmatrix} V_{11} - \Delta E_n & V_{12} & \dots & V_{1N} \\ V_{21} & V_{22} - \Delta E_n & \dots & V_{2N} \\ \dots & \dots & \dots & \dots \\ V_{N1} & V_{N2} & \dots & V_{NN} - \Delta E_n \end{vmatrix} = 0$$

The  $N$  eigenvalues obtained by solving this equation give the shifts in energy due to the perturbation, and the eigenvectors give the perturbed states  $|n\rangle$  in the unperturbed basis set  $|m^{(0)}\rangle$ .

### 3.2 Notes

- The perturbed eigenstates of  $\hat{H}$  are linear combinations of degenerate eigenstates of  $\hat{H}_0$ . This means that they too are eigenstates of  $\hat{H}_0$  from a different eigenbasis.
- If  $\hat{H}_0$  is compatible with  $\hat{V}$ , i.e.  $[\hat{H}_0, \hat{V}] = 0$ , then there is no mixing with non-degenerate states and the analysis above is exact.
- Notice how the mathematics mimics the quantum mechanics. Without the perturbation the eigenbasis of  $\hat{H}_0$  is not unique. When we try to determine its energy shift we find a matrix equation which can *only* be solved for *specific* values of  $\Delta E_n$ . These  $\Delta E_n$  in turn correspond to specific choices for the coefficients  $c_{nm}$ . Thus to solve the equations we are forced to collapse the wavefunction onto an eigenstate of  $\hat{V}$ .  $V_{km}$  is a Hermitian matrix, and consequently has real eigenvalues.

### 3.3 Example of degenerate perturbation theory: Stark Effect in Hydrogen

The change in energy levels in an atom due to an external electric field is known as the Stark effect. The perturbing potential is thus  $\hat{V} = eEz = eEr \cos \theta$ . Ignoring spin, we examine this effect on the fourfold degenerate  $n=2$  levels. We will label these by their appropriate quantum number:  $|l, m\rangle$ .

$$\begin{aligned} u_{00} &= (8\pi a_0^3)^{-1/2} (1 - r/2a_0) e^{-r/2a_0}; & u_{10} &= (8\pi a_0^3)^{-1/2} (r/2a_0) \cos \theta e^{-r/2a_0} \\ u_{11} &= (\pi a_0^3)^{-1/2} (r/8a_0) \sin \theta e^{i\phi} e^{-r/2a_0} & u_{1-1} &= (\pi a_0^3)^{-1/2} (r/8a_0) \sin \theta e^{-i\phi} e^{-r/2a_0} \end{aligned}$$

From the analysis above, we need to calculate the matrix elements.

$$V_{lm, l'm'} = \langle l, m | eEz | l', m' \rangle = eE \int \int \int u_{lm}^*(r \cos \theta) u_{l'm'} r^2 \sin \theta d\theta d\phi dr$$

It turns out that many of these are zero, since if any of the integrals are zero their product will be. Looking first at parity, it is clear that  $eEz$  has odd parity ( $eE(-r) \cos \theta = -eEr \cos \theta$ ),  $u_{00}$  has even parity and  $u_{1m}$  have odd parity. Since the integral over all space of any odd function is zero,  $V_{00,00} = V_{1m,1m'} = 0$ . Secondly,  $\int_0^{2\pi} e^{\pm i\phi} d\phi = 0$ , so  $V_{00,11} = V_{00,1-1} = V_{11,00} = V_{1-1,00} = 0$ .

Since the perturbation is real,  $V_{00,10} = V_{10,00}$  and the only remaining non-zero matrix element is:

$$\langle 00 | eEr \cos \theta | 10 \rangle = (8\pi a_0^3)^{-1} \int_0^{2\pi} d\phi \int_0^\pi \cos^2 \theta \sin \theta d\theta \int_0^\infty (1 - r/2a_0) e^{-r/a_0} r^4 / 2a_0 dr = -3eEa_0$$

the determinantal equation is then:

$$\begin{vmatrix} -\Delta E_n & -3eEa_0 & 0 & 0 \\ -3eEa_0 & -\Delta E_n & 0 & 0 \\ 0 & 0 & -\Delta E_n & 0 \\ 0 & 0 & 0 & -\Delta E_n \end{vmatrix} = (\Delta E_n)^4 - (\Delta E_n)^2 (3eEa_0)^2 = 0$$

The solutions to this are  $\Delta E_n = \pm 3eEa_0, 0, 0$ . The degeneracy of the states  $u_{11}$  and  $u_{1-1}$  is not lifted, but the new non-degenerate eigenstates corresponding to  $\Delta E_n = \pm 3eEa_0$  are  $(u_{00} \mp u_{10})/\sqrt{2}$ . Consequently, the spectral line corresponding to the  $n = 2 \rightarrow n = 1$  Lyman- $\alpha$  transition is split into three if the hydrogen atom is in an electric field.

A curious aspect of these eigenstates is that they are not eigenstates of  $\mathbf{L}^2$ , although they are eigenstates of  $L_z$ . Nor do they have definite parity. In an electric field, therefore, the total angular

momentum is not a good quantum number. Note that this effect is specific to hydrogen, since in other elements the  $s$  and  $p$  levels are not degenerate.

Experimental results confirm this theory beautifully - the splitting of levels in hydrogen varies linearly with the applied field strength, while in all other atoms it varies quadratically: the first order perturbation is zero.

Looking at the electrostatics: the energy of a spherically symmetric charge density in a uniform field is clearly independent of orientation. To have any orientation dependence the object must have a dipole moment. The combination of  $2s$  and  $2p$  wavefunctions achieves this.

### 3.4 Symmetry and Degeneracy

In real systems degeneracy almost always related to symmetry. In general if the probability density has lower symmetry than the Hamiltonian, the wavefunction will be degenerate.

There is a clear physical reason behind this. Consider the  $2p_x$  orbital in hydrogen: it has a lobe along the  $x$ -axis. However, there is no measurable quantity which defines an  $x$ -axis - the coordinate system is just introduced by physicists to help solve the equations. The lobe could just as well point in the  $y$  or  $z$  or  $(27, 43.2, -12)$  direction. Thus the  $p_x$  orbital has lower symmetry than the Hamiltonian (spherically symmetric potential), and is degenerate with  $p_y$  and  $p_z$ . Likewise the spin: we talk about ‘spin up’, but there is no way to define ‘up’ from the Hamiltonian. Thus there is degeneracy between spin states ‘up’ and ‘down’.

If we reduce the symmetry of the Hamiltonian, we now ‘lift’ the degeneracy. (i.e. the levels no longer have the same energy). For example, an applied magnetic field defines an axis and lowers the symmetry of the Hamiltonian. If the field is weak, we can use perturbation theory and assume we still have  $p$  orbitals (Zeeman effect). Now, the orbitals must be eigenstates not only of  $\hat{H}_0$ , but also of  $\mu \cdot \mathbf{B}$  where  $\mu$  is the magnetic dipole moment. The degenerate energy level splits into several different energy levels, depending on the relative orientation of the moment and the field: The degeneracy is lifted by the reduction in symmetry.

### 3.5 Time-variation of expectation values: Degeneracy and constants of motion

The time variation of the expectation value of an operator  $\hat{A}$  which commutes with the Hamiltonian is:

$$\frac{d}{dt} \langle \Phi | \hat{A} | \Phi \rangle = \int d^3r \frac{d\Phi^*}{dt} \hat{A} \Phi + \Phi^* \hat{A} \frac{d\Phi}{dt}$$

but since  $i\hbar \frac{d\Phi}{dt} = \hat{H} \Phi$  and  $-i\hbar \frac{d\Phi^*}{dt} = \hat{H}^* \Phi^*$

$$-i\hbar \frac{d}{dt} \langle \Phi | \hat{A} | \Phi \rangle = \int (\hat{H}^* \Phi^* \hat{A} \Phi - \Phi^* \hat{A} \hat{H} \Phi) d^3r = \langle \Phi | [\hat{H}, \hat{A}] | \Phi \rangle$$

Where we also use the fact that  $\hat{H}$  is Hermitian. Thus if  $\hat{H}$  commutes with  $\hat{A}$  ( $[\hat{H}, \hat{A}] = 0$ ), the expectation value of  $A$  is independent of time. It is a conserved quantity.

As we have seen above, if we have degenerate eigenstates of the Hamiltonian,  $\hat{H}$ , then there must be some other operator  $\hat{A}$  which commutes with the Hamiltonian for which there are energy-degenerate eigenstates with different eigenvalues  $A$ . These eigenvalues,  $A$ , are then constants of the motion. Moreover, if  $\Phi$  is an eigenfunction of  $\hat{H}$ , then  $\hat{A}\Phi$  is also an eigenfunction of  $\hat{H}$ .

$$\hat{H}(\hat{A}\Phi) = \hat{A}\hat{H}\Phi = \hat{A}(E\Phi) = E(\hat{A}\Phi)$$

There is a three way link between symmetry, degeneracy and conserved quantities.

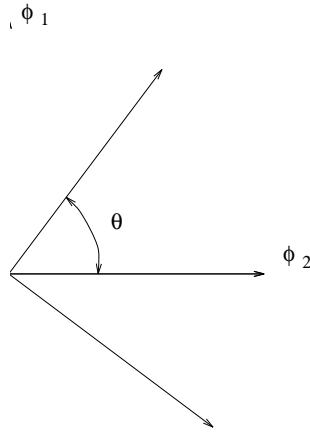


Figure 1: Any linear combination of two degenerate eigenstates produces another eigenstate.

### 3.6 Wavefunction Collapse onto degenerate levels

Refer back to the postulates of quantum mechanics: We know that acting with an operator  $\hat{A}$  on an eigenstate  $|\alpha_n\rangle$  of that operator gives us an eigenvalue  $A_n$ , which corresponds to a measurable quantity.

There is no guarantee that  $|\alpha_n\rangle$  is the only eigenstate of  $\hat{A}$  which has this eigenvalue (e.g. energy levels in hydrogen). Different states with the same eigenvalue are referred to as degenerate.

Assume we find two orthogonal, degenerate eigenstates of  $\hat{A}$ :  $|\alpha_1\rangle$  and  $|\alpha_2\rangle$ . i.e.  $\hat{A}|\alpha_1\rangle = A_1|\alpha_1\rangle$  and  $\hat{A}|\alpha_2\rangle = A_1|\alpha_2\rangle$ . We also see that

$$\hat{A}(\cos\theta|\alpha_1\rangle + \sin\theta|\alpha_2\rangle) = A_1(\cos\theta|\alpha_1\rangle + \sin\theta|\alpha_2\rangle)$$

for any  $\theta$ . We use  $\cos\theta$  for the expansion instead of the normal  $c_i$  to emphasise the similarity between eigenstates and vectors. It also allows for easy normalisation since  $\cos^2\theta + \sin^2\theta = 1$ .

Thus any linear combination of degenerate eigenstates produces another eigenstate. There is still only twofold degeneracy, because there are only two orthogonal states,  $(\sin\theta|\alpha_1\rangle - \cos\theta|\alpha_2\rangle)$  being the other one. The complete set of orthonormal eigenstates for  $\hat{A}$  is thus not a unique quantity, since we can choose any  $\theta$  to generate a pair of degenerate eigenstates.

A consequence of this is that when a measurement is made of  $\hat{A}$  which finds  $A_1$ , there is not a complete collapse of the wavefunction.

Consider measuring observable  $A$  in a system in a general state  $|\Phi\rangle$ . By expanding  $|\Phi\rangle$  in the eigenstates of  $\hat{A}$ :  $|\Phi\rangle = \sum_i c_i|\alpha_i\rangle$  we find the probability that the measurement will yield result  $A_1$  is

$$|\langle\alpha_1|\Phi\rangle|^2 + |\langle\alpha_2|\Phi\rangle|^2 \equiv |c_1|^2 + |c_2|^2$$

The measurement has determined that we are either in state  $\alpha_1$  or  $\alpha_2$ , but not which. Thus there is a partial collapse of the wavefunction onto a linear combination of them:

$$(\cos\theta|\alpha_1\rangle + \sin\theta|\alpha_2\rangle); \quad \cos\theta = \frac{c_1}{\sqrt{|c_1|^2 + |c_2|^2}}$$

which is itself an eigenvector of  $\hat{A}$ .

Thus, in the case of degenerate final states, the final wavefunction after the measurement *does* depend on the initial wavefunction. The generalisation of this to the case of many degenerate states is straightforward.

## 4 Degeneracy, Symmetry and Conservation Laws

### 4.1 Distinguishing between eigenstates, Quantum numbers as labels

How can we distinguish between quantum states  $|\alpha_n\rangle$  which have degenerate values of A? The obvious way is to measure the quantised observables and use them to label the state. We must be sure not to make measurements which change the state. Thus all measurements should correspond to commuting operators (Compatible observations: see QP3). In the non-degenerate case measuring energy is sufficient, but in hydrogen, for example, that we used quantum numbers  $n$  (for energy, operator  $\hat{H}$ ),  $l$  (for total angular momentum,  $\hat{\mathbf{L}}^2$ ) and  $m_l$  (one component of angular momentum,  $\hat{L}_z$ ).

Continuing the example of twofold degeneracy (3.6), suppose that some operator  $\hat{B}$  is compatible with  $\hat{A}$ . This means that  $[A, B] = 0$  and  $\hat{A}$  and  $\hat{B}$  have a common eigenbasis. i.e. some  $\theta$  and  $\theta + \pi/2$  give eigenstates of both  $\hat{A}$  and  $\hat{B}$  in the form  $|\alpha(\theta)\rangle = (\cos\theta|\alpha_1\rangle + \sin\theta|\alpha_2\rangle)$ .

To find the appropriate value of  $\theta$ , we have a similar problem to that encountered in 3.1 and must solve for the eigenvectors of:

$$\begin{pmatrix} \langle\alpha_1|\hat{B}|\alpha_1\rangle & \langle\alpha_1|\hat{B}|\alpha_2\rangle \\ \langle\alpha_2|\hat{B}|\alpha_1\rangle & \langle\alpha_2|\hat{B}|\alpha_2\rangle \end{pmatrix}$$

The eigenvalues of this equation are the quantised measurable values of  $\hat{B}$ . If both of these are equal, there must be another measurable C which will distinguish the two states.

The generalisation to many degenerate levels is straightforward. If there are  $n$  orthogonal degenerate eigenstates of  $\hat{A}$ , (therefore an  $n$ -dimensional space in which every unit vector is an eigenstate of  $\hat{A}$ ), compatibility of eigenbases means there are at least  $n$  eigenstates of  $\hat{B}$ . It is now possible that all these have different B eigenvalues, or that at least two have the same eigenvalue, in which case if we want a specific set of orthogonal eigenstates, we must look for another compatible operator  $\hat{C}$ .

When the set of operators is sufficiently large that there is a unique set of eigenvalues for each eigenstate, we call it a complete commuting set of operators. An example is  $\hat{H}$ ,  $\hat{\mathbf{L}}^2$ ,  $\hat{S}_z$  and  $\hat{L}_z$  in hydrogen. The complete commuting set is not unique for a given Hamiltonian, for hydrogen we could have used  $\hat{H}$ ,  $\hat{\mathbf{L}}^2$ ,  $\hat{S}_x$  and  $\hat{L}_x$  or  $\hat{H}$ ,  $\hat{\mathbf{L}}^2$ ,  $\hat{J}$  and  $\hat{L}_z$ . If one of the quantum numbers can be written in terms of the others then it is redundant. If two of the quantum numbers come from non-commuting operators, then the set does not define a state since the full set of measurements could not be performed without changing the wavefunction.

### 4.2 Example

Consider the 2D harmonic oscillator  $V_0 = \frac{1}{2}m\omega^2(x^2 + y^2)$ . If we measure the energy and find it to be  $2\hbar\omega$ , then the state could be  $|n_x = 1, n_y = 0\rangle$  or  $|n_x = 0, n_y = 1\rangle$  or any linear combination. To fully define any state we require any two quantum numbers:  $n_x, n_y$  and  $E = (n_x + n_y + 1)\hbar\omega$ .

Suppose we measure the energy and find  $3\hbar\omega$ : there are three degenerate possibilities. Suppose we then apply a perturbation  $\Delta V = \lambda x^2$  (see 2.6). This breaks the symmetry and collapses the wavefunction onto either  $|1, 1\rangle$ ,  $|2, 0\rangle$  or  $|0, 2\rangle$ . The perturbation matrix (see 3.1)  $\langle n_x, n_y | \Delta V | n_x, n_y \rangle$  is diagonal provided we choose the basis with  $x$  along the direction of the perturbation. it has eigenvalues  $(2n_x + \frac{1}{2})\lambda\hbar/m\omega$ . If we then measure the energy and find  $E = 3\hbar\omega + \lambda\hbar/2m\omega$  then we know that the state is  $|0, 2\rangle$ .

**Aside:** Consider mixing with the non-degenerate states. By symmetry the perturbation does not mix  $n_x = 0$  and  $n_x = 1$  states, nor does it affect  $n_y$  (see 2.3). Thus applying the perturbation may induce a transition from  $|0, 2\rangle$  to  $|2, 2\rangle$ ,  $|4, 2\rangle$  etc. but not to  $n_x = \text{odd}$  or  $n_y \neq 2$ .

### 4.3 Conservation of Particles

Recall that the probability density for finding a particle at position  $\mathbf{r}$  and time  $t$  is given by:

$$P(\mathbf{r}, t) = \Phi^*(\mathbf{r}, t)\Phi(\mathbf{r}, t) = |\Phi(\mathbf{r}, t)|^2$$

Where  $\Phi(\mathbf{r}, t)$  is the value of the wavefunction for the particle at that point. Thus if we want the wavefunction to contain one particle per volume  $V$  it must be normalised to give:

$$\int_V |\Phi(\mathbf{r}, t)|^2 d^3\mathbf{r} = 1$$

This contrasts with the normalisation condition for a single particle:

$$\int |\Phi(\mathbf{r}, t)|^2 d^3\mathbf{r} = \langle \Phi | \Phi \rangle = 1$$

which is taken over all space. The problem with the second form arises for wavefunctions for which the normalisation integral over all space does not converge. A simple example of this is the free particle wavefunction  $Ae^{ikx}$ . If we consider only a finite region of space, however, as for a bound particle, the normalisation can be done.

We now consider the conservation of particles. This requires that the change of the normalisation integral in time be equal to the flow of particles out of the volume  $\Omega$ . This derivative is:

$$\begin{aligned} \frac{\partial}{\partial t} \int_V |\Phi(\mathbf{r}, t)|^2 d^3\mathbf{r} &= \int_V \Phi^* \frac{\partial \Phi}{\partial t} + \frac{\partial \Phi^*}{\partial t} \Phi d^3\mathbf{r} && \text{using the TDSE } \frac{\partial \Phi}{\partial t} = \frac{i\hbar}{2m} \nabla^2 \Phi - \frac{i}{\hbar} \hat{V} \Phi \\ &= \frac{i\hbar}{2m} \int_{\Omega} \Phi^* \nabla^2 \Phi - (\nabla^2 \Phi^*) \Phi d^3\mathbf{r} && \text{Terms in } V \text{ cancel for real } V \\ &= \frac{i\hbar}{2m} \int_{\Omega} \nabla \cdot [\Phi^* \nabla \Phi - (\nabla \Phi^*) \Phi] d^3\mathbf{r} && \text{Easiest to show in reverse} \\ &= \frac{i\hbar}{2m} \int_A \Phi^* \nabla \Phi - (\nabla \Phi^*) \Phi \cdot d\mathbf{S} && \text{Integrate by parts, assume } \Phi(r \rightarrow \infty) \rightarrow 0 \end{aligned}$$

The final integral is obtained by integration by parts (Green's theorem) in exactly the same way as for Gauss' theorem, so that  $\mathbf{S}$  is the surface area bounding the volume  $\Omega$  and the integral is over the component of the vector normal to the area  $d\mathbf{S}$ .

We define a *probability current density*  $\mathbf{J}$ . (beware: some books call it  $\mathbf{S}$ )

$$\mathbf{J}(\mathbf{r}, t) = -\frac{i\hbar}{2m} [\Phi^* \nabla \Phi - (\nabla \Phi^*) \Phi]$$

This combination of  $\Phi$ ,  $\Phi^*$  and their differentials is called the *Wronskian*. It now follows that:

$$\frac{\partial}{\partial t} \int_V |\Phi(\mathbf{r}, t)|^2 d^3\mathbf{r} = - \int_V \nabla \cdot \mathbf{J}(\mathbf{r}, t) d^3\mathbf{r} = - \int_A \mathbf{J} \cdot d\mathbf{S}$$

This gives a differential relation  $\frac{\partial |\Phi|^2}{\partial t} = -\nabla \cdot \mathbf{J}$

This relation looks like the various continuity equations in fluid flow and electromagnetism. *Rate of change in amount of stuff = rate that stuff flows away.* Since  $\mathbf{J}(\mathbf{r}, t)$  is the real part of  $\frac{\hbar}{im} \Phi^* \nabla \Phi$ , this is consistent with the velocity operator being  $\hat{\mathbf{v}} = \hat{\mathbf{p}}/m = \frac{\hbar}{im} \nabla$ .

It is important to be careful about  $\mathbf{J}(\mathbf{r}, t)$  because it is *not measurable*. This is because it implies simultaneous knowledge on position and velocity, at variance with the uncertainty principle. Nevertheless, if  $\mathbf{J}$  does not depend on  $\mathbf{r}$  (as for a free particle  $\Phi = Ae^{ik\cdot\mathbf{r}}$ ) an accurate velocity measurement can be obtained.

## 4.4 Translational Symmetry and Conservation of Momentum

Consider a transformation operator in 1 dimension  $\hat{D}$  which acts on the coordinates of a system as a displacement  $\hat{D}[f(x)] = f(x+l)$ . The eigenfunctions of  $\hat{D}$  satisfy  $\hat{D}|\phi(x)\rangle = d|\phi(x)\rangle = |\phi(x+l)\rangle$ . The general solutions to this equation are  $\phi(x) = e^{ikx}u(x)$  where  $u(x)$  satisfies  $u(x) = u(x+l)$  and  $k$  is complex.

This kind of translational symmetry exists when we have a crystal structure. Now consider a 1D closed loop of  $N$  atoms: Uniqueness of the wavefunction requires that  $\phi(x) = \phi(x + Nl) \Rightarrow e^{ikx} = e^{ik(x+Nl)}$ . Thus possible wavefunctions must have real  $k$  and the form

$$\phi(x) = e^{2\pi nix/Nl}u(x)$$

This is the 1D statement of Bloch's Theorem, the basis of study of electrons in solids. Notice how  $k$ , which we associate with the momentum, is a quantum number associated with translational symmetry, which in turn has an operator  $\hat{D}$  which commutes with the Hamiltonian and is thus a constant of the motion. Translational symmetry is associated with conservation of momentum.

## 4.5 The Kronig-Penney Model

In the above  $u(x)$  is still completely general. In the *Kronig-Penney model* we consider a periodically repeating square potential defined in one cell by  $V(x) = 0$  ( $0 < x < b$ );  $V(x) = V_0$  ( $b < x < l$ ), then we can solve for  $u(x)$  in one cell. Like the finite square well, this is a tedious boundary condition problem and leads to an equation the LHS of which is drawn below:

$$\cos k_1 b \cos k_2(l - b) - \frac{k_1^2 + k_2^2}{2k_1 k_2} \sin k_1 b \sin k_2(l - b) = \cos kl$$

where  $k_1 = \sqrt{2mE}/\hbar$  and  $k_2 = \sqrt{2m(E - V_0)}/\hbar$ , the appropriate free particle wavevectors, thus for  $E < V_0$ ,  $k_2$  is imaginary.

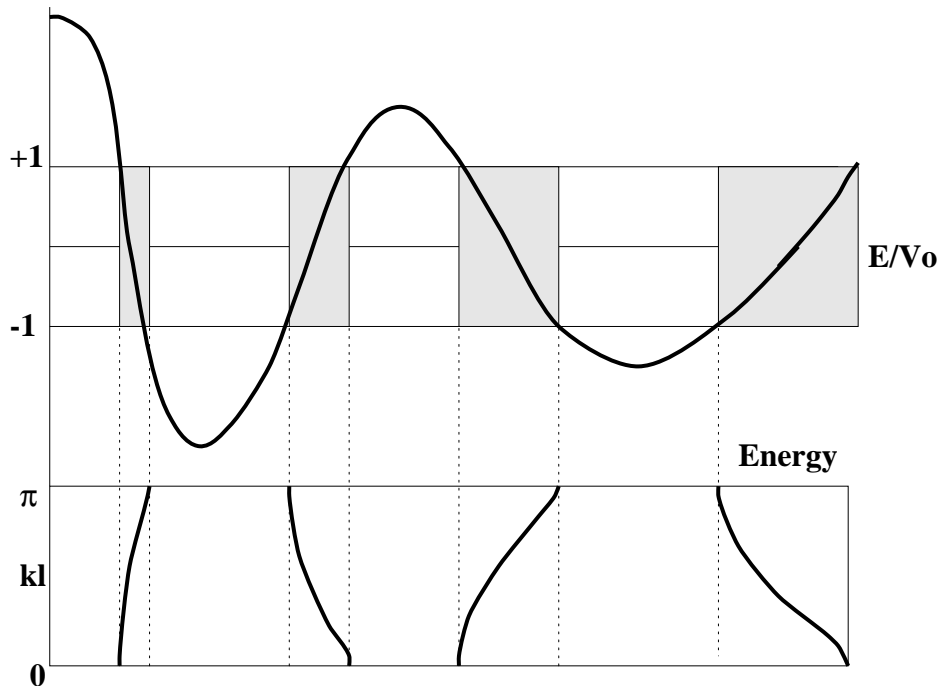


Figure 2: Graph of function arising from multiple square-well problem: Allowed energy solutions exist only where  $|\cos kl| \leq 1$  as is illustrated by the simple 'band structure' in the lower figure.

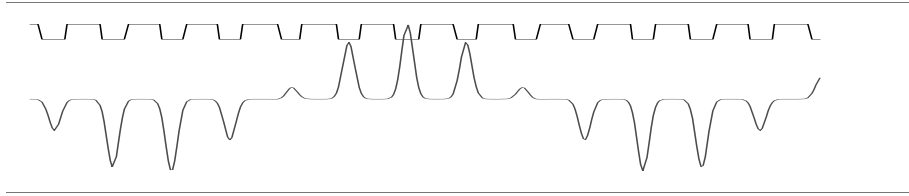


Figure 3: The Kronig-Penney potential and a Bloch function

The key point about this equation is that it cannot be solved for certain values of  $E$ , around  $k_1 b = m\pi$ . A plot of the left hand side of the equation against  $E/V_0$  illustrates this, solutions for some value of  $k$  can be found only in the shaded regions of  $E$ . Moreover each shaded region contains  $N$  allowed  $k = 2\pi n/Nl$  values. Thus if each atom contributes two electrons the lower ‘valence’ band will be filled (one of each spin in each state) and the upper ‘conduction’ band will be empty. To get an electron to move (change to a different  $k$ -state) requires a lot of energy, so this represents an insulator.

In three dimensions, the topology of the bands becomes much more complicated: this is a topic for solid state physics.

#### 4.6 Radioactive decay and imaginary potentials

If the number of particles in a given state is reduced in time, then the total intensity of that state is reduced. Consider a particle moving in a region of imaginary potential  $V(r) = -iV_0$ . The TDSE is:

$$i\hbar \frac{\partial}{\partial t} |\Phi, t\rangle = [H_0 - iV_0] |\Phi, t\rangle$$

Assume that the time independent part of the state is an combination of eigenstates of the real part of the Hamiltonian:

$$|\Phi, t\rangle = \sum_n c_n(t) \exp(-iE_n t/\hbar) |\phi_n\rangle; \quad \text{where} \quad H_0 |\phi_n\rangle = E_n |\phi_n\rangle$$

Following the same analysis as for TDSE, premultiplying by  $\langle m|$ , and for constant  $V_0$ ,  $V_{mn} = \delta_{mn} V_0$  we obtain:

$$i\hbar \dot{c}_m = -iV_0 c_m \quad \Rightarrow \quad |c_m(t)|^2 = |c_m(0)|^2 e^{-2V_0 t/\hbar}$$

Thus the probability amplitude of the state decreases in time. An imaginary potential can be used to represent destruction of particles, either by absorption (in a scattering process, perhaps) or by radioactive decay. Obviously the ket is not a full description of the system, since that should include information about the decay products. The lifetime of the state is  $\tau = \hbar/2V_0$ .

Notice that  $-iV_0$  is not a Hermitian operator, and so it is not possible to perform a single measurement of half life.

## 5 Time-dependence

### 5.1 Time-dependent Hamiltonians

Recall that for a system described by a Hamiltonian,  $\hat{H}_0$ , which is time-independent, the most general state of the system can be described by a wavefunction  $|\Psi, t\rangle$  which can be expanded in the energy eigenbasis  $\{|n\rangle\}$  as follows:

$$|\Psi, t\rangle = \sum_n c_n \exp(-iE_n t/\hbar) |n\rangle$$

where the coefficients,  $c_n$ , are time-independent, and  $E_n$  denotes the eigenvalue corresponding to the energy eigenstate  $|n\rangle$ .

When we generalise to the case where the Hamiltonian is of the form

$$\hat{H} = \hat{H}_0 + \hat{V}(t)$$

we can again expand in the eigenbasis of  $\hat{H}_0$

$$|\Psi, t\rangle = \sum_n c_n(t) \exp(-iE_n t/\hbar) |n\rangle$$

but the coefficients,  $c_n$ , will now in general be time-dependent.

The wavefunction satisfies the time-dependent Schrödinger equation;

$$i\hbar \frac{\partial}{\partial t} |\Psi, t\rangle = \hat{H} |\Psi, t\rangle$$

so that we can substitute the expansion of  $|\Psi, t\rangle$  to determine the equations satisfied by the coefficients  $c_n(t)$ . Writing  $E_n = \hbar\omega_n$  and denoting the time derivative of  $c_n$  by  $\dot{c}_n$  we obtain

$$i\hbar \sum_n (\dot{c}_n - i\omega_n c_n) \exp(-i\omega_n t) |n\rangle = \sum_n (c_n \hbar\omega_n + c_n \hat{V}) \exp(-i\omega_n t) |n\rangle$$

which simplifies immediately to give

$$\sum_n (i\hbar \dot{c}_n - c_n \hat{V}) \exp(-i\omega_n t) |n\rangle = 0$$

We now take the scalar product of this equation with  $\langle m|$  to give

$$i\hbar \dot{c}_m \exp(-i\omega_m t) - \sum_n c_n V_{mn} \exp(-i\omega_n t) = 0$$

giving the following set of coupled, first-order differential equations for the coefficients:

$$i\hbar \dot{c}_m = \sum_n c_n V_{mn} \exp(i\omega_{mn} t)$$

where  $\omega_{mn} = \omega_m - \omega_n$  and  $V_{mn} = \langle m|\hat{V}|n\rangle$ .

This tells us how the coefficient  $c_m$  varies with time, i.e. the probability that a measurement will show the system to be in the  $m^{\text{th}}$  eigenstate. It is exact, but not terribly useful because we must, in general, solve an infinite set of coupled differential equations.

It is worth dwelling on the importance of the quantity  $V_{mn}$ . This ‘matrix element’ is an integral which tells us how much the potential  $\hat{V}$  mixes states  $|m\rangle$  and  $|n\rangle$ . If it is zero (which it often is, by symmetry) then  $\hat{V}$  cannot induce a transition between states  $|m\rangle$  and  $|n\rangle$ .

## 5.2 Time-dependent Perturbation Theory

Consider the Hamiltonian

$$\hat{H} = \hat{H}_0 + \lambda \hat{V}(t)$$

and assume that we can expand the time dependent coefficients  $c_n$  in a power series in  $\lambda$ :

$$c_n = c_i + \lambda c_n^{(1)} + \lambda^2 c_n^{(2)} + \dots$$

Where  $c_i$  is the value of  $c_n$  at  $t=0$  ( $c_n^{(0)}$ ) We substitute in the equation for  $\dot{c}_m$  derived above, remembering to replace the arbitrary sized  $\hat{V}$  by the small  $\lambda \hat{V}$ , to give

$$\dot{c}_m^{(0)} + \lambda \dot{c}_m^{(1)} + \dots = (i\hbar)^{-1} \lambda \sum_n c_i V_{mn} \exp(i\omega_{mn}t) + \dots$$

We can now equate terms of the same degree in  $\lambda$ . The zeroth and first order terms give:

$$\begin{aligned} \dot{c}_m^{(0)} &= 0 \\ \dot{c}_m^{(1)} &= (i\hbar)^{-1} \sum_n c_i V_{mn} \exp(i\omega_{mn}t) \end{aligned}$$

The first equation is trivial; to zeroth order the coefficients are time-independent, since to this order the Hamiltonian is time-independent, and we recover the unperturbed result.

The second equation allows us to obtain the first-order contribution by integrating the first-order differential equation to give:

$$c_m^{(1)} = (i\hbar)^{-1} \sum_n c_i \int_0^t V_{mn} \exp(i\omega_{mn}t) dt$$

In the special case where the system is known to be in an eigenstate of  $\hat{H}_0$ , say  $|k\rangle$ , at  $t = 0$ , then  $c_k^{(0)} = 1$  and all other  $c_m^{(0)} = 0$ ,  $m \neq k$ , giving

$$c_m^{(1)} = (i\hbar)^{-1} \int_0^t V_{mk} \exp(i\omega_{mk}t) dt$$

The probability of finding the system at a later time,  $t$ , in the state  $|m\rangle$  where  $m \neq k$  is given by

$$p_m(t) = |c_m^{(1)}|^2$$

Since we have assumed a small perturbation, this result is only reliable if  $p_m(t) \ll 1$ .

## 5.3 Time-independent Perturbations

The results obtained in the last section can also be applied to the case where the perturbation,  $\hat{V}$ , is actually independent of time (strictly, 'switched on' at  $t=0$ ).

In this case we obtain

$$\begin{aligned} c_m^{(1)} &= (i\hbar)^{-1} V_{mk} \int_0^t \exp(i\omega_{mk}t) dt \\ &= \frac{V_{mk}}{\hbar\omega_{mk}} [1 - \exp(i\omega_{mk}t)] \end{aligned}$$

giving for the transition probability

$$p_m(t) = |c_m^{(1)}|^2 = \frac{|V_{mk}|^2}{\hbar^2} \frac{\sin^2(\omega_{mk}t/2)}{(\omega_{mk}/2)^2}.$$

For sufficiently large values of  $t$ , the function

$$f(t, \omega_{mk}) \equiv \frac{\sin^2(\omega_{mk}t/2)}{(\omega_{mk}/2)^2}$$

consists essentially of a large peak, centred on  $\omega_{mk} = 0$ , of height  $t^2$  and width  $\approx 4\pi/t$ , as indicated in Fig. 4. Thus there is only a significant transition probability if  $E_m \approx E_k$ . That is, if  $|\omega_{mk}| < 2\pi/t$ .

Note that we are assuming that the system was prepared in some eigenstate of  $\hat{H}_0$  which is not an eigenstate of  $\hat{V}$ : if it were, then the matrix element  $V_{nm}$  would be zero and  $p_m(t) = 0$ . Thus although the analysis treats the perturbation as time independent, it is applied to cases where the perturbation is switched on at  $t = 0$ . Moreover only perturbations which are incompatible with the Hamiltonian can induce transitions.

#### 5.4 Transitions to a group of states

We are often interested in the situation where transitions take place not to a single final state but to a group,  $G$ , of final states with energy in some range about the initial state energy

$$E_k - \Delta E \leq E_m \leq E_k + \Delta E$$

Then the total transition probability is obtained by summing the contributions of all the final states. The number of final states in the interval between  $E_m$  and  $E_m + dE_m$  is  $g(E_m) dE_m$ , where the function  $g(E_m)$  is known as the *density of final states*. The total transition probability for transitions to  $G$  is then given by

$$p_G(t) = \frac{1}{\hbar^2} \int_{E_k - \Delta E}^{E_k + \Delta E} |V_{mk}|^2 f(t, \omega_{mk}) g(E_m) dE_m.$$

For sufficiently large  $t$ , and  $\Delta E \gg 2\pi\hbar/t$ , we observe that essentially the only contributions to the integral come from the energy range corresponding to the narrow central peak of the function  $f(t, \omega_{mk})$ . Within this range we can neglect the variation of  $g(E_m)$  and  $V_{mk}$ , which can therefore be taken out of the integral to give

$$p_G(t) = \left[ \frac{|V_{mk}|^2}{\hbar^2} g(E_m) \right]_{E_m=E_k} \int_{E_k - \Delta E}^{E_k + \Delta E} f(t, \omega_{mk}) dE_m.$$

Furthermore, we can extend the limits on the integration to  $\pm\infty$ . Noting that  $dE_m = \hbar d\omega_{mk}$  and using the result that

$$\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} dx = \pi$$

we obtain for the first-order transition probability

$$p_G(t) = \frac{2\pi t}{\hbar} \left[ |V_{mk}|^2 g(E_m) \right]_{E_m=E_k}$$

The transition rate,  $R$ , is just the derivative of this with respect to  $t$  and is thus given by the so-called Fermi Golden Rule:

$$R = \frac{2\pi}{\hbar} \left[ |V_{mk}|^2 g(E_m) \right]_{E_m=E_k}$$

The Fermi Golden Rule is probably the single most widely used result in quantum mechanics. The factor of  $\frac{2\pi}{\hbar}$  depends on the choice of perturbing potential, but the  $|V_{n1}|^2 g(E_m)$  term appears for any applied perturbation. Be careful about the density of energy states - one sometimes encounters density of frequency states (which differs by a factor of  $\hbar$ ) or of wavevector states.

It may appear that need to know the density of final states,  $g(E_m)$ , but this is not always true. In cases where  $|V_{mk}| = 0$  transitions are forbidden, and in some cases we can deduce  $g(E_m)$  from the relative rates of related transitions.

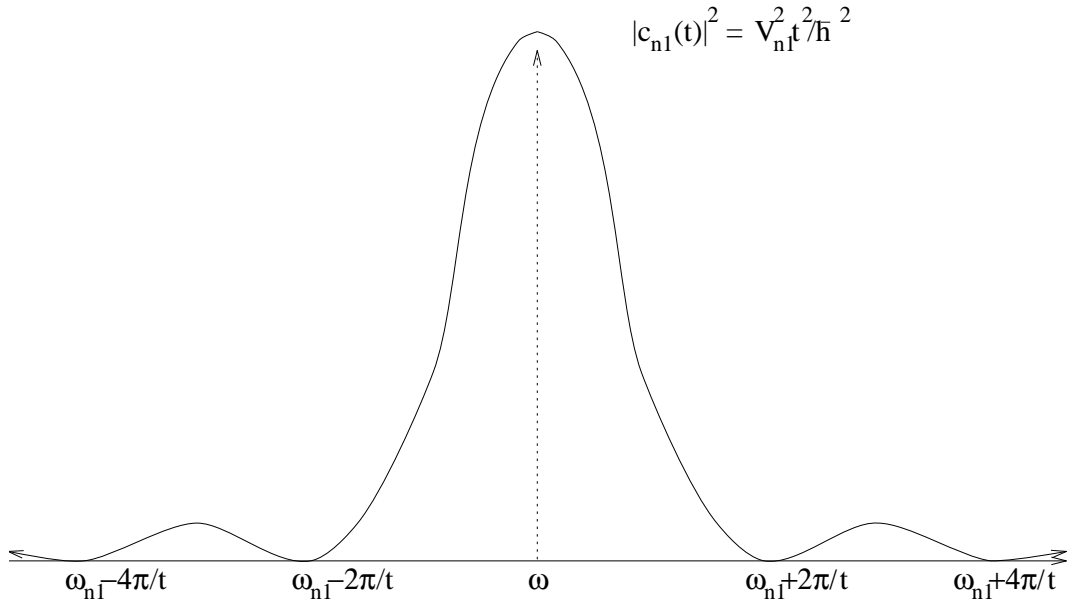


Figure 4: Transition probability as a function of applied harmonic perturbation frequency

## 5.5 Harmonic Perturbation

This is generally useful since by Fourier analysis we can decompose any periodic perturbation into harmonic components.

Let the perturbing potential be  $V(\mathbf{r}, t) = V(\mathbf{r}) \cos \omega t$

If the initial state at  $t = 0$  is  $k$ , and the final state  $m$  then

$$c_m \approx \frac{-i}{\hbar} V_{mk} \int_0^t e^{-i\omega_{mk}t} \frac{1}{2} (e^{i\omega t} + e^{-i\omega t}) dt = \frac{V_{mk}}{2\hbar} \left( \frac{e^{i(\omega_{mk}-\omega)t} - 1}{\omega_{mk} - \omega} + \frac{e^{-i(\omega_{mk}+\omega)t} - 1}{\omega_{mk} + \omega} \right)$$

where  $V_{mk}$  is the time independent part of the matrix element  $\langle m | \hat{V} | k \rangle$ . This function is dominated by the first term in the region around  $\omega_{mk} = \omega$ , so we can consider only the first term to obtain an estimate for the transition probability:

$$|c_m(t)|^2 = \frac{V_{mk}^2 \sin^2[(\omega_{mk} - \omega)t/2]}{\hbar^2 (\omega_{mk} - \omega)^2} = \frac{4}{\hbar^2} V_{mk}^2 f(t, \omega_{mk} - \omega)$$

Where the function  $f$  is the same as we encountered earlier. Thus an external perturbation at a given frequency most strongly induces transitions between energy levels separated by  $\hbar\omega$ .

This is another manifestation of an uncertainty principle. If the potential is electromagnetic, the most probable transition is the absorption of a  $\hbar\omega_{mk}$  photon as the system changes energy by  $\hbar\omega_{mk}$ . But if the transition happens very fast, the peak is broad and the photon could have a wide range of energies, contrariwise, if the transition occurs after a long time the photon frequency is well defined:  $\Delta E \Delta t \geq \hbar/2$ . This uncertainty gives rise to the ‘natural linewidth’ of a particular transition, and causes a limit to the accuracy of certain experiments. There is a slight difference from the Heisenberg Uncertainty in non-relativistic quantum mechanics because time is not an operator so one cannot define the commutator of time with the Hamiltonian.

Note the extraordinary result that the transition probability at small times is  $(4V_{mk}^2/\hbar^2) t^2$ . Consider what happens if the state is measured frequently compared to if measurements are made infrequently: frequent measurement tends to inhibit the transition!

## 6 Two state systems

### 6.1 Time Dependence

The exact expression for the time dependence of a system required a set of as many simultaneous differential equations as there are states in the system. One case where we can solve this time dependent problem exactly is when we have a small number of basis states. Consider, for example, a system which requires only two basis states. Say we prepare it in initial state  $|1\rangle$  and we want to know how long it will take to go to the other state  $|2\rangle$ . We have two coupled equations in the time dependent  $c_1$  and  $c_2$ :

$$\begin{aligned}i\hbar\dot{c}_1 &= V_{11}c_1 + V_{12}c_2e^{i\omega_{12}t} \\i\hbar\dot{c}_2 &= V_{22}c_2 + V_{21}c_1e^{i\omega_{21}t}\end{aligned}$$

where  $c_1(0) = 1$  and  $c_2(0) = 0$ .

If the change is slow, we can use first order time-dependent perturbation theory. In this we ignore the possibility that the transition has already occurred when we calculate the transition rate  $\dot{c}_1$ . We thus replace the  $c_n(t)$  by  $c_n(0)$ , and integrate whence:

$$\begin{aligned}c_1 &\approx \exp(iV_{11}t/\hbar) \\|c_1|^2 &\approx 1 \\c_2 &\approx \frac{-i}{\hbar} \int_0^t V_{21}e^{i\omega_{21}t} dt\end{aligned}$$

Including the constant of integration for  $c_1(0) = 1$ .

### 6.2 Notes

- The ‘Matrix element’  $V_{21}$  determines whether there is a transition from an initial state 1 to a final state 2 even if  $\hat{V}$  is *independent* of time. It also determines the rate of the transition.
- If the states  $|1\rangle$  and  $|2\rangle$  are eigenstates of the perturbation  $\hat{V}$  then  $V_{21} = V_{12} = 0$  and no transition occurs.
- Over a long period of time, the system will oscillate between the two states.
- The mathematics is the same as for two coupled pendula, where the energy moves back and forth between the two bobs.
- The states can represent *anything*, and oscillation will occur whenever there are off diagonal terms in the matrix.
- Examples: (see Feynman III Ch.9-11) Nitrogen atom in ammonia, electron in  $\text{H}_2^+$ , pion exchange, benzene, electron spins, photon polarisation, neutrino oscillations, neutral kaons.

### 6.3 Oscillation in a two state system

Consider the expectation value of a quantity  $S$  in a system which has two non-degenerate energy eigenstates  $|1\rangle$  and  $|2\rangle$ , and where the Hermitian operator  $\hat{S}$  is defined by  $\hat{S}|1\rangle = |2\rangle$ ,  $\hat{S}|2\rangle = |1\rangle$ .

The general state can be written:

$$|\phi\rangle = c_1 \exp(-iE_1t/\hbar)|1\rangle + c_2 \exp(-iE_2t/\hbar)|2\rangle$$

from which it follows that the expectation value  $\langle \hat{S} \rangle$  will be:

$$\begin{aligned} \langle \hat{S} \rangle &= \langle \phi | \hat{S} | \phi \rangle \\ &= [c_1 e^{iE_1t/\hbar} \langle 1 | + c_2 e^{iE_2t/\hbar} \langle 2 |] [c_1 e^{-iE_1t/\hbar} |2\rangle + c_2 e^{-iE_2t/\hbar} |1\rangle] \\ &= c_1 c_2 [e^{i\omega_{21}t} + e^{-i\omega_{21}t}] \\ &= 2c_1 c_2 \cos(\omega_{21}t) \end{aligned}$$

Thus the expectation value of  $\hat{S}$  oscillates in time at frequency  $\omega_{21} = (E_2 - E_1)/\hbar$ . This arises because  $\hat{S}$  is not compatible with the hamiltonian, and hence does not define a constant of the motion.

### 6.4 Example of oscillation in a two state system - the neutral kaons

The phenomenon of expectation value oscillation is actually seen in neutral kaons, although the picture is further complicated by radioactive decay. Full description of fundamental particles is beyond the scope of this course, but for this example all you need to know about kaons are these three points:

- Neutral kaons are particles which interact via the strong and weak nuclear forces.
- There are two of them: particle and antiparticle.
- The weak and strong forces can be represented as quantum operators.

Neutral kaons are produced by pion decay via the strong interaction. Consequently, kaons are produced in eigenstates of the strong interaction operator  $\hat{S}$  (strangeness). These eigenstates are called  $|K^0\rangle$  ( $S=1$ ) and  $|\bar{K}^0\rangle$  ( $S=-1$ ).

Neutral kaons decay via the weak interaction. The operator for this is called  $\hat{C}\hat{P}$ . It is a combination of charge conjugation ( $\hat{C}$ ), which reverses the sign of the charge and the magnetic moment of a particle, and parity ( $\hat{P}$ ) which inverts space  $\mathbf{r} \rightarrow -\mathbf{r}$ .

$\hat{S}$  and  $\hat{C}\hat{P}$  are not compatible operators. Acting on  $|K^0\rangle$  and  $|\bar{K}^0\rangle$  with  $\hat{C}\hat{P}$  gives:

$$\hat{C}\hat{P}|K^0\rangle = |\bar{K}^0\rangle; \quad \hat{C}\hat{P}|\bar{K}^0\rangle = |K^0\rangle$$

Clearly  $|K^0\rangle$  and  $|\bar{K}^0\rangle$  are not eigenstates of CP. When looking at kaon decay, we need a different representation of the kaons; the appropriate eigenstates are:

$$|K_1\rangle = [ |K^0\rangle + |\bar{K}^0\rangle ] / \sqrt{2}; \quad |K_2\rangle = [ |K^0\rangle - |\bar{K}^0\rangle ] / \sqrt{2}$$

Now it turns out that the decay of  $|K_1\rangle$  is 800 times faster than  $|K_2\rangle$ , so for practical purposes, we could ignore the decay of  $|K_2\rangle$  so as time passes with kaons in vacuo, soon only state  $|K_2\rangle$  remains. In their rest frame, the  $|K_1\rangle$  kaons have energy  $m_1c^2$ . From TDSE, we can write the time dependence of the amplitude  $a_1$  of the  $|K_1\rangle$  state, and similarly the amplitude  $a_2$  of the  $|K_2\rangle$  state, as:

$$a_1(t) = a_1(0)e^{-t/2\tau_1}e^{-im_1c^2t/\hbar}; \quad a_2(t) = a_2(0)e^{-im_2c^2t/\hbar}$$

This corresponds to the radioactive decay law with decay constant  $\tau$  for  $|K_1\rangle$  and no decay for  $|K_2\rangle$ . - recall 4.6 how radioactive decay can be represented by an imaginary potential in  $\hat{H}$ :

$$\text{Intensity } I = |a(t)|^2 = |a(0)|^2e^{-t/\tau}$$

Now, the expectation value of the strangeness of a beam of kaons is:

$$\langle \hat{S} \rangle = [a_1(t)\langle K_1| + a_2(t)\langle K_2|] [a_1^*(t)|K_2\rangle + a_2^*(t)|K_1\rangle]$$

If the initial beam were prepared in state  $|K_0\rangle$  then  $a_1(0) = a_2(0) = 1/\sqrt{2}$ , the strangeness oscillates in time while decaying to zero.

$$S = I(K^0) - I(\bar{K}^0) = e^{-t/2\tau_1} \cos(m_{12}t)$$

defining  $m_{12} = (m_2 - m_1)c^2/\hbar$ . This can actually be measured, though for details you'll have to go to the particle physics course.

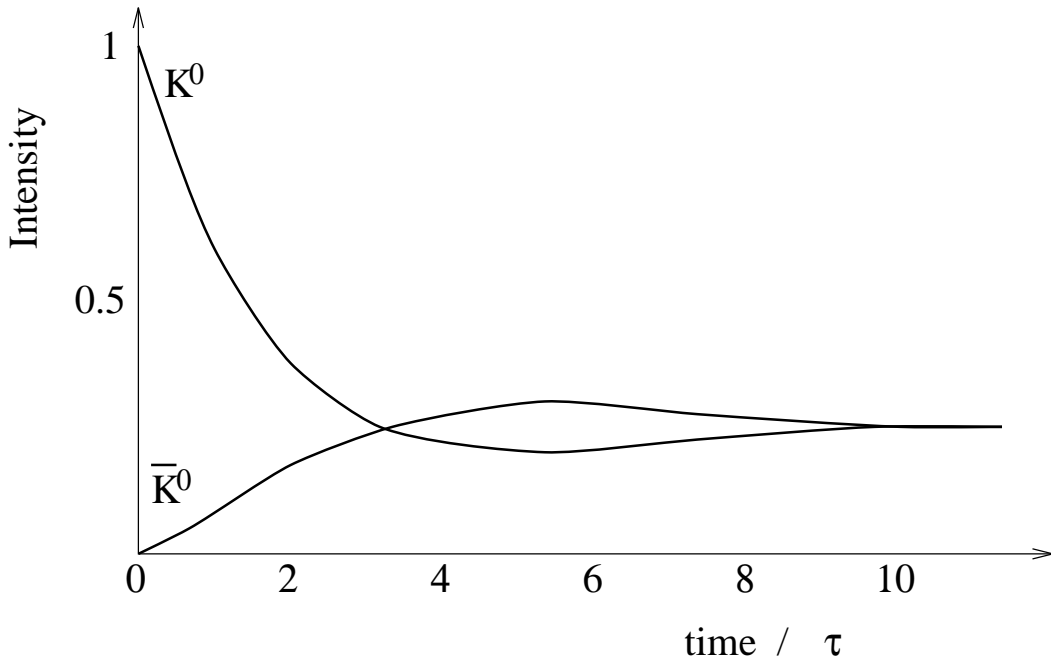


Figure 5: The variation in intensity of  $K^0$  and  $\bar{K}^0$  as decay proceeds via the weak interaction for which they are not eigenstates. This assumes  $m_{12}\tau_1 \approx 0.5$ .

## 6.5 Regeneration through wavefunction collapse

If the beam of  $|K_2\rangle$  kaons passes through some matter, they interact via the strong force. This effectively measures  $S$  for each kaon and the wavefunction is collapsed from  $|K_2\rangle$  back onto pure  $|K_0\rangle$  or pure  $|\bar{K}_0\rangle$ . In either case, some component of  $|K_1\rangle$  is *regenerated*, and on leaving the matter the decay process via  $\hat{C}P$  can start again.

Although it has no classical analogue, this regeneration is mathematically similar to the regeneration of spin  $S_z = \frac{1}{2}$  electrons from a polarised  $S_z = -\frac{1}{2}$  beam by passing through a  $S_x$ -oriented Stern-Gerlach apparatus.

There is also some similarity to a system of three polarisers, the middle one at  $45^\circ$  to the other two. In that case inserting the middle polariser reduces the intensity by a factor of four: halving the intensity itself and causing a further halving by the final polariser. Here the interposition of matter does not reduce the beam intensity, but it allows for a subsequent halving by  $|K_1\rangle$  decay.

Exactly the same analysis can be done for *neutrinos*. An electron neutrino turns into a muon neutrino and back at a rate proportional to their mass difference (but not if they have no difference). No direct measurement of neutrino mass has been made, but one explanation of there being fewer atmospheric neutrinos than expected detected after passing through the earth is that they change from detectable muon neutrinos into undetectable tau neutrinos. If this were true, the experiment suggests a mass difference of about  $10^{-37}$ kg.

## 6.6 Strong force - Two state system or degenerate perturbation

The fundamental forces can be thought of as manifestations of two state systems. Consider a system comprising a proton and a neutron. The proton can decay into a neutron plus a pion, while the neutron can absorb the pion and become a proton. We can think of the system as two neutrons and a pion: the pion having two degenerate states  $|a\rangle$  and  $|b\rangle$  depending on which neutron it is located. The off-diagonal terms are now  $\langle a|\hat{V}_a|b\rangle$ , where  $V_a$  is the potential energy of the pion due to neutron  $a$ . The two state analysis shows that we can think of the pion hopping back and forth between the neutrons (the pion exchange mechanism). Or we can treat the system by degenerate perturbation theory and diagonalise the 2x2 matrix to find energies:  $V_{aa} \pm V_{ab}$ . The true ground state energy is thus  $E_g = V_{aa} - V_{ab}$ .

Note that  $V_{ab}$  is the overlap between the state with the pion on one site and the state with the pion on the other site. Obviously this depends on the separation ( $R$ ), and so there is a force between the neutrons  $dE_g/dR$ . Less obviously, the force depends on the tails of the wavefunctions, which in turn are exponentially dependent on the pion mass. Thus the strength of the strong force falls off exponentially with distance.

Note also that we have described the basis states of our two state system as ‘a proton and a neutron’, but the actual ground state is a mixture of the two. When interacting via the strong force, the nucleons lose their well defined identity.

This picture of forces arising from exchange of ‘virtual’ particles (the pion is not observed as a free particle here) is the standard way of thinking about fundamental forces - the electromagnetics force involves ‘exchange of virtual photons’, the gravitational force ‘gravitons’ etc. These force are long ranged (not exponentially decaying) because the particles involved have zero mass.

All of this is analogous to covalent bonding: ‘exchange of electrons’: and in each case there is still a level of understanding lurking beneath: QED (photons) for electron-ion bonding and QCD (gluons and quarks) for nucleon binding.

## 7 The Variational Principle

### 7.1 Approximate solution of the Schroedinger equation

If we can't find an analytic solution to the Schroedinger equation, a trick known as the variational principle allows us to estimate the energy of the ground state of a system. We choose some trial function  $\phi(a_n)$  which depends on some *variational parameters*,  $a_n$  and minimise

$$E[a_n] = \frac{\langle \phi | \hat{H} | \phi \rangle}{\langle \phi | \phi \rangle}$$

with respect to those parameters. This gives an approximation to the wavefunction whose accuracy depends on the number of parameters and the clever choice of  $\phi(a_n)$ . For more rigorous treatments, a set of basis functions may be used.

The proof is as follows, if we expand a normalised guess wavefunction  $|\phi(a_n)\rangle$  in terms of the true (unknown) eigenbasis  $|i\rangle$  of the Hamiltonian:

$$E_{\phi(a_n)} = \sum_{ij} \langle \phi | i \rangle \langle i | \hat{H} | j \rangle \langle j | \phi \rangle = \sum_i |\langle \phi | i \rangle|^2 E_i$$

and define the true ground state of the system by  $\hat{H}|i_0\rangle = E_0|i_0\rangle$ , then we can write:

$$\text{Min}(E[a_i]) = \sum_i |\langle \phi | i \rangle|^2 E_i = E_0 + \sum_i |\langle \phi | i \rangle|^2 (E_i - E_0) \geq E_0$$

Thus the lower we can make the energy  $E[a_i]$ , the closer  $|\phi\rangle$  will be to the actual ground state wavefunction  $|i_0\rangle$ . If the trial wavefunction consists of a complete basis set of orthonormal functions  $|\chi_i\rangle$ , each multiplied by  $a_i$ :  $|\phi\rangle = \sum_i a_i |\chi_i\rangle$  then the solution is exact and we just have the usual trick of expanding a wavefunction in a basis set. Alternately, since complete basis sets are often infinite, we might just use a few basis functions to get an answer close to the ground state  $|i_0\rangle$ . In practice, this is how most quantum mechanics problems are solved when an analytic solution cannot be found.

Note that in principle this is valid only for the ground state. However, once the ground state is known it may be possible to find the first excited state by minimising  $E[a_i]$  subject to the constraint  $\langle \phi | \phi_0 \rangle = 0$ . However, the error in this approach will be larger than for  $E_0$  because not only is the basis incomplete, but also  $|i_0\rangle$  is not known exactly so that the constraint  $\langle \phi | \phi_0 \rangle = 0$  may not be exactly satisfied (sometimes symmetry will suffice). For this reason, the variational method is usually used only for ground state calculations.

### 7.2 Kinetic Energy

The expectation value of the kinetic energy  $\langle \hat{T} \rangle$  is always positive. This can be shown by an integration by parts in which the first term vanishes provided the wavefunction tends to zero at infinity (which it will for a bound state). In 1D:

$$\langle \hat{T} \rangle = \frac{-\hbar^2}{2m} \int \Phi^* \frac{d^2}{dx^2} \Phi dx = \frac{-\hbar^2}{2m} [\Phi^* \frac{d}{dx} \Phi]_{-\infty}^{\infty} + \frac{\hbar^2}{2m} \int \frac{d}{dx} \Phi^* \frac{d}{dx} \Phi dx = \frac{\hbar^2}{2m} \int \left| \frac{d}{dx} \Phi \right|^2 dx$$

The second term integrand is positive everywhere, so the kinetic energy is always positive.

### 7.3 Analytic example of variational method - Binding of the deuteron

Say we want to solve the problem of a particle in a potential  $V(r) = -Ae^{-r/a}$ . This is a model for the binding energy of a deuteron due to the strong nuclear force, with  $A=32\text{MeV}$  and  $a=2.2\text{fm}$ . The strong nuclear force does not exactly have the form  $V(r) = -Ae^{-r/a}$ , unlike the Coulomb interaction we don't know what the exact form should be, but  $V(r) = -Ae^{-r/a}$  is a reasonable model.

The potential is spherically symmetric, most attractive at  $r = 0$  and falls rapidly to zero at large  $r$ , so we choose a trial wavefunction which does the same, say  $\phi = ce^{-\alpha|r|/2a}$ . This has only one dimensionless variational parameter,  $\alpha$ . The value of  $c$  follows from normalisation  $\int c^2 e^{-\alpha r/a} 4\pi r^2 dr = 1$ ; which gives  $c^2 = \alpha^3/8\pi a^3$ . (The  $4\pi r^2$  comes from the problem being three dimensional).

According to the variational principle, our best estimate for the ground state using this trial function comes from minimising  $\langle \phi | \hat{H} | \phi \rangle$  with respect to  $\alpha$ . Using the kinetic energy expression from the previous section:

$$\begin{aligned} \langle \phi | H | \phi \rangle / \langle \phi | \phi \rangle &= \frac{\hbar^2}{2m} \int_0^\infty c^2 \frac{\alpha^2}{4a^2} \exp(-\alpha r/a) 4\pi r^2 dr - A \int_0^\infty c^2 \exp[-(\alpha + 1)r/a] 4\pi r^2 dr \\ &= \frac{\hbar^2 \alpha^2}{8ma^2} - A \left( \frac{\alpha}{\alpha + 1} \right)^3 \end{aligned}$$

From this we find the minimum for  $E(\alpha)$  at  $\alpha_0$

$$\frac{dE}{d\alpha} = \frac{\hbar^2 \alpha}{4ma^2} - 3A \left( \frac{\alpha^2}{(\alpha + 1)^4} \right) = 0 \quad \implies \quad \frac{(\alpha_0 + 1)^4}{\alpha_0} = 12Ama^2/\hbar^2 = 22.3$$

Solving for  $\alpha_0$  gives  $\alpha_0 = 1.34$ , and substituting back into  $\langle \phi | H | \phi \rangle$  gives  $E_0 = -2.14\text{MeV}$ .

This is fairly close to the exact solution for this potential, which can be obtained analytically as a Bessel function of  $\sqrt{8mA}(a/\hbar)e^{-r/2a}$  if you manage to spot that change of variables! The exact solution gives  $E_0 = -2.245\text{MeV}$ .

### 7.4 Lattice methods: Variational method by computer

The variational method transposes the problem of solving a differential equation onto the problem of minimising a function of many variables. It is therefore good for use with computers.

One of the simplest ways of solving for the ground state of a system with a computer is to use as a basis set consisting of the values of  $|\phi\rangle$  defined on a lattice. In 1D such a solution is simply a histogram where we adjust the wavefunction at each point until the energy of the whole system is minimised. The kinetic energy (second derivative of the wavefunction) must then be obtained by some interpolation method. The weights of  $|\phi\rangle$  at each point can be regarded as a basis set. It is not complete, but it becomes more and more complete as the lattice gets finer.

Another common way of solving the Schroedinger equation numerically is to write the wavefunction as a Fourier series in a confined space of length  $L$ .

$$\Phi(x) = \sum_k a_k e^{ik\pi x/L}$$

where  $a_k$  are the variational parameters. Using Fourier series (also called plane wave expansions) has several advantages. Increasing accuracy can be obtained by adding more Fourier components (because each plane wave is orthogonal to all the others), the value of  $\Phi(x)$  can be quickly found by a Fourier transform of  $a(k)$  and the kinetic energy has particularly simple form:

$$\frac{d^2\Phi}{dx^2} = - \sum_k a_k \frac{k^2 \pi^2}{L^2} e^{ik\pi x/L}$$

which requires no numerical differentiation if used on a computer. Note also that this method is exactly the same as expanding the wavefunction in a basis set.

## 7.5 The density-functional pseudopotential plane wave method

This is probably the most commonly used method for quantum mechanical study in solid state physics. The problem is to solve the Schroedinger equation for a collection of interacting electrons. The density functional theorem, proved in the 1960s, showed that the problem can be transformed to that of a system of non-interacting ‘quasiparticles’, with the cost that the Hamiltonian depends on the electron density  $\rho(\mathbf{r})$ :

$$H[\rho(\mathbf{r})]\phi_i = E_i\phi_i \quad \text{where} \quad \rho(\mathbf{r}) = \sum_i |\phi_i(\mathbf{r})|^2$$

Thus the Schroedinger equation is a nonlinear differential equation of many variables. And we don’t know how to solve them. Thus we must turn to the variational method. In a solid, we also have no idea what the wavefunction should look like. The most general approach here is to use a Fourier Series (plane wave expansion). The wavefunction for the  $i$ th electron is then written as

$$\phi_i = \sum_k c_{ik} \exp(-i\mathbf{k}\cdot\mathbf{r}) \quad \text{and the variational equation becomes :} \quad E_0 = \text{Min} \sum_i \langle \phi_i | \hat{H}(\rho) | \phi_i \rangle$$

It is thus possible to solve for the ground state of the electrons to an accuracy determined by the number of Fourier components used. Once this is known most structural properties of the material can be calculated.

The overall scheme is simple: the wavefunctions are expanded in a computer-friendly basis set and the variational principle is used to transform the problem from a set coupled non-linear differential equations into a minimisation of a single function of many variables.

A drawback to this scheme is that electron wavefunctions don’t actually look much like plane waves, so the basis set is very different from the wavefunctions, and very many Fourier components are required. One way around this is to use a ‘pseudopotential’ which attempts to describe the potential due to the nucleus and tightly bound shells of ‘core’ electrons which do not take part in bonding. In silicon for example the pseudopotential describes the nucleus and the  $1s2s2p$  shells of electrons. An interesting thing about pseudopotentials is that they are different for  $s$  and  $d$  electrons (because they must include the fact that  $3s$  must be radially orthogonal to  $1s$  and  $2s$ , while  $3d$  are automatically so because of the angular dependence). This is called non-locality.

(For details see <http://www.cse.clrc.ac.uk/Activity/UKCP+994>)

## 7.6 Density functional theory and the Kohn-Sham functional

If we consider the total probability density of a system of many interacting particles  $\rho(\mathbf{r})$ , there may be several possible wavefunctions which could give rise to it: call this set  $S(\Phi)$ .

Now, consider the expectation value of the energy  $\langle \hat{H} \rangle$ . We know from the variational principle that  $\langle \hat{H} \rangle \geq E_o$ . If we define a functional  $F[\rho(\mathbf{r})] = \text{Min}_{S(\Phi)} \langle \hat{H} \rangle$ , then it follows that  $F[\rho] \geq E_o$ .

Consequently we can use the variational principle to find the  $\rho(\mathbf{r})$  which minimises the value of  $F$ , and this may give us the ground state energy *without having to evaluate the wavefunction*. This is especially useful when the wavefunction consists of complex combinations of many different single-particle wavefunctions, as with the many electrons in a solid or molecule.

The drawback is that the functional may not be known. A popular form of functional was introduced by Kohn and Sham:

$$F(\rho) = T[\rho] + \frac{1}{2} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|} + E_{xc}[\rho] + \sum_i \int \frac{Z e \rho(\mathbf{r}')}{4\pi\epsilon_0|\mathbf{R}_i-\mathbf{r}'|}$$

where  $T$  is the kinetic energy of a non-interacting electron gas, the integrals represent electrostatic interactions between the electrons and between electrons and ions, and  $E_{xc}$  is ‘everything else’. The advantage of this form is that it can be recast to give a set of one-particle equations with non-interacting fermions moving in an effective potential:

$$V_{eff} = \sum_{ion} \frac{Ze}{4\pi\epsilon_0|\mathbf{R}_{ion} - \mathbf{r}'|} + \int \frac{\rho(\mathbf{r}')}{4\pi\epsilon_0|\mathbf{r} - \mathbf{r}'|} + \frac{\delta E_{xc}[\rho]}{\delta\rho(\mathbf{r})}$$

Since  $V_{eff}$  depends on  $\rho(\mathbf{r})$  these equations must be solved self-consistently.

## 7.7 Quantum forces: the Hellmann-Feynman Theorem

For molecules and solids one is often interested in *forces* as well as energies. If we can write the energy as of a many electron system in state  $\phi$ :  $E = \langle\phi|\hat{H}|\phi\rangle$  and differentiate with respect to some quantity  $\alpha$  then

$$\frac{dE}{d\alpha} = \left\langle\frac{d\phi}{d\alpha}|\hat{H}|\phi\right\rangle + \langle\phi|\frac{d\hat{H}}{d\alpha}|\phi\rangle + \langle\phi|\hat{H}|\frac{d\phi}{d\alpha}\rangle$$

But since  $\hat{H}|\phi\rangle = E|\phi\rangle$  and  $\langle\phi|\phi\rangle$  is 1 for normalisation:

$$\frac{dE}{d\alpha} = \langle\phi|\frac{d\hat{H}}{d\alpha}|\phi\rangle + E\frac{d}{d\alpha}\langle\phi|\phi\rangle = \langle\phi|\frac{d\hat{H}}{d\alpha}|\phi\rangle$$

This result is the Hellmann-Feynman theorem. If  $\alpha$  represents the position of a nucleus in a solid, then the force on that nucleus is the expectation value of the force operator  $\frac{d\hat{H}}{d\alpha}$ . It can be applied to any quantity which is a differential of the Hamiltonian provided the basis set does not change.

## 7.8 Unresolved issues

The exact form of  $E_{xc}$  and the meaning of the eigenfunctions of the Kohn-Sham Hamiltonian are still not agreed upon. For solid-state systems a ‘local density approximation’ for  $E_{xc}$  is remarkably good. This just says that the total  $E_{xc}[\rho(\mathbf{r})]$  can be found by integrating the contribution at each point in space as if it were equivalent to  $E_{xc}$  for a uniform density (which can be calculated).

At any given time the electrons tend to stay out of each others way more than the averaged probability density might suggest. Calculating the energy associated with this effect is still a major research effort. The flavour of the problem can be seen by comparing two models for the spatial wavefunction of the hydrogen molecule. Since we are dealing with Fermions, the wavefunction must be antisymmetric with respect to swapping the electron labels, the spin part is antisymmetric so the spatial part is symmetric (See Atomic and Molecular course). Using functions  $\phi_1$  and  $\phi_2$  centred on atoms 1 and 2 respectively. The ‘strongly correlated’ Heitler London model assumes the electrons are on opposite atoms.

$$\Phi_{HL} = \phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2) + \phi_1(\mathbf{r}_2)\phi_2(\mathbf{r}_1)$$

While the ‘uncorrelated’ Hartree Fock model includes ionic configurations:

$$\Phi_{HF} = \phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2) + \phi_1(\mathbf{r}_2)\phi_2(\mathbf{r}_1) + \phi_1(\mathbf{r}_1)\phi_1(\mathbf{r}_2) + \phi_2(\mathbf{r}_2)\phi_2(\mathbf{r}_1)$$

And the truth lies somewhere in between.

## 7.9 Localised Basis Sets

An alternative to the plane wave method of 7.6 is to use a basis set comprising atom-centred functions (e.g. atomic orbitals and spherical harmonics). The two approaches mimic the free electron vs ionic pictures of condensed matter.

The advantage of localised bases is that far fewer basis functions are needed to get a good description of the wavefunction. The disadvantages are that basis functions centred on adjacent atoms are not orthogonal, and that if one wants to move an atom, the basis set has to move as well. This means that the Hellmann-Feynman theorem does not apply.

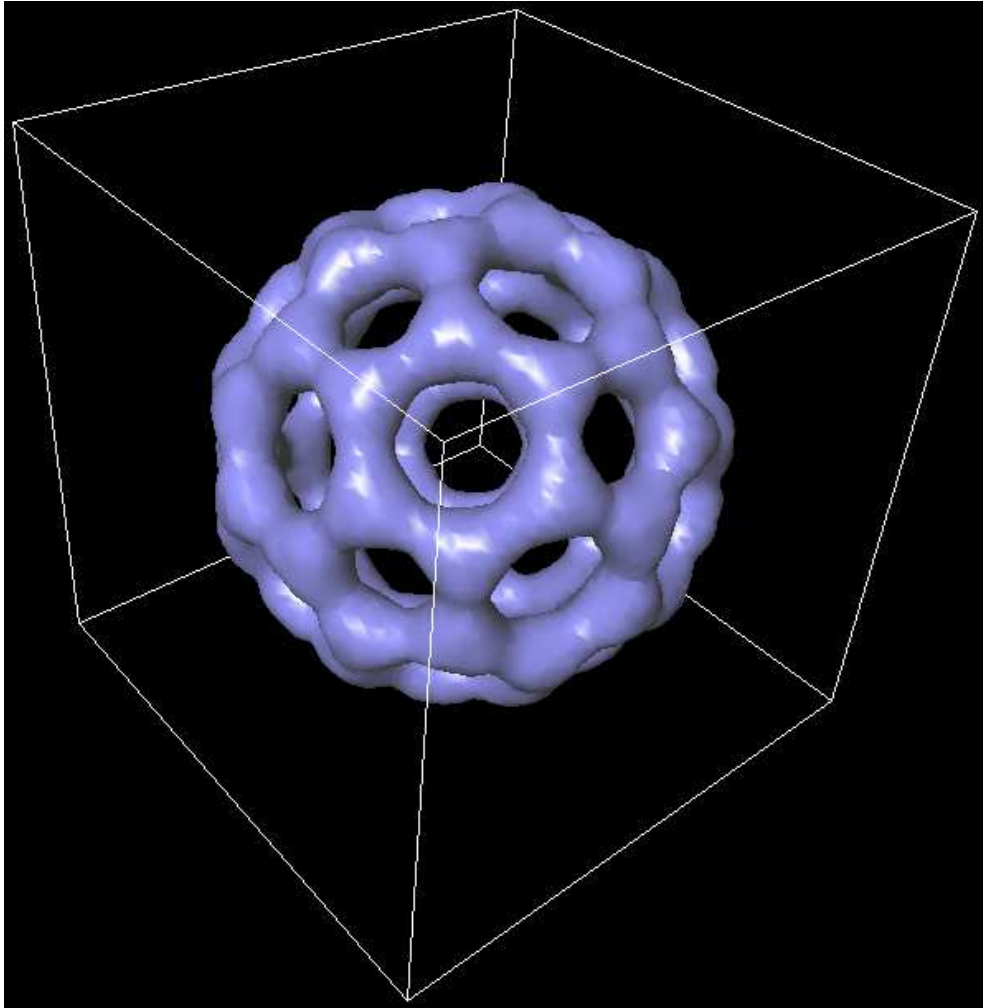


Figure 6: The electron probability density of a buckyball ( $C_{60}$  molecule), illustrating the complexity of problem for which the Schroedinger equation can be solved numerically.

## 8 Fundamentals of Quantum Scattering Theory

### 8.1 Centre of Mass Frame and the Two-body Problem

The problem of a particle in a given potential can be solved classically from Newton's equations. The Schroedinger equation can be used to describe the behaviour of one particle in a field.

The problem of two particles interacting via conservative fields can be reformulated into two parts: the behaviour of the centre of mass and the behaviour of the relative velocities of the particles. If we work in the centre of mass frame (COM), then the behaviour of the centre of mass is trivial, and we need worry only about the relative motions. This can be described by a *single* effective particle with effective mass  $\mu = \frac{m_1 m_2}{m_1 + m_2}$ . This effective particle can then be treated with one particle equations.

The problem of three interacting particles cannot be reduced in this way. Hence the 'three-body-problem' is in general insoluble.

The COM transformation allows us to treat the scattering problem as a one body problem. For scattering problems we work in the COM frame, describing two real particles as an effective particle moving in a potential. Do not forget that for any experiment we will have to apply the above transformation to relate theory to the experimental results, though if the target particle is much heavier than the other the transformation may be slight. Note also that this transformation is invalid if there is an external field.

### 8.2 Some terminology for general scattering

The incident flux (I) of particles with momentum  $\mathbf{p} = \hbar\mathbf{k}$  is the number of incident particles crossing unit area perpendicular to the beam direction per unit time.

The scattered flux (S) of particles with momentum  $\mathbf{p}' = \hbar\mathbf{k}'$ , is the number of scattered particles scattered into the element of solid angle  $d\Omega$  about the direction  $\theta, \phi$  per unit time per unit solid angle.

The differential cross section is the ratio of the scattered flux in direction  $\theta, \phi$  to the incident flux.

$$\frac{d\sigma}{d\Omega} = S/I$$

The total cross section is the ratio of the scattered flux in any direction to the incident flux.

$$\sigma_T = \int \int \frac{d\sigma}{d\Omega} \sin\theta d\theta d\phi$$

### 8.3 Scattering in one dimension- Step function

Firstly, we review the problem of scattering by a step function in one dimension. Consider a particle moving from a region ( $x < 0$ ) where the potential is  $V = 0$  to a region ( $x > 0$ ) where the potential is  $V = V_0$ .

Assuming the particle energy  $E > V_0$ , this is simply the free particle problem, the spatial solution to which is:

$$\Phi = A \exp(ikx) + B \exp(-ikx) \quad (x < 0); \quad \Phi = C \exp(ik'x) + D \exp(-ik'x) \quad (x > 0)$$

where  $k = \sqrt{2mE}/\hbar$  and  $k' = \sqrt{2m(E - V_0)}/\hbar$

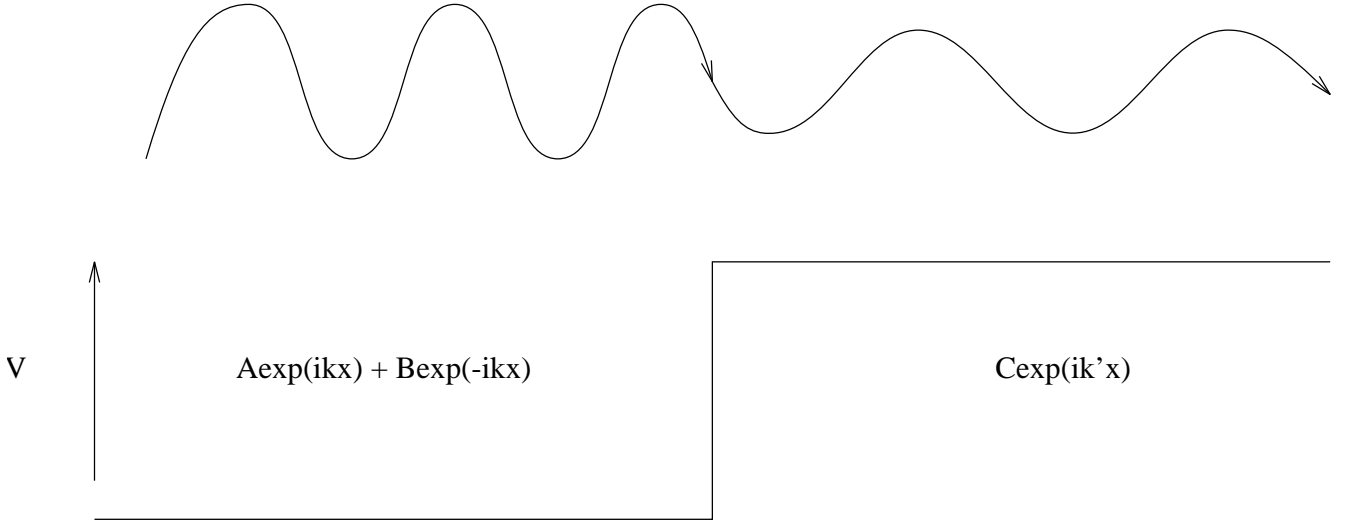


Figure 7: Scattering at a step function.

From the boundary condition that all particles start from  $x = -\infty$ , we can immediately set  $D=0$ . From the condition of continuity of  $\Phi$  and  $d\Phi/dx$  at  $x = 0$  we also require  $A + B = C$  and  $k(A - B) = k'C$

This gives the reflected amplitude  $B/A = (k - k')/(k + k')$  and the transmitted amplitude  $C/A = 2k/(k + k')$

The reflected flux is thus

$$\frac{\hbar k}{m} A^2 \left( \frac{k - k'}{k + k'} \right)^2$$

and the transmitted flux is

$$\frac{\hbar k'}{m} A^2 \left( \frac{2k}{k + k'} \right)^2$$

Note that  $A^2 \neq B^2 + C^2$ . The conserved quantity is the *flux* of particles, not the probability density. In this case the transmitted particles are moving more slowly than the incident ones.

Notice that if  $V_0$  is negative, the transmitted flux gets smaller as  $|V_0|$  gets larger: it is difficult to fall off a big cliff! This anomaly is due to the unphysical potential - the discontinuous first derivative at  $x = 0$ .

We have not considered the case of  $E < V_0$ . Now  $\Phi(x > 0) = Ce^{-\kappa'x}$  where  $\kappa' = \sqrt{2m(V_0 - E)}/\hbar$ . The boundary conditions are then  $A + B = C$  and  $ik(A - B) = \kappa'C$ , which gives the reflected amplitude  $B/A = (ik - \kappa')/(ik + \kappa')$  and the transmitted amplitude  $C\kappa'/Ak = 2ik/(ik + \kappa')$ .

Now the reflected flux is equal to the incident flux, and although the wavefunction penetrates the region  $x > 0$ , it decays exponentially and there is no propagating wave.

## 8.4 Scattering in one dimension - Square Well

The square well potential has  $V(x < 0) = V(x > a) = 0$ ;  $V(0 < x < a) = V_0$ . As with the step function, we can write the wavefunction as a plane wave in each of the three regions.

$$\begin{aligned}\Phi(x < 0) &= A \exp(ikx) + B \exp(-ikx) \\ \Phi(0 < x < a) &= F \exp(ik'x) + G \exp(-ik'x) \\ \Phi(x > a) &= C \exp(ikx) + D \exp(-ikx)\end{aligned}$$

Once again there is no wave coming back from  $x = \infty$  ( $D = 0$ ).

There are now four boundary conditions from continuity of the wave function and its derivative at  $x=0$  and  $x=a$ . The solving of four equations in four unknowns is straightforward but tedious. Eventually one can obtain ratios for reflected and transmitted flux:

$$\begin{aligned}B/A &= \frac{(k^2 - k'^2)(1 - e^{2ik'a})}{(k + k')^2 - (k - k')^2 e^{2ik'a}} \\ C/A &= \frac{4kk' e^{i(k'-k)a}}{(k + k')^2 - (k - k')^2 e^{2ik'a}}\end{aligned}$$

where  $k^2 = 2mE/\hbar^2$  and  $k'^2 = 2m(E - V_0)/\hbar^2$ . Since the wavenumber is the same on both sides of the barrier, the reflection and transmission coefficients are just:

$$\begin{aligned}|B/A|^2 &= \left[ 1 + \frac{4k^2 k'^2}{(k^2 - k'^2)^2 \sin^2 k'a} \right]^{-1} = \left[ 1 + \frac{4E(E - V_0)}{V_0^2 \sin^2 k'a} \right]^{-1} \\ |C/A|^2 &= \left[ 1 + \frac{(k^2 - k'^2)^2 \sin^2 k'a}{4k^2 k'^2} \right]^{-1} = \left[ 1 + \frac{V_0^2 \sin^2 k'a}{4E(E - V_0)} \right]^{-1}\end{aligned}$$

We get complete transmission when  $k'a = n\pi$ , i.e. when an exact number of half waves fit in the well.

Assuming that  $E > V_0$ . Looking at the limits of this, we see that as  $E \rightarrow V_0$  then  $\sin^2(k'a) \rightarrow k'a$  and the transmission coefficient

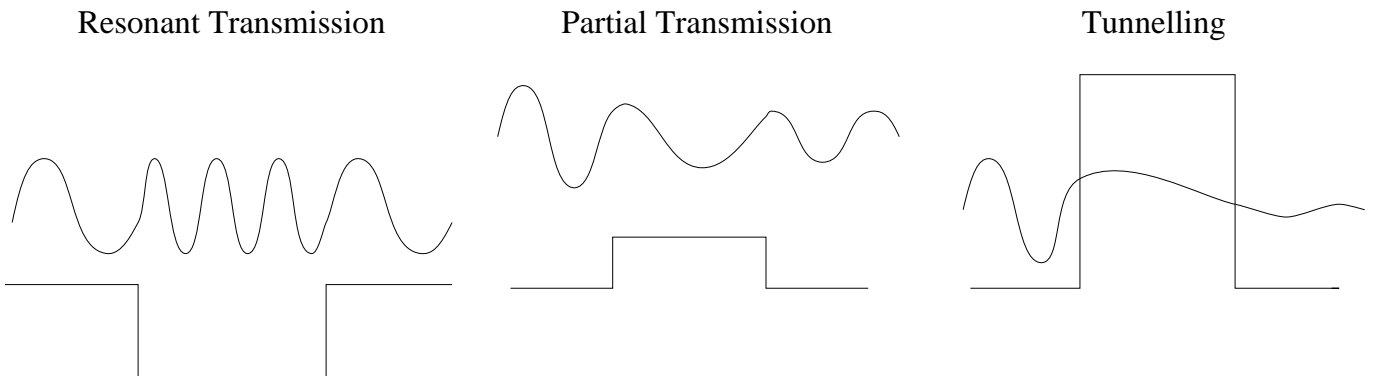


Figure 8: Forward moving wavefunctions passing a square well potential

$$|C/A|^2 \rightarrow \left[ 1 + \frac{mV_0a^2}{2\hbar^2} \right]^{-1}$$

As the incoming particle energy is increased, the transmission oscillates between  $\left[ 1 + \frac{V_0^2}{4E(E-V_0)} \right]^{-1}$  and 1 at  $k'a = n\pi$ . The lower limit itself increases to 1 as E increases.

For the tunnelling case where  $E < V_0$  we can use these solutions for B/A and C/A, except that  $k'$  is now imaginary. This gives

$$|C/A|^2 = \left[ 1 + \frac{4E(E - V_0)}{V_0^2 \sinh^2 |k'|a} \right]^{-1}$$

which decreases monotonically with decreasing E. Thus a small change in  $V_0$  can give a large change in  $|C/A|^2$ . This is the principle on which the transistor and the tunnelling electron microscope are based.

Note that the transmitted wave  $\Phi(x > a) = C \exp(ikx)$ , differs from the incident wave only by a phase - it has the same wavevector. Thus the only effect of the potential on the transmitted particles is to change their *phase*, an idea we shall meet again.

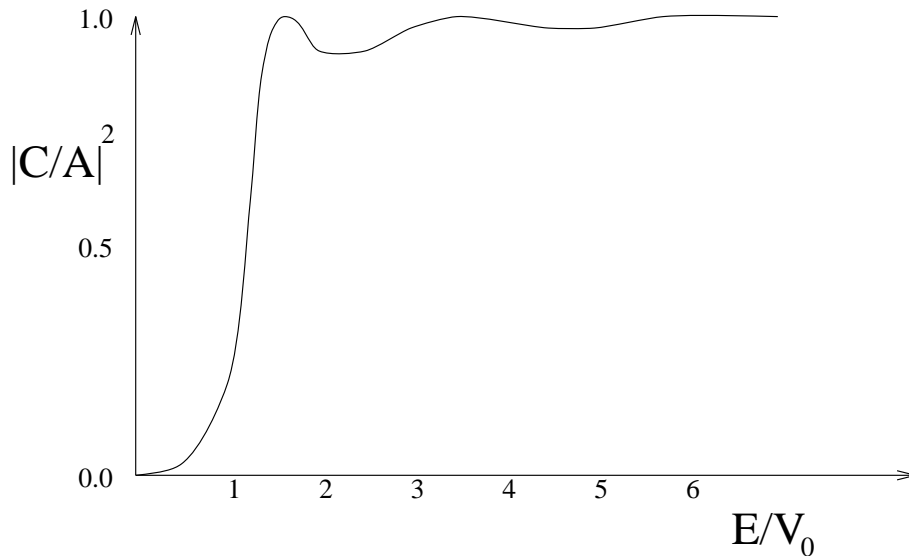


Figure 9: Transmission coefficient across a 1D square well

## 8.5 The transistor

Transistors can be modelled as a barrier potential, with the voltage across them represented by different potentials on either side.

The rapid variation in transmission coefficient (current) with change in potential barrier (voltage) is the basis of the transistor. The name comes from 'transfer resistor'. The resistance to motion of electrons past the barrier is determined by the voltage  $V_0$  in the barrier region more than the voltage difference across the transistor.

Actual behaviour also depends on the availability of electrons for conduction, which depends in turn on the material in question, since there must be available electron states of appropriate energy on each side of the barrier.

## 9 Scattering in three dimensions

### 9.1 Cross sections and geometry

Most experiments in physics consist of sending one particle to collide with another, and looking at what comes out.

The quantity we can usually measure is the scattering cross section: by analogy with classical scattering of hard spheres, we assuming that scattering occurs if the particles ‘hit’ each other. The cross section is the apparent ‘target area’. The total scattering cross section can be determined by the reduction in intensity of a beam of particles passing through a region on ‘targets’, while the differential scattering cross section requires detecting the scattered particles at different angles.

We will use spherical polar coordinates, with the scattering potential located at the origin and the plane wave incident flux parallel to the  $z$  direction. In this coordinate system, scattering processes are symmetric about  $\phi$ , so  $\frac{d\sigma}{d\Omega}$  will be independent of  $\phi$ .

We will also use a purely classical concept, the impact parameter  $b$  which is defined as the distance of the incident particle from the  $z$ -axis prior to scattering.

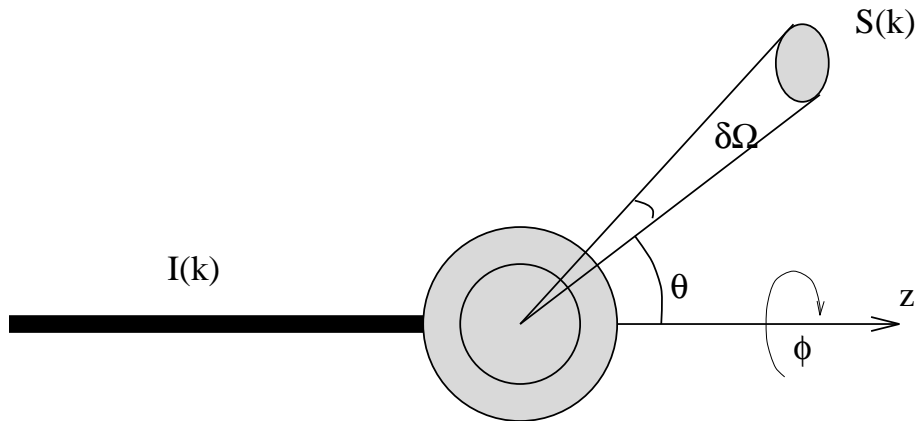


Figure 10: Standard spherical coordinate geometry for scattering

### 9.2 The Born Approximation

We can use the results of time-dependent perturbation theory to do an approximate calculation of the cross-section. Provided that the interaction between particle and scattering centre is *localised* to the region around  $r = 0$ , we can regard the incident and scattered particles as free when they are far from the scattering centre. We just need the result that we obtained for a *constant* perturbation, Fermi’s Golden Rule, to compute the rate of transitions between the initial state (free particle of momentum  $\mathbf{p}$ ) to the final state (free particle of momentum  $\mathbf{p}'$ ).

We write the Hamiltonian as

$$\hat{H} = \hat{H}_0 + \hat{V}(\mathbf{r}) \quad \text{where} \quad \hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2m}, \text{ the kinetic energy operator}$$

and treat the potential energy operator,  $\hat{V}(\mathbf{r})$ , as the perturbation inducing transitions between the eigenstates of  $\hat{H}_0$ , which are plane waves.

If we label the initial and final plane-wave states by their respective wave-vectors we have for the rate of transitions

$$R = \frac{2\pi}{\hbar} |\langle \mathbf{k}' | \hat{V} | \mathbf{k} \rangle|^2 g(E_k)$$

where  $g(E_k)$  is the density of final states;  $g(E_k)dE_k$  is the number of final states with energy in the range  $E_k \rightarrow E_k + dE_k$ .

The quantity  $\langle \mathbf{k}' | \hat{V} | \mathbf{k} \rangle$  is known as the matrix element of the perturbation and is usually abbreviated thus

$$V_{\mathbf{k}'\mathbf{k}} \equiv \langle \mathbf{k}' | \hat{V} | \mathbf{k} \rangle = \int \int \int u_{\mathbf{k}'}^*(\mathbf{r}) V(\mathbf{r}) u_{\mathbf{k}}(\mathbf{r}) d\tau.$$

### 9.3 Box Normalisation

Plane-wave states have wavefunctions of the form:  $u_{\mathbf{k}}(\mathbf{r}) = C \exp(i\mathbf{k}\cdot\mathbf{r})$  with  $C$  a normalisation constant. Because plane-wave states are not properly normalisable we employ the trick of normalising them in a large (relative to potential range) cubic box of side  $L$  with periodic boundary conditions. We then take the limit  $L \rightarrow \infty$  at the end of the calculation.

Thus we require that

$$\int \int \int_{\text{box}} u_{\mathbf{k}}^*(\mathbf{r}) u_{\mathbf{k}}(\mathbf{r}) d\tau = |C|^2 \int \int \int_{\text{box}} d\tau = |C|^2 L^3 = 1$$

giving for the normalised eigenfunctions:  $u_{\mathbf{k}}(\mathbf{r}) = L^{-3/2} \exp(i\mathbf{k}\cdot\mathbf{r})$

Of course, enclosing the system in a finite box has the consequence that the allowed momentum eigenvalues are no longer continuous but *discrete*. With periodic boundary conditions

$$u(-\frac{L}{2}, y, z) = u(\frac{L}{2}, y, z), \quad \text{etc.}$$

the momentum eigenvalues are forced to be of the form

$$\mathbf{p} \equiv \hbar\mathbf{k} = \frac{2\pi\hbar}{L}(n_x, n_y, n_z), \quad \text{with } n_x, n_y, n_z = 0, \pm 1, \pm 2, \dots$$

For sufficiently large  $L$ , we can approximate the continuous spectrum arbitrarily closely.

### 9.4 Density of Final States

Any possible final-state wave-vector,  $\mathbf{k}$ , corresponds to a point in *wave-vector* space with coordinates  $(k_x, k_y, k_z)$ . The points form a cubic lattice with lattice spacing  $2\pi/L$ . Thus the volume of  $k$ -space per lattice point is  $(2\pi/L)^3$ , and the number of states in a volume element  $d^3\mathbf{k}$  is

$$\left(\frac{L}{2\pi}\right)^3 d^3\mathbf{k}' = \left(\frac{L}{2\pi}\right)^3 k^2 dk d\Omega$$

We require  $g(E_k)$ , the density of states per unit energy, where:

$$E_k = \frac{\hbar^2 k^2}{2m}$$

is the energy corresponding to wave-vector  $\mathbf{k}'$ . Now, the wave-vectors in the range  $\mathbf{k}' \rightarrow \mathbf{k}' + d^3\mathbf{k}'$  correspond to the energy range  $E_k \rightarrow E_k + dE_k$ , so that

$$g(E_k) dE_k = \left(\frac{L}{2\pi}\right)^3 k^2 dk d\Omega$$

is the number of states with energy in the desired interval and with wave-vector,  $\mathbf{k}'$ , pointing into the solid angle  $d\Omega$  about the direction  $(\theta, \phi)$ . Noting that

$$dE_k = \frac{\hbar^2 k}{m} dk$$

yields the final result for the density of states,

$$g(E_k) = \frac{L^3 m k}{8\pi^3 \hbar^2} d\Omega$$

## 9.5 Incident and Scattered Flux

The box normalisation corresponds to one particle per volume  $L^3$ , so that the number of particles crossing unit area perpendicular to the beam per unit time is just given by the magnitude of the incident velocity divided by  $L^3$ :

$$\text{incident flux} = \frac{|\mathbf{p}|/m}{L^3} = \frac{\hbar k}{mL^3}$$

Using the Golden Rule, we have that the rate of transitions between the initial state of wave-vector  $\mathbf{k}$  and final states whose wave-vectors  $\mathbf{k}'$  lie in the element of solid angle  $d\Omega$  about the direction  $(\theta, \phi)$  of the wave-vector  $\mathbf{k}'$ , is given by

$$R = \frac{2\pi}{\hbar} |V_{\mathbf{k}'\mathbf{k}}|^2 \frac{L^3}{8\pi^3} \frac{mk}{\hbar^2} d\Omega$$

but this is just the number of particles scattered into  $d\Omega$  per unit time. To get the scattered flux we simply divide by  $d\Omega$  to get the number per unit time per unit solid angle.

## 9.6 The Differential Cross-Section

We now have all the ingredients, the scattered flux and the incident flux, to compute the cross-section:

$$\frac{d\sigma}{d\Omega} \equiv \frac{\text{scattered flux}}{\text{incident flux}} = \frac{mL^3}{\hbar k} \frac{2\pi}{\hbar} |V_{\mathbf{k}'\mathbf{k}}|^2 \frac{L^3}{8\pi^3} \frac{mk}{\hbar^2}$$

Noting that, for *elastic scattering*,  $k' = k$ , we obtain finally the so-called Born approximation for the differential cross-section:

$$\frac{d\sigma}{d\Omega} = \frac{m^2}{4\pi^2\hbar^4} L^6 |\langle \mathbf{k}' | \hat{V} | \mathbf{k} \rangle|^2$$

where the matrix element  $V_{\mathbf{k}'\mathbf{k}} \equiv \langle \mathbf{k}' | \hat{V} | \mathbf{k} \rangle$  is given by

$$\langle \mathbf{k}' | \hat{V} | \mathbf{k} \rangle = \frac{1}{L^3} \int \int \int V(\mathbf{r}) \exp(-i\chi \cdot \mathbf{r}) d\tau$$

with  $\chi \equiv \mathbf{k}' - \mathbf{k}$ , the so-called *wave-vector transfer*. Thus the required matrix element in the Born approximation is just the 3-dimensional Fourier transform of the potential energy function.

Observe that the final result for the differential cross-section is independent of the box size,  $L$ , which we used to normalise the plane-wave states.

## 9.7 Further Simplification to 1D for Conservative, Central Potential

Consider a central potential  $V(\mathbf{r}) = V(|\mathbf{r}|)$  where energy is conserved  $|\mathbf{k}'|^2 = |\mathbf{k}|^2$ . Here  $\chi$  is a vector of length  $2k \sin \frac{\theta}{2}$  where  $\theta$  is the scattering angle.

We can make some progress with the matrix element integral if we choose a polar coordinate system with  $\chi$  along the  $z$ -axis, so that  $\chi \cdot \mathbf{r} = \chi r \cos \theta$ . Since we are trying to integrate over all space this change does not affect the limits of the integral.

$$\begin{aligned} V_{\mathbf{k}'\mathbf{k}} &= \int_0^{2\pi} d\phi \int_{-1}^{+1} \int_0^\infty V(r) e^{-i\chi r \cos \theta} r^2 dr d(\cos \theta) \\ &= 2\pi \int_0^\infty \frac{e^{-i\chi r} - e^{i\chi r}}{-i\chi r} V(r) r^2 dr = \frac{4\pi}{\chi} \int_0^\infty r V(r) \sin(\chi r) dr \end{aligned}$$

But since  $|\mathbf{k}| = |\mathbf{k}'|$ ,  $|\chi| = 2k \sin \frac{\theta}{2}$ , Whence we obtain the most useful form of the Born approximation:

$$\frac{d\sigma}{d\Omega} = \frac{m^2}{(k \sin \frac{\theta}{2})^2 \hbar^4} \left| \int_0^\infty rV(r) \sin(2kr \sin \frac{\theta}{2}) dr \right|^2$$

Thus the scattering cross-section is independent of  $\phi$  (due to cylindrical symmetry of the problem). Note that this shows that the differential cross section does not depend on scattering angle and beam energy independently, but on a single parameter  $\chi$ . By using a range of energies for the incoming particles,  $k$ , this dependence can be used to test whether experimental data can be well described by the Born Approximation.

The most common use of the Born approximation is, of course, in reverse. Having found  $\frac{d\sigma}{d\Omega}$  experimentally, a reverse Fourier transform can be used to obtain the form of the potential.

## 9.8 Example of Born Approximation

Consider scattering of particles interacting via a 3D square well potential:  $V(r < a) = V_0$ ;  $V(r > a) = 0$ .

The integral required here is then (with  $\chi = 2k \sin \frac{\theta}{2}$ ):

$$\int_0^a rV_0 \sin(\chi r) dr = \left[ \frac{\sin(\chi r) - \chi r \cos(\chi r)}{\chi^2} \right]_0^a$$

whence:

$$\frac{d\sigma}{d\Omega} = \left[ \frac{2\mu V_0}{\chi \hbar^2} \right]^2 \left[ \frac{\sin(\chi a) - \chi a \cos(\chi a)}{\chi^2} \right]^2$$

Using a Maclaurin expansion, the low energy limit is:

$$\frac{d\sigma}{d\Omega} = \left[ \frac{2\mu V_0}{\chi \hbar^2} \right]^2 \frac{1}{9} \left[ 1 - \frac{1}{5} \chi^2 a^2 \right]$$

From integrating over  $\theta$  and  $\phi$  the low and high energy limits for the total cross section are

$$\sigma(E \rightarrow \infty) = 2\pi \left[ \frac{\mu}{\hbar^2} \right]^2 \left[ \frac{V_0 a^3}{ka} \right]^2 \quad \sigma(E \rightarrow 0) = 2\pi \left[ \frac{\mu}{\hbar^2} \right]^2 \left[ \frac{V_0 a^3}{ka} \right]^2 \frac{8}{9} (k^2 a^2 - \frac{2}{5} k^4 a^4 + \dots)$$

## 9.9 General Notes on Scattering in the Born Approximation

The square well illustrates some general feature of scattering in the Born approximation:

- Born approximation is based on perturbation theory, so works best for high energy particles.
- Scattering depends on  $V_0^2$ , so both attractive and repulsive potentials behave the same.
- At high energy, cross section is inversely proportional to the energy ( $E = \hbar^2 k^2 / 2m$ )
- Dependence on  $k$  and  $\theta$  arises only through the combination  $\chi = 2k \sin \frac{\theta}{2}$ . Thus as energy increases, the scattering angle  $\theta$  is reduced and the scattered beam becomes more peaked in the 'straight on' direction.
- Angular dependence depends on the range of the potential  $a$  but not on the strength  $V_0$ .
- Total cross section depends on both range  $a$  and depth  $V_0$  of the potential.

## 10 Partial Wave Analysis

### 10.1 Using conservation of angular momentum

One method of solving the scattering problem is to use the Born Approximation. This is a perturbation method based on the Fermi Golden Rule and is therefore valid for short-ranged, weak potentials. For the case of a central potential a more general method exists: Partial waves.

It is well known that a particle moving in a central potential experiences no torque (about the potential origin), and therefore conserves angular momentum. In the quantum case angular momentum is quantised, so the scattering does not change the angular momentum quantum number  $l$ . Partial waves proceeds by a ‘divide and conquer’ strategy of expanding the incident and scattered fluxes in a basis set of distinct angular momentum. The angular parts of this basis are unaffected by the scattering, thus we need consider only the 1D radial problem. Moreover, for a conservative potential, the energy of the particles is unaffected, thus  $|k| = |k'|$ , so all that a central potential can achieve is a change of *phase*.

Like the Born approximation, the derivation of the Partial Wave equation is complicated, but it can be done for a general potential, so once we have the result we can just use it.

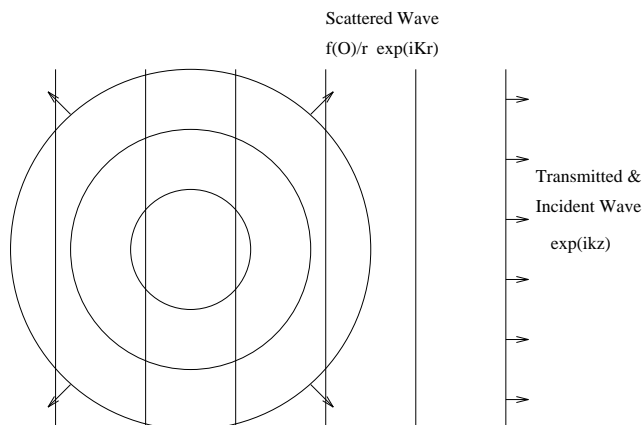


Figure 11: Plane wave in, radial wave out.  $|\Psi\rangle = e^{ikz} + f(\theta)e^{ikr}/r$

### 10.2 Digression: Expanding a plane wave in an angular momentum basis set

A preliminary step in partial wave analysis is to expand the incident plane wave in a set of partial waves. The solution to the Schroedinger equation for a free particle in spherical polars can be separated into three parts:

$$\text{Incident free particle wavefunction} = e^{iKz} = e^{iKr \cos \theta} = R(r)\Theta(\theta)\Phi(\phi)$$

Solving for  $\Phi$  is trivial - there is no  $\phi$  dependence of  $e^{iKr \cos \theta}$  - so  $\Phi(\phi) = 1$ .  $\Theta$  is also straightforward, since we have already solved the Schroedinger equation for hydrogen, and the potential has the same  $\theta$  dependence (i.e. none), so the  $\Theta$  functions must be the Legendre polynomials  $P_l(\cos \theta)$  which are orthogonal and normalised to  $\langle P_l P_k \rangle = \frac{2}{2l+1} \delta_{kl}$ . By analogy with hydrogen (though with  $V=0$ ),  $R(r)$  must now be a solution to the equation:

$$\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) - \frac{l(l+1)}{r^2} R + \frac{2\mu}{\hbar^2} ER = 0$$

Because we are describing free particles,  $E > 0$  and  $E$  is not quantised. Incoming beams of any energy can be treated. The solutions to this equation are well known to mathematicians, and can be found in most quantum mechanics books. They are the Spherical Bessel Functions:  $j_l(Kr)$ .

Thus we can expand  $e^{iKr \cos \theta}$  in spherical harmonics and spherical Bessel functions - a complete set of orthonormal basis states which are eigenstates of the angular momentum:  $\hat{L}^2 |j_l(Kr) P_l(\cos \theta)\rangle = l(l+1)\hbar |j_l(Kr) P_l(\cos \theta)\rangle$ . Of the complete set of spherical harmonics, we need only the Legendre polynomials  $P_l(\cos \theta) = Y_{l0}$  because of the cylindrical symmetry ( $\Phi(\phi) = 1$ ). We can write the plane wave along  $z$  in this basis set, with coefficients  $a_n$

$$\exp(iKr \cos \theta) = \sum_{n=0}^{\infty} a_n j_n(Kr) P_n(\cos \theta)$$

whence, multiplying by  $P_l(\cos \theta)$  and integrating over  $\theta$  to pick out a specific component:

$$\langle P_l | \exp(iKr \cos \theta) \rangle = \sum_{n=0}^{\infty} a_n j_n(Kr) \langle P_l | P_n \rangle = \frac{2}{2l+1} a_l j_l$$

Integrating the left hand side by parts gives a term proportional to  $r^{-1}$  and a series of terms of higher order  $r^{-n}$ . Taking the limit as  $r \rightarrow \infty$ , where only the  $r^{-1}$  term is significant:

$$\frac{2}{2l+1} a_l j_l = \frac{1}{iKr} [P_l(\cos \theta) \exp(iKr \cos \theta)]_{\cos \theta = -1}^{\cos \theta = 1} \quad r \rightarrow \infty$$

This boundary condition is sufficient to determine the  $a_l$ .

Now, for Legendre polynomials  $P_l(\cos \theta = 1) = 1$  and  $P_l(\cos \theta = -1) = (-1)^l = e^{il\pi}$  so that:

$$\frac{2}{2l+1} a_l j_l = \frac{1}{iKr} [e^{iKr} - e^{il\pi - iKr}] \quad r \rightarrow \infty$$

which after a bit of manipulation (use  $e^{il\pi/2} = i^l$ ) becomes:

$$a_l j_l(Kr) = (2l+1) i^l \frac{\sin(Kr - l\pi/2)}{Kr} \quad r \rightarrow \infty$$

For the Kr dependence to be correct, we must have:

$$j_l(Kr) = \frac{\sin(Kr - l\pi/2)}{Kr} \quad r \rightarrow \infty$$

which can be confirmed by comparing the form of the Bessel function at  $r \rightarrow \infty$  with the plane wave. Thus  $a_l = (2l+1) i^l$ .

Finally, we can write:

$$\exp(iKr \cos \theta) = \sum_{l=0}^{\infty} i^l j_l(Kr) (2l+1) P_l(\cos \theta)$$

which is the representation of a plane wave as a linear combination of partial waves with distinct angular momentum. This is the starting point for partial wave analysis. Because it is independent of the scattering potential we do not need to repeat this expansion (or one like it) every time we do a partial wave calculation: we just use the result.

Note the term  $(2l+1)$ . This can be related to the classical ‘impact parameter’ mentioned above. The angular momentum of a particle of velocity  $v$  is  $mbv = \sqrt{l(l+1)}\hbar$ . Thus a classical (large  $l$ ) particle with angular momentum  $l\hbar$  would pass between a ring of radius  $b = l\hbar/mv$  and one of radius  $b = (l+1)\hbar/mv$ . The area between these rings is  $(2l+1)\pi(\hbar/mv)^2$  so for a uniform beam the probability of a particle having angular momentum  $l$  is proportional to  $(2l+1)$ .

### 10.3 Incident and Scattered Flux

If we can write the solution to the Schrödinger equation  $|\Psi\rangle$  at large  $r$  in the form of an incident plane wave and a scattered radial wave:

$$|\Psi\rangle = \text{IncidentWave} + \text{ScatteredWave} = \exp iKz + \frac{f(\theta)}{r} \exp iKr$$

Then the incident flux is the product of the probability density and the particle velocity  $I = v e^{iKz} e^{-iKz} = v = \hbar k/m$ . Likewise the scattered flux must have a radial function which gives a normalisable plane wave ( $e^{-iKr}/r$ ), and a  $\theta$  dependence arising from the scattering, which we call  $f(\theta)$ . By symmetry, there is no  $\phi$  dependence. Thus the scattered flux per unit area will be:  $v f^*(\theta) f(\theta)/r^2$ .

Thus  $d\sigma/d\Omega = S(\theta)/I = f^*(\theta) f(\theta)$ , and all we need do is calculate  $f(\theta)$ .

### 10.4 Solving the Schrödinger Equation - the Phase Shifts

In the previous section we solved this problem without the potential (i.e.  $f(\theta) = 0$ ), we now solve it with the central potential. For a spherically symmetric potential, the angular part of the Hamiltonian is the same as the previous section, because the potential is independent of  $\phi$  and  $\theta$ , so the analysis is the same except the equation for  $R(Kr)$ , which becomes:

$$\frac{d^2 u_l(r)}{dr^2} - \frac{l(l+1)}{r^2} u_l(r) + \frac{2\mu}{\hbar^2} [E - V(r)] u_l(r) = 0$$

where we set  $u_l(r) = r R_l(r)$ , the same substitution as in the atomic hydrogen problem.

If we look at the limit of large  $R(Kr \rightarrow \infty)$ , where the detector in any experiment would be stationed, we find  $V(r \rightarrow \infty) = 0$  and so  $R_l(Kr \rightarrow \infty)$  describes a free particle. The solution must therefore tend to the same limit as  $j_l(Kr \rightarrow \infty)$ , though perhaps with different phase:

$$R_l(Kr) = \sin(Kr - l\pi/2 + \delta_l)/Kr$$

This is always a solution at  $r \rightarrow \infty$  provided that  $V(r) \rightarrow 0$  faster than  $1/r$  (Localised potential).

Thus to perform partial wave analysis we need to solve the radial Schrödinger equation for  $R_l(Kr)$  for each angular momentum component.

The wavefunction at long range describes a free particle with the same  $-K-$  as the incident beam. The only possible effect of the conservative, central potential is to change the *phase* of plane wave by  $\delta_l$ . The limit of  $r \rightarrow \infty$  for each solution can be expressed by a single number: the phase shift  $\delta_l$ .

### 10.5 Obtaining Cross sections

Recalling that to get cross sections we need to find  $f(\theta)$ , we express  $\Psi$  in the appropriate form at  $R \rightarrow \infty$ .

$$\Psi = e^{iKz} + f(\theta) \frac{e^{iKr}}{r} = \sum_{l=0}^{\infty} i^l j_l(Kr) (2l+1) P_l(\cos\theta) + f(\theta) \frac{e^{iKr}}{r} = \sum_{l=0}^{\infty} b_l R_l(Kr) P_l(\cos\theta)$$

where  $b_l$  are expansion coefficients for the expression of  $\Psi$  in the partial wave basis, the equivalent of the  $a_l$  for the expression of a free particle in the partial wave basis.

We already know the  $r \rightarrow \infty$  values for  $j_l$  and  $R_l$ . Using the limiting values of  $j_l(Kr \rightarrow \infty)$ , writing the equation above in terms of complex exponentials and multiplying by  $2iKr$  we find:

$$\sum_{l=0}^{\infty} i^l [e^{i(Kr-l\pi/2)} - e^{-i(Kr-l\pi/2)}] (2l+1) P_l(\cos \theta) + 2iK f(\theta) e^{iKr} = \sum_{l=0}^{\infty} b_l [e^{i(Kr-l\pi/2+\delta_l)} - e^{-i(Kr-l\pi/2+\delta_l)}] P_l(\cos \theta)$$

Comparing the coefficients of the  $e^{-iKr}$  term, we can solve for the expansion coefficients  $b_l$

$$(2l+1)i^l e^{il\pi/2} = b_l e^{il\pi/2 - i\delta_l} \quad ; \quad b_l = (2l+1)i^l e^{i\delta_l}$$

and now, using these values for  $b_l$  we can compare the coefficients of the  $e^{iKr}$  terms, and after a little manipulation of complex exponentials, we find:

$$f(\theta) = K^{-1} \sum_{l=0}^{\infty} (2l+1) e^{i\delta_l} \sin \delta_l P_l(\cos \theta)$$

From this we can calculate  $d\sigma/d\Omega = |f(\theta)|^2$  and  $\sigma = 2\pi \int |f(\theta)|^2 d\theta$ . Note that  $d\sigma/d\Omega$  involves the product of two series, and thus contains many cross terms. In general it is very complicated.

However, when integrated over all  $\theta$  these cross terms vanish due to orthogonality of the Legendre polynomials  $\langle P_l | P_{l'} \rangle = 0$  ( $l \neq l'$ ). Thus the total cross section, expressed in partial waves, has a particularly simple form:

$$\sigma = \frac{4\pi}{K^2} \sum_{l=0}^{\infty} (2l+1) \sin^2 \delta_l$$

Hence the scattering cross sections are completely determined by  $|K|$  and the phase shifts  $\delta_l$ . For a given problem, all we must calculate are the  $\delta_l$ , then we can simply apply the result above for  $\sigma$ .

Thus all the effect of the potential on a given partial wave is contained in a single number - the phase shift. This is the amount by which can be imagined as the amount a given partial wave is pulled in by the potential. The phase shifts must be obtained by solving the radial equation for  $R_l(Kr)$  and comparing with  $j_l(Kr)$  at large  $r$  for each  $l$ . Although the analysis done so far is valid for all problems, for a specific problem, one must still evaluate the phase shifts and solve the radial equation.

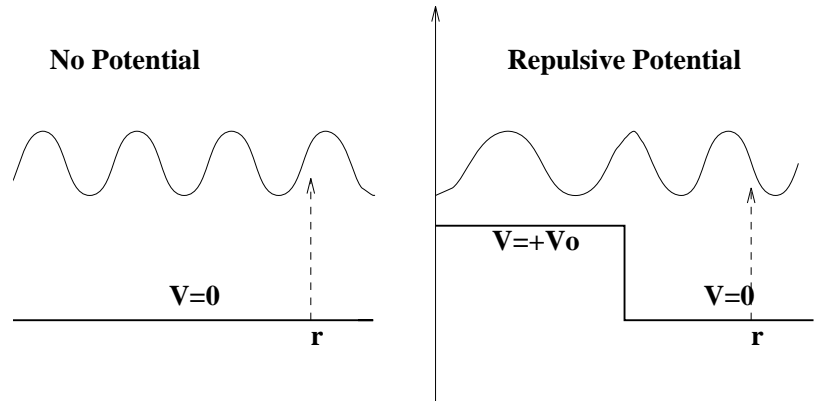


Figure 12: Radial wavefunctions,  $u_l(r) = rR(r)$  showing phase difference at  $r$  due to short-ranged potential. The attractive potential pulls in the wave giving negative  $\delta_l$ , while the repulsive potential pushes out the wave for positive  $\delta_l$

## 11 Using Partial Waves

### 11.1 Impact Parameter and Classical Analogies

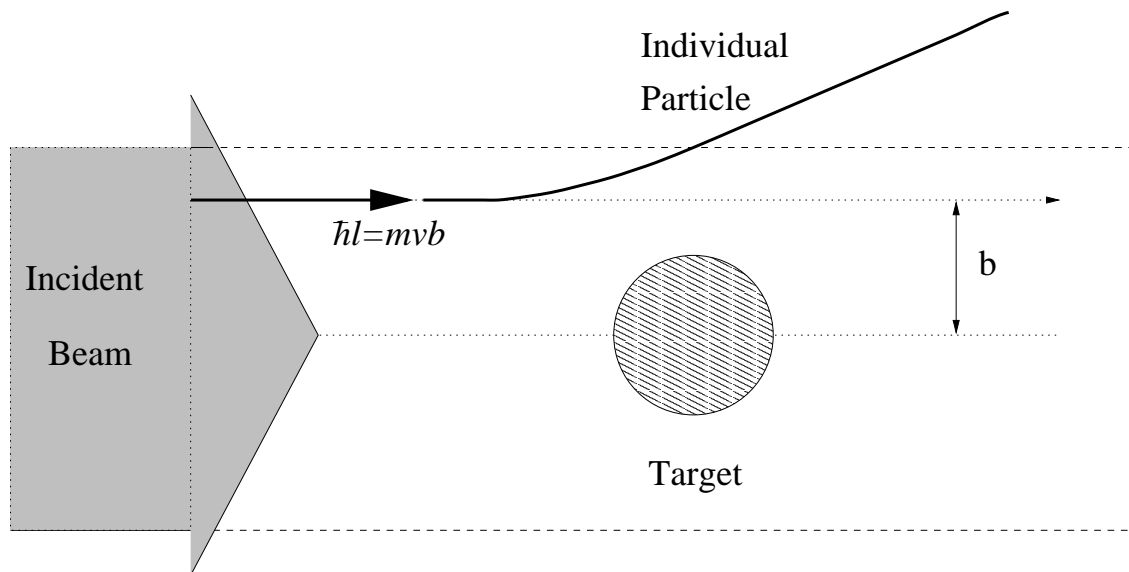


Figure 13: Relation between classical and quantum angular momentum

Knowing the impact parameter gives us some classical idea of whether a scattering event is likely. If the impact parameter is larger than the range of the potential, then classically the particles would miss. In the quantum case, we expect this to mean that the phase shift for that angular momentum is zero, and hence that the contribution from that term in the expansion is zero. Thus at a given incoming momentum,  $\hbar k$ , we can determine how many terms in the partial wave expansion to consider from  $\hbar k b_{max} \approx l_{max} \hbar$ , where  $b_{max}$  is the maximum impact parameter for classical collision, i.e. the range of the potential.

### 11.2 S-wave scattering

Although exact at all energies, the partial wave method is most useful for dealing with scattering of low energy particles. This is because for slow moving particles to have large angular momentum ( $\hbar k b$ ) they must have large impact parameters  $b$ . Classically, particles with impact parameter larger than the range of the potential miss the potential. Thus for scattering of slow-moving particles we need only consider a few partial waves, all the others are unaffected by the potential ( $\delta_l \approx 0$ ). Thus partial waves and the Born approximation are complementary methods, good for slow and fast particles respectively.

For very low energy we need consider only the first term in the partial wave expansion. This is known as *S-wave scattering*. In this case it is possible to solve for the differential cross section, since only the first term in the series for  $f(\theta)$  is involved: Since the angular variation is  $P_0(\cos \theta) = 1$  the scattering is isotropic.

$$\frac{d\sigma}{d\theta} = |f(\theta)|^2 = k^{-2} \sin^2 \delta_0$$

At higher energies, other angular momentum components come into play. For a given  $l$  component, scattering is maximised for  $\delta_l = \pi/2$ .

### 11.3 Resonance

In some cases where a potential has a bound state of particular angular momentum, the scattering of particle with that angular momentum will be especially enhanced. In such cases the total scattering cross section will show a peak, and the angular distribution will be characteristic of the appropriate  $P_l(\cos \theta)$ . This very strong scattering is known as resonance and is a powerful method for studying bound states.

### 11.4 Example of S-wave scattering - Attractive square well potential

An example where we can solve for the phase shift is the 3D-square well potential:

$$(V(r < R) = -V_0; V(r > R) = 0).$$

For the  $l = 0$  case the radial equation with  $U_0 = R_0 r$  is

$$\frac{d^2 u_0(r)}{dr^2} + \frac{2\mu}{\hbar^2} [E - V(r)] u_0(r) = 0$$

The solutions to this are familiar from the 1D square well. If we write

$$K_0 = \sqrt{2\mu[E + V_0]}/\hbar; \quad K = \sqrt{2\mu E}/\hbar$$

then for  $r < R$ ,  $u(r) = A \sin K_0 r + B \cos K_0 r$ .

and for  $r > R$ ,  $u(r) = C \sin Kr + D \cos Kr$ . which can easily be written in a different form to show the appropriate phase shift  $\delta_0$ :  $u(r) = F \sin(Kr + \delta_0)$  where ( $C = F \cos \delta_0$ ;  $D = F \sin \delta_0$ )

As with the 1D square well, the boundary conditions are that  $u$  and  $\frac{du}{dr}$  are continuous at  $R$ , which lead to:

$$K \tan K_0 R = K_0 \tan(KR + \delta_0) \quad \text{or} \quad \delta_0 = \tan^{-1} \left( \frac{K}{K_0} \tan K_0 R \right) - KR$$

In the low energy case  $KR \ll 1$ , we obtain maximum scattering ( $\sin^2 \delta_0 \rightarrow 1$ ) when  $K_0 R = (n + \frac{1}{2})\pi$ , when the scattering cross section is  $\sigma = 4\pi/K^2$ . This is an example of *s-wave resonance*.

In the same slow particle limit  $K \ll K_0$ , and assuming that  $\tan K_0 R$  is not very large:  $\delta_0 \approx \sin \delta_0$ .

$$\sigma \approx 4\pi R^2 \left( \frac{\tan K_0 R}{K_0 R} - 1 \right)^2$$

This correctly predicts that when  $\tan K_0 R = K_0 R$  the scattering cross section will be zero.

There are a few features of the square-well which also apply in more general cases. Assuming  $K_0$  is basically a measure of the potential depth.

- For weak coupling  $K_0 R \ll 1$ ,  $\delta_0(K) \rightarrow 0$  as  $K \rightarrow 0$
- When  $K_0 R$  approaches  $\pi/2$  the potential is almost able to bind an *s-wave* bound state. Now the phase shift  $\delta_0(K) \rightarrow \pi/2$  and the cross section *diverges* like  $K^{-2}$  as  $K \rightarrow 0$ . This is known as zero energy resonance.
- If  $E$  is high enough that  $\delta_l = (n + \frac{1}{2})\pi$  for  $l \neq 0$  the scattering cross section can become especially high due to another angular momentum component - *p-wave* resonance for  $l = 1$ , *d-wave* resonance for  $l = 2$  etc. In these cases the eigenfunction becomes large near to the potential. The potential is said to have *virtual states* at the resonance energies.

- *Levinson's Theorem* states that

$$\lim_{k \rightarrow 0} \delta_l(k) = n_l \pi$$

where  $n_l$  is the number of bound states with angular momentum  $l$ .

- Whenever  $\delta_0(K) = n\pi$ , for  $s$ -wave scattering,  $\sigma = 0$ . Thus for certain energies of the incoming particle, the scattering is extremely small. This condition can only be consistent with the condition for  $s$ -wave scattering ( $KR \ll 1$ ) if the potential is attractive ( $V_0 < 0$ ).
- $\delta_0(K)$  tends to decrease with increasing  $K$ . This can be understood physically as the faster particles having less time to interact and thus experiencing smaller phase shifts. As  $K \rightarrow \infty$ ,  $\delta_l(K) \rightarrow 0$  because the potential is now weak relative to the particle energy. Of course  $\sigma(K \rightarrow \infty)$  decreases even more quickly because of the  $K^{-2}$  term.

## 11.5 Partial Waves in the Classical Limit - Hard Spheres

Consider the scattering of a small hard sphere (radius  $x_m$ , mass  $m$ ) by a large hard sphere ( $X_M$ ,  $M$ ). Firstly we transform the problem to the centre of mass reference frame where it becomes that of a single effective particle of mass  $\mu = mM/(m + M)$  moving in a hard sphere potential ( $V(r < r_H = X_M + x_m) = \infty$ ). Thus the boundary condition is  $R_l(r_H) = 0$ .

Consider the classical limit, where the sphere radius is much larger than the de Broglie wavelength,  $kr_H \gg 1$ . Up to  $l = Kr_H$  the phase shift is enormous and  $\sin \delta_l$  could have any value. For  $l > Kr_H$  the impact parameter is so large that the particles miss and  $\delta_l = 0$ . Thus we can write the scattering cross section:

$$\sigma = \frac{4\pi}{K^2} \sum_{l=0}^{l=Kr_H} (2l+1) \frac{1}{2}$$

where we replace  $\sin^2 \delta_l$  with its average value of  $\frac{1}{2}$ .

Since  $Kr_H$  is large, we can replace the sum by an integral and take only the leading term;  $(Kr_H)^2 \gg Kr_H$ :

$$\sigma \approx \frac{2\pi}{K^2} \int_{l=0}^{l=Kr_H} (2l+1) dl \approx 2\pi r_H^2$$

This result should send us rushing back to look for the extra factor of 2, since the cross-section of a sphere might be expected to be  $\pi r_H^2$ . In fact, though, the analysis is correct and closer analysis of the  $\theta$  dependence of the wavefunction shows that half the amplitude is diffracted into the classical 'shadow' of the sphere to cancel the amplitude of the unscattered wave there.

## 11.6 Ramsauer-Townsend effect

This is the name given to the fact that electrons with energy about 1eV can pass almost freely through Xe, Kr, and Ar:- there is a sharp minimum in electron scattering cross-section for these noble gases.

Due to polarisation of these atoms by the incoming electron the potential appears to increase as  $K$  increases (more localised electrons are better able to polarise the atom). Thus  $\delta_0(k \rightarrow 0) = n\pi$ , in accordance with Levinson's theorem, and  $\delta_0$  initially increases as  $k$  increases, before eventually decreasing. Thus at a certain value of  $k$ , the phase shift is again  $\delta_0(k) = n\pi$ , and the total scattering cross section  $\sigma_T$  has an abrupt minimum. Although there are subsequent  $s$ -wave minima at e.g.  $\delta_0(k) = (n-1)\pi$ , these occur at sufficiently large values of  $k$  that  $s$ -wave scattering is no longer dominant.

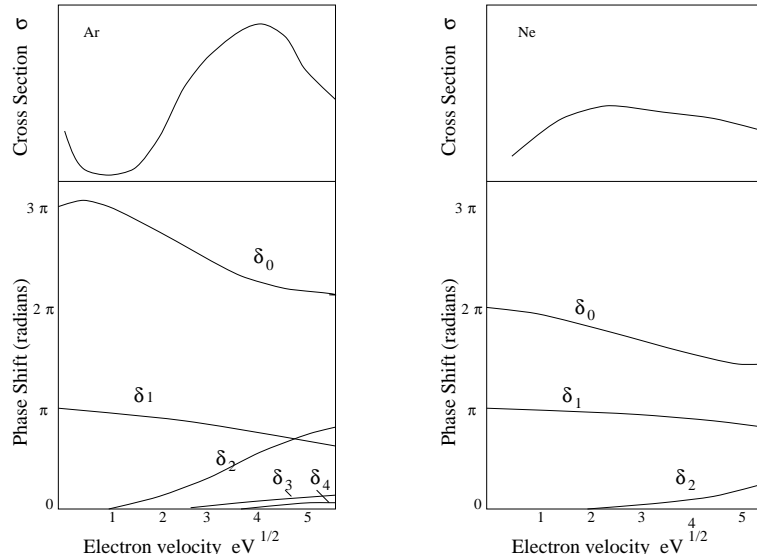


Figure 14: Minimum in scattering cross section in Ar due to  $\delta_0 = 3\pi$ ; No such effect in Ne due to weaker polarisation.

By contrast, neon and helium have lower polarisability, due to fewer bound electrons. Thus the phase shift  $\delta_0$  decreases monotonically with  $k$  from  $n\pi$  at  $k = 0$  at there is no low-energy minimum. Higher  $l$  phase shifts may increase with  $k$  because higher  $k$  implies smaller impact parameter (classically, more chance of hitting the atom). The cross section increases more slowly due to the additional  $K^{-2}$  dependence. The maximum in the Ar cross section at about 13eV is mainly due to the  $d$ -wave  $\delta_2 = \pi/2$ .

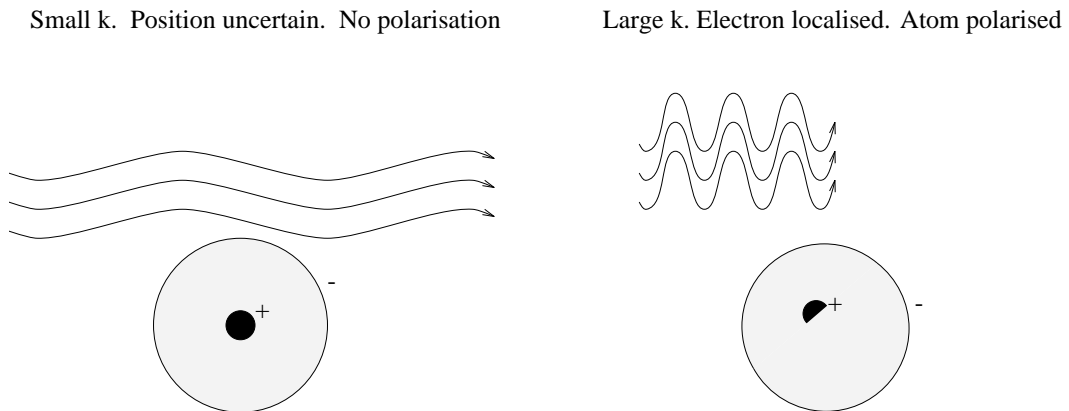


Figure 15: More localised electrons polarise atoms and thus increase the attractive potential

## 12 Further Concepts in Quantum Mechanics

### 12.1 Born Series - A Hint of Quantisation of the Field

Solving the Schroedinger equation using Green Functions automatically gives a solution in a form appropriate for scattering. Although this method is not directly useful for problems at the level of this course, it provides an alternate insight into scattering processes and hints at the idea of quantisation of the potential field.

By making the substitution  $E = \hbar^2 k^2 / 2\mu$  and  $U(r) = (2\mu/\hbar^2)V(r)$  we can write the TISE as:

$$[\nabla^2 + k^2]\Phi = U(r)\Phi$$

For  $U(r) = 0$  the solution of this differential equation (complementary function) is simply  $\phi_0(r) = Ae^{ik \cdot r}$ , a travelling wave. We also introduce a 'Green's Function' for the operator  $[\nabla^2 + k^2]$ , which is the solution to the equation:

$$[\nabla^2 + k^2]G(r) = \delta(r) \qquad G(r) = -\exp(ikr)/4\pi r$$

$\delta(r)$  is the Dirac delta-function.  $G(r)$  has the property that for any function  $\Phi$  which satisfies

$$\Phi(r) = \phi_0(r) + \int G(r - r')U(r')\Phi(r')d^3r'$$

where  $\phi_0(r)$  is the free particle solution, will be a solution to the TISE. Since  $\phi_0(r)$  is the unscattered incoming wave, the second term must represent the scattered wave.

Thus the general solution to the TISE is given by:

$$\Phi(r) = Ae^{ik \cdot r} + \int G(r - r')U(r')\Phi(r')d^3r'$$

In this expression,  $\Phi$  appears on both sides. We can substitute for  $\Phi$  using the same equation:

$$\Phi(r) = Ae^{ik \cdot r} + \int G(r - r')U(r')Ae^{ik \cdot r}d^3r' + \int \int G(r - r')U(r')G(r' - r'')U(r'')\Phi(r'')d^3r'd^3r''$$

Repeated substitutions gives the *Born series*.  $\Phi$  still appears on both sides. The first term is linear in the potential, the second term quadratic, and by further substitutions we can get cubic, quartic terms etc. If the potential is weak, these higher order terms can be ignored. The first order term is just the matrix element between the incoming plane wave and the Green function. Taking only the first order term is just the first Born approximation again!

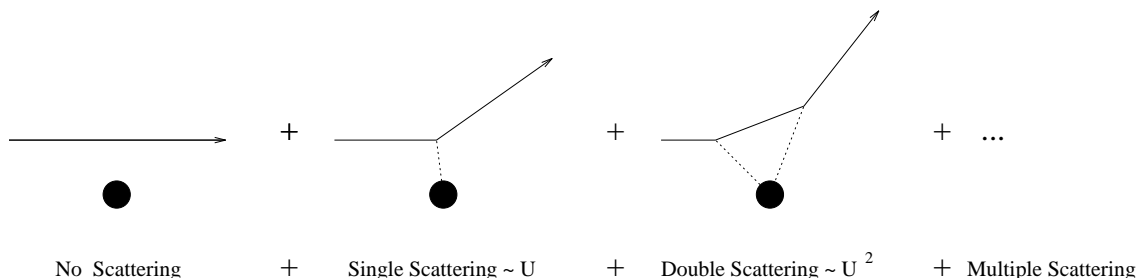


Figure 16: Born Series - scattering as series of terms

If we think of the potential  $U$  as an operator, the first term represents the incoming wavefunction being operated on once. The second term represents the incoming wavefunction being operated on twice. And so forth. This suggests a way of *quantising* the effect of the field: The first order term corresponds to a single scattering event, the second order term to double scattering etc.

## 12.2 Identical Particles

Quantum mechanics allows us to predict the results of experiments. If we conduct an experiment with indistinguishable particles a correct quantum description cannot allow anything which distinguishes between them.

For example, if the wavefunctions of two particles overlap, and we detect a particle, which one is it? The answer to this is not only that we don't know, but that we can't know. Quantum mechanics can only tell us the probability of finding *a* particle in a given region. The wavefunction must therefore describe *both* particles. The Hamiltonian is then:  $-\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) + V(r)$  where the subscripts label each particle, and there are six coordinates, three for each particle. The solution is  $\Phi(x_1, y_1, z_1, x_2, y_2, z_2)$  a wave in six dimensions! This should contain the information we can measure: the probability of finding a particle.

What basis states would be appropriate for  $\Phi$ ? We could use products of one particle wavefunctions such as  $|a(\mathbf{r}_1)b(\mathbf{r}_2)\rangle$ , where  $a(\mathbf{r}_1)$  and  $b(\mathbf{r}_2)$  are one-particle wavefunctions of atoms 1 and 2, which allow us to separate the two particle equation into two one particle equations:

$$\left[\frac{-\hbar^2}{2m}\nabla_1^2 + V(r_1)\right]|a(\mathbf{r}_1)\rangle = E_1|a(\mathbf{r}_1)\rangle; \quad \left[\frac{-\hbar^2}{2m}\nabla_2^2 + V(r_2)\right]|b(\mathbf{r}_2)\rangle = E_2|b(\mathbf{r}_2)\rangle$$

provided that there is no interaction between the particles (see exercises and example of helium in Atomic and Molecular Physics).

Note that  $\nabla_1^2$  does not act on  $b(\mathbf{r}_2)$ . This allows the separation. Unfortunately, by doing this we have introduced unphysical *labels* to the indistinguishable particles. And this is wrong: the effect of it is that the particles do not interfere with each other because they are in different dimensions (six dimensional space - remember?).

When we construct a two particle wavefunction out of two one-particle wavefunctions we must ensure that the probability density (The *measurable* quantity  $|\Phi|^2$ ) is independent of the *artificial* labels.

We introduce the *exchange operator*  $\hat{P}_{12}$ : an operator which permutes the labels of the particles. This is a rather strange operator, because it only changes the *unphysical* labels which we have attached to the one-particle wavefunctions in order to make the maths more easy. For a meaningful solution we must have a wavefunction which has a probability amplitude unchanged by  $\hat{P}_{12}$ . The wavefunction must be symmetric or antisymmetric with respect to exchange:  $|\Phi(\mathbf{r}_1, \mathbf{r}_2)\rangle = \pm|\Phi(\mathbf{r}_2, \mathbf{r}_1)\rangle$ .

Physical solutions must not only be eigenfunctions of the Hamiltonians, but also be eigenfunctions of  $\hat{P}_{12}$  with eigenvalues  $\pm 1$ .  $\langle\Phi|\Phi\rangle = \langle\hat{P}_{12}\Phi|\hat{P}_{12}\Phi\rangle$ . Also, the Hamiltonian must commute with  $\hat{P}_{12}$ , otherwise  $\hat{H}$  and  $\hat{P}_{12}$  could not have common eigenfunctions.

A simple product  $|a(\mathbf{r}_1)b(\mathbf{r}_2)\rangle$  will not satisfy this (unless  $a=b$ ). A linear combination of all permutations is required, for two particles:

$$|\Phi^-\rangle = |a(\mathbf{r}_1)b(\mathbf{r}_2) - a(\mathbf{r}_2)b(\mathbf{r}_1)\rangle/\sqrt{2}$$

$$|\Phi^+\rangle = C_{ab}|a(\mathbf{r}_1)b(\mathbf{r}_2) + a(\mathbf{r}_2)b(\mathbf{r}_1)\rangle + C_{aa}|a(\mathbf{r}_2)a(\mathbf{r}_1)\rangle + C_{bb}|b(\mathbf{r}_2)b(\mathbf{r}_1)\rangle$$

where the  $C_{ab}$  terms are expansion and normalisation parameters. Note that the antisymmetric combination cannot include terms where both particles are in the same state, and there are three possibilities for the symmetric state: most simply  $C_{ab} = 1$ ,  $C_{bb} = 1$  or  $C_{aa} = 1$  although any linear combinations are possible. If the  $a$  and  $b$  represent spin  $\frac{1}{2}$  states then  $|\Phi^-\rangle$  is a singlet and the three  $|\Phi^+\rangle$ s form a triplet, which will be degenerate if  $a$  and  $b$  are.

### 12.3 The Pauli exclusion principle and exchange interaction

This states that wavefunctions describing systems of many fermions must be antisymmetric with respect to exchange, i.e.  $\hat{P}_{12}|\Phi\rangle = -|\Phi\rangle$ . Note that in an atom or molecule  $\Phi$  includes both spin and spatial parts, even though the Hamiltonian (Coulomb potential) acts only on the spatial part.

The expectation value of the potential energy is different for symmetric and antisymmetric combinations. Using  $|\Phi^\pm\rangle$  from above (with  $C_{ab} = 1$ ).

$$\langle\Phi^\pm|\hat{V}|\Phi^\pm\rangle = \langle a(\mathbf{r}_1)b(\mathbf{r}_2)|V(r)|a(\mathbf{r}_1)b(\mathbf{r}_2)\rangle \pm \langle a(\mathbf{r}_1)b(\mathbf{r}_2)|V(r)|a(\mathbf{r}_2)b(\mathbf{r}_1)\rangle$$

The first term is called the direct interaction and the second term is known as the exchange interaction, and is comparable in size to the first. For an  $N$ -particle system there are  $N$  direct terms and  $N(N - 1)$  exchange terms.

### 12.4 Scattering of distinguishable particles and identical particles

Consider two beams of distinguishable particles with the same mass colliding, and scattering through some angle  $\theta$ . Let the intensity of the scattered particles have angular dependence  $|f(\theta)|^2$ . Conservation of energy and momentum ensure that the scattering angles are the same for both particles in the COM frame. As usual, the radial part of the wavefunction far from the region of interaction is simply a plane wave so the wavefunction can be written as a function of  $\theta$ .

The intensity for the process in which both particles are scattered through an angle  $(\pi - \theta)$  is  $|f(\pi - \theta)|^2$ . Note that this process results in particles arriving in the same places as with  $f(\theta)$  - it is just the other particles (see diagram).

If the two particle beams are distinguishable they cannot interfere and differential cross section for either particle to be detected at  $\theta$  is:

$$I_{dis} = |f(\theta)|^2 + |f(\pi - \theta)|^2$$

If, however, the particles are indistinguishable boson(fermions), they can interfere and the combined wavefunction must (anti)symmetric under exchange of labels:

$$\Phi_{fer}^{bos} = f(\theta) \pm f(\pi - \theta) \qquad I_{fer}^{bos} = |f(\theta) \pm f(\pi - \theta)|^2$$

Taking the specific extreme example of scattering through  $\pi/2$ , the differential cross section is  $2|f(\pi/2)|^2$  for distinguishable particles,  $4|f(\pi/2)|^2$  for identical bosons, and 0 for identical fermions.

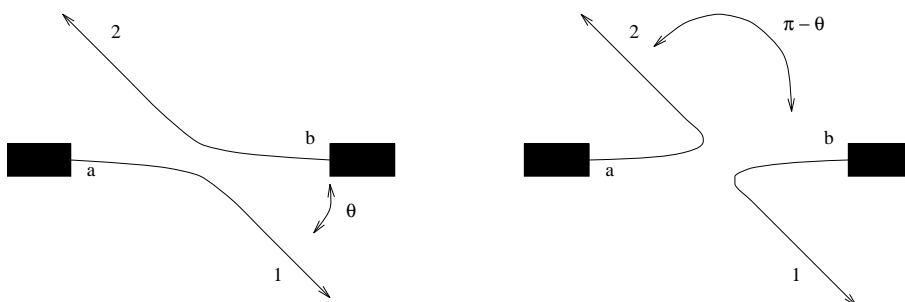


Figure 17: Two indistinguishable scattering processes.

## 12.5 Collision between two unpolarised electron beams

In this case, half the collisions will be between like-polarised electrons, so will involve interference, and half will be between unlike electrons: so there would be no interference. In both cases  $|f(\theta)|^2$  represents Rutherford scattering. The differential cross section of finding an electron scattered through an angle  $\theta$  is thus:

$$I = \frac{1}{2}(I_{dis} + I_{ind}) = \frac{1}{2}(|f(\theta)|^2 + |f(\pi - \theta)|^2) + \frac{1}{2}|f(\theta) - f(\pi - \theta)|^2$$

Consider  $\theta = \pi/2$ . The like polarised beams give zero probability, so unpolarised beams give only half what we would expect from Coulomb scattering of distinguishable particles. Furthermore, the spins of pairs of electrons scattered through  $\theta = \pi/2$  are always observed to be opposite.

An alternate philosophy is that we should treat the spins as a symmetric triplet and an antisymmetric singlet, with probabilities  $\frac{3}{4}$  and  $\frac{1}{4}$ . Then the spatial scattering process must be antisymmetric in the first case and symmetric in the second. This gives the same answer!

## 12.6 Scattering of particles into the same state

Consider scattering of two indistinguishable bosons by an external potential. We know that the wavefunction of this system must be *symmetric* with respect to exchange: *amplitudes* for indistinguishable boson processes add. Thus the cross section for scattering of both through the same angle is:  $|2f(\theta)|^2$

Thus two bosons are twice as likely to be scattered into the same state as two distinguishable particles. For many bosons the effect is even more pronounced, and the probability of scattering *out* of the state is reduced.

The tendency for bosons to clump into one state leads to superfluid behaviour in He<sup>4</sup> and superconductivity:  $\alpha$  particles and Cooper pairs behave as bosons. All the particles are in the same state and cannot be scattered out.

Note that for fermions, the cross section for being scattered into the same state is  $|f(\theta) - f(\theta)|^2 = 0$ , as we would expect from the exclusion principle.

## 12.7 Perturbation of identical free particles with a periodic potential

For a free particle moving in a 1D region of space there are two degenerate wavefunctions ( $\Phi = e^{\pm ikx}$ ). If there is a weak periodic potential,  $V \cos ax$ , to evaluate the energy shift to first order in degenerate perturbation theory the relevant matrix elements are:

$$\int e^{\pm ikx} V \cos ax e^{\mp ikx} dx = \int V \cos ax dx = 0; \quad \int e^{\pm ikx} V \cos ax e^{\pm ikx} dx = \int V \cos ax \cos 2kx dx$$

The second term is also zero, except in the case  $2k = a$ . This gives rise to the remarkable result:

*To first order, free particles are unaffected by a periodic potential unless it has half the wavelength.*

This goes some way towards explaining why free electron models work so well for solids. To second order, another non-degenerate term (Section 2.5) comes into play:  $\int e^{\pm i(a/2+\delta)x} V \cos ax e^{\pm i(a/2-\delta)x} \neq 0$  which gives a significant energy shift for small  $\delta$  where the energy difference between the states ( $\hbar^2 a \delta / m$ ) is small.

For a system containing many weakly bound fermions, the exclusion principle means that levels “fill up”. In a solid, typically, there will be more electrons than ions and so states around  $2k = a$  will be occupied. Only where the Fermi Energy is coincident with  $2k = a$  then the splitting structure can lower the total energy since the lower level is occupied but the higher is not.

## 13 The quantum mechanics nobody really understands

### 13.1 Diffraction from two slits

Interference from two slits of a single particle with itself remains a difficult concept to understand.

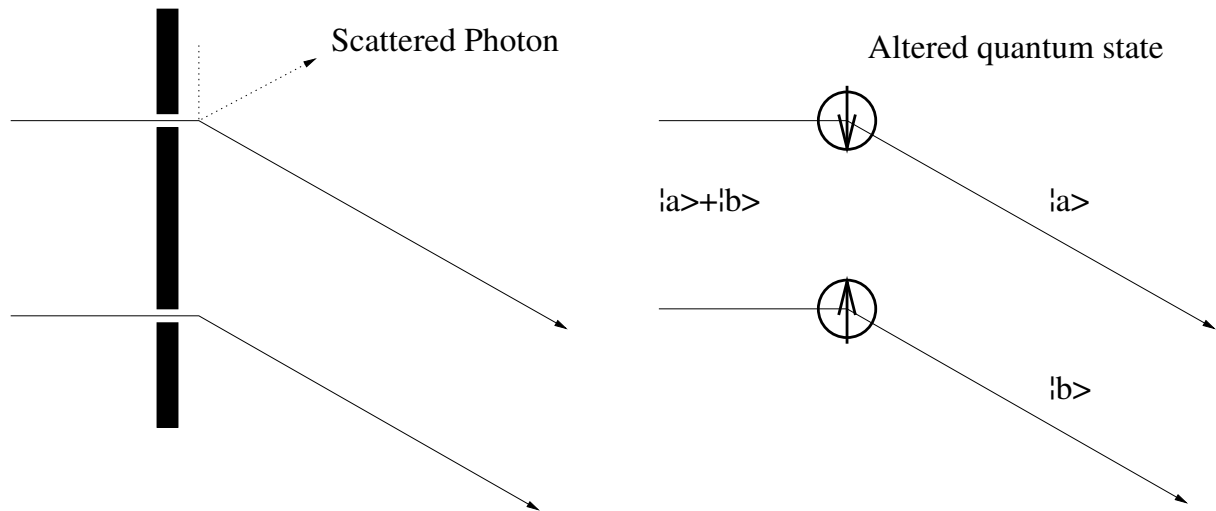


Figure 18: Feynman's 'classical' explanation of the destruction of the interference pattern by measurement, and two separate demonstrations that it is really a quantum effect

Feynman introduced a nice argument based on the uncertainty principle. He argued that the wavelength of light required to detect which way a particle went must be smaller than the slit separation. From the uncertainty principle, it follows that the momentum transfer must be so large that it would destroy the interference pattern. Thus the measurement device destroyed the interference. Unfortunately, more recent experiments show things are more complicated than that.

Eichmann *et al* (Phys.Rev.Lett, 1993) set up a 'two slit' experiment using photon with lead atoms as the scatterers. With careful choice of energy, he was able to arrange that the scattering event changed the internal electronic state of the atom: a process which requires negligible momentum transfer but would allow subsequent measurement of the atomic state and determination which way the particle went. As a consequence, the interference fringes vanish.

Durr *et al* (Nature, 1998) used a standing light wave to scatter rubidium atoms. Added to this was a microwave source which changed the hyperfine state of the atoms, which could in principle be measured but supplies negligible momentum. Again the interference pattern was washed out.

Again, quantum mechanics has been shown to give a correct description non-identical wavefunctions do not interfere even if they describe the same particle! It does not matter whether the measurement of the internal states is actually performed: the mere fact that it could be is enough to destroy the interference.

## 13.2 What does quantum mechanics mean: Hidden Variables

The interpretation of collapsing wavefunctions is often regarded as unphysical, or philosophically problematic. There appears to be a contradiction with relativity in the idea that the wavefunction collapses instantaneously throughout space, although the wavefunction is not measurable.

An attractive contrary view to the idea of ‘measurement collapsing the wavefunction’ is that for a particular system the value of a observable is a property of the particle, and the wavefunction only expresses averages over many particles. This kind of property is known as a hidden variable. As we shall see, this interpretation of quantum mechanics can be tested, and is inconsistent with experimental results.

Consider a two-photon decay from a source (e.g.  $^{40}\text{Ca}$ ). Two polarisers are oriented along the z-direction, and we detect whether or not the photons pass through the polariser.

The decay is one in which angular momentum is conserved, so the photons must be either both right-polarised ( $\mathbf{e}_R$ ) or both left-polarised ( $\mathbf{e}_L$ ) (they travel in opposite directions). We are dealing with bosons, so the wavefunction can be written as a superposition:

$$|12\rangle = \sqrt{\frac{1}{2}} (\mathbf{e}_{1R}\mathbf{e}_{2R} + \mathbf{e}_{1L}\mathbf{e}_{2L})$$

Now convert into x and y polarisation using  $\mathbf{e}_R = (\mathbf{e}_x - i\mathbf{e}_y)$  and  $\mathbf{e}_L = (\mathbf{e}_x + i\mathbf{e}_y)$  to give

$$|12\rangle = \sqrt{\frac{1}{2}} (\mathbf{e}_{1x}\mathbf{e}_{2x} + \mathbf{e}_{1y}\mathbf{e}_{2y})$$

From this we can clearly see that the quantum probability of the photon 1 passing through its detector is  $\frac{1}{2}$ , and if so the wavefunction collapses onto  $|12\rangle = \mathbf{e}_{1x}\mathbf{e}_{2x}$  and the conditional probability of the second photon passing through its detector is then 1. Thus quantum mechanics tells us that the probability of both detectors counting is  $\frac{1}{2}$ .

Contrariwise, a hidden variables argument might say that on production the photons were polarised in a random direction, say  $\theta$  to the x-axis. In this case the probability of passing through either detector would be  $\cos^2\theta$ , and the probability of simultaneous counts will be  $\langle \cos^4\theta \rangle = 3/8$ . The mathematics for particles with correlated spins is similar.

Since the wavefunction collapse and hidden variable approach give different answers, we can do an experiment to see which is correct.

## 13.3 Bell’s Inequality and Aspect’s experiment

Consider extending the experiment described above to the case of analysers at arbitrary angles which detect all photons. We define measurables  $a(\theta)$  and  $b(\phi)$  as +1 if the photon is aligned with the analyser and -1 if it is opposed. What, then, is the ensemble average value of  $P(\theta, \phi) = \langle a(\theta)b(\phi) \rangle$ ? Clearly, if  $a(\theta)$  and  $b(\phi)$  are uncorrelated  $P=0$ , but since they come from a common source, this is not the case: their wavefunctions are sometimes referred to as ‘entangled’.

If the photons start out with ‘hidden variable’ polarisation  $\chi$ , then it is easily shown that:

$$P_{HV}(\theta, \phi) = \frac{1}{2\pi} \oint (\cos^2(\theta - \chi) - \sin^2(\theta - \chi)) (\cos^2(\phi - \chi) - \sin^2(\phi - \chi)) d\chi = \frac{1}{2} \cos 2(\theta - \phi)$$

Meanwhile if the wavefunction collapses at the first measurement, taken arbitrarily as A:

$$P_{QM}(\theta, \phi) = \frac{1}{2\pi} \oint (\cos^2(\theta - \chi) - \sin^2(\theta - \chi)) (\cos^2(\theta - \phi) - \sin^2(\theta - \phi)) d\chi = \cos 2(\theta - \phi)$$

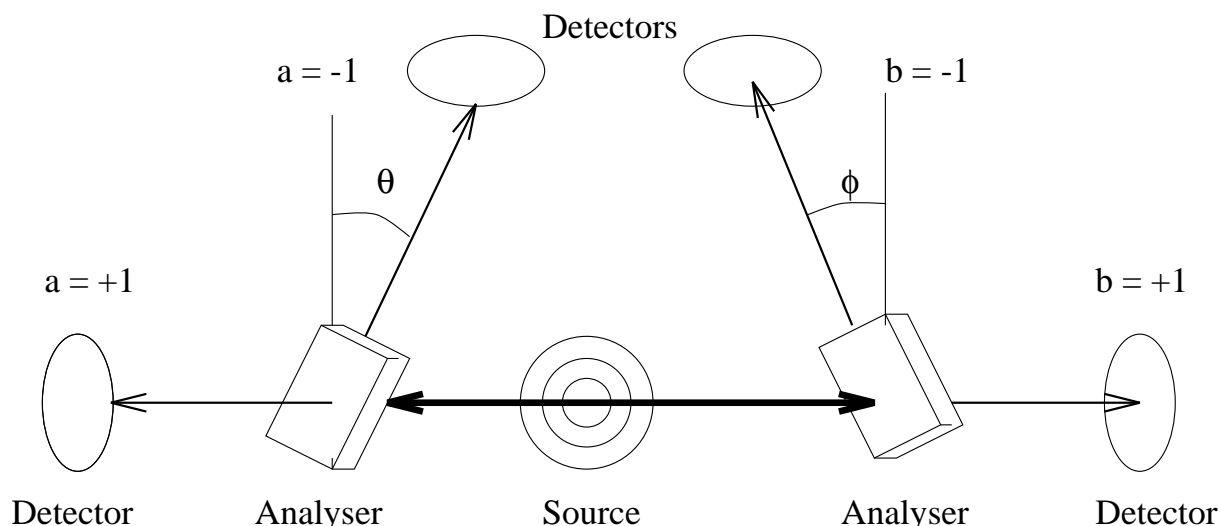


Figure 19: Aspect's Experiment: The polarisations of both photons from the two-photon  $^{40}\text{Ca}$  source are measured by analysers at angles of  $\theta$  and  $\phi$ .

In 1982, to test this Aspect carried out measurements on  $^{40}\text{Ca}$  decays using two different angles for both  $\theta$  and  $\phi$ . The quantity he evaluated was:

$$S(\theta_1, \phi_1, \theta_2, \phi_2) = P(\theta_1, \phi_1) + P(\theta_2, \phi_2) + P(\theta_2, \phi_1) - P(\theta_1, \phi_2)$$

Where he chose the values which give the largest S:  $\theta_1 = \phi_1 + \frac{\pi}{8} = \theta_2 + \frac{2\pi}{8} = \phi_2 + \frac{3\pi}{8}$

The hidden variables theory suggests the result should be  $S = \sqrt{2}$ , while the wavefunction collapse suggests  $S = 2\sqrt{2}$  with perfect measurement devices. Imperfections in the measurement will reduce the measured correlation in each case. Aspect measured  $S = 2.697 \pm 0.015$ , confirming the quantum prediction.

The apparent complexity of Aspect's experiment is needed to eliminate sources of error due to detector, analyser and source imperfections.

There is an apparent contradiction between quantum mechanics and relativity, in that the *interpretation* of quantum mechanics requires *instantaneous* collapse of the wavefunction. However, there is no *measurable* quantity for which the two theories give different predictions.

Most of the wavefunctions we have solved are from Schrodinger's equation, which treats time and space in different ways. For a properly relativistic approach, they should be equivalent. This discrepancy between quantum and relativity is easily resolved: the Dirac equation provides a fully relativistic wave equation for which the Schrodinger equation is a low energy approximation. A nice thing about the Dirac equation is it can only be solved by spinors: as with quantisation the observed physics turns out to be the only way to solve the mathematics.

## 13.4 Back to the beginning - waves on a string

There are a finite number of fundamental particles, with specific masses. Theorists are tempted to view these as quantised states of something still more fundamental: so called superstrings. In string theories the fundamental particles arise as standing waves of elementary strings: the lengthscale of which is the Planck length  $\approx 10^{-35}\text{m}$ .

Now the two types of particles, bosons and fermions, can emerge from the same theory provided they are degenerate in the sense (section 4) that an operator exists which acts on a boson and converts it to a fermions. This degeneracy is associated with a symmetry called supersymmetry.

This all looks very elegant, but when you put in the appropriate masses none of the known fermions map onto any own the known bosons. Also, the theory can only give the observed families of particles if framed in 10 dimensions (time plus a 9D space). This can be argued away by having all the supersymmetric ‘partner’ states be of too high energy to be observed, and the additional space dimensions fold up on themselves (compactification).

## 13.5 What we haven’t covered: Quantisation of Excitations

In this course we have dealt with quantum states of single particles. Although there is no time to go further, it is worth noting that quantisation is a far more general phenomena than this. Almost any excitation can be quantised. Often, similar to the simple harmonic oscillator, there is still some energy associated with the lowest quantum state. This zero point energy is real, and can be measured in the ‘Casimir effect’: there is a force between two metallic plates in a vacuum, because moving them would change the wavelength/energy of the zero-point quantised electromagnetic waves between them: this change in energy in response to a move equates to a force.

In solids, the quantum theory of phonons quantises the normal mode vibrations of a lattice, the quantum theory of magnons quantises the excitations in magnetic media, Quantum electrodynamics describes quantisation of electromagnetic fields, with interactions mediated by ‘lumps’ of electromagnetism (photons). Quantum chromodynamics describes quantisation of the strong force with interactions mediated by gluons. And the hunt is on for the graviton.

Finally, as quantum mechanics demands that the universe be described by discrete ‘lumps’ of stuff, people are now questioning whether space and time themselves should be immune from quantisation, or whether time ticks by in discrete intervals and space can only be divided into smallest possible lengths. This would wreck continuum theories like general relativity, but only at incredibly high energies and momenta.

All these have never yet been seen—  
But Scientists, who ought to know,  
Assure us that they must be so....  
Oh! let us never, never doubt  
What nobody is sure about!

—from The Microbe, by Hillaire Belloc

*“If it is correct, it signifies the end of physics as a science”* Albert Einstein on Quantum Mechanics