

ALAN H. BRIDLE

August 1962

NOTES ON

QUANTUM MECHANICS

Operator Theory

Uncertainty Principle

Schrödinger's Equation

Sturm-Liouville Theory

Matrix Mechanics

Perturbation Theory

Any object of physical enquiry which, in classical physics, is thought of as an entity as distinct from a mere property, will be termed a PHYSICAL SYSTEM. Such objects may be capable of direct observation (large-scale bodies, measuring devices, light-sources, etc.) or may indicate their presumed existence by a sequence of inferences based upon direct observations (molecules, atoms, photons, etc.)

With a system are associated a number of properties capable of measurement, such as position, energy, momentum, and the like. These will be called OBSERVABLES, although this is not intended to imply that they are observable directly. Every measurement of an observable yields a number. In classical physics this number was assigned to the observable as its only characterisation and referred to as its magnitude. Hence the notion of an observable was relatively unimportant because it was synonymous with the result of observation. In quantum mechanics we accord to this notion a more independent character inasmuch as we shall associate it with a more general construct (an operator) which determines numbers, but not a single number. The relation between these numbers and the results of measurements performed on an observable will form the subject of a later discussion.

We now define the state of a physical system. Classical physics used several unrelated definitions of state, the principal one being the designation of all momenta and all co-ordinates of the particles composing the system, amounting to an enumeration of $2s$ numbers, if the system has s degrees of freedom. In our present definition we renounce all attempts of defining the configuration of parts and take a function ϕ of the s co-ordinates as representing the state. In quantum mechanics the state of a system is thus no longer defined by a number of variables having an immediate intuitive appeal and recalling exact configurations of constituent parts; it is simply a function in configuration space.

The laws of physics in general connect states; in quantum mechanics they will be laws connecting ϕ -functions at different instants of time.

The independent variables which ϕ contains as arguments are the co-ordinates defining the degrees of freedom of the system. This number of degrees of freedom must be considered as arrived at by trial and error, without reference to the ideas of classical mechanics. They will naturally in general turn out to be the degrees of freedom of classical mechanics, but need not be regarded as such from an axiomatic point of view.

We now postulate some restrictions upon the nature of the ϕ -functions. ϕ , which may in general be a complex function (ϕ^* denotes its complex conjugate), must be quadratically integrable, so that

$$\int \phi \phi^* d\Gamma \text{ exists. (A)}$$

This integration is to be extended over the fundamental domain of all the variables of which ϕ is a function. $d\Gamma$, therefore, represents the element of configuration space. If the fundamental domain of any of the variables extends to infinity, (A) implies in general that ϕ should vanish at infinity in a suitable manner. (Functions which do not vanish at infinity but have an integrable square have hitherto played no part in quantum mechanics). Condition (A), although necessary, is not always sufficient. The second condition must be

$$\phi \text{ is single-valued. (B)}$$

Cases in which this condition is important arise whenever one of the variables in ϕ is an angle, for then (B) requires that $\phi(2\pi + \beta) = \phi(\beta)$.

A notion which proves very useful in quantum mechanical discussions is that of

an operator, or a mathematical operation represented by a characteristic symbol.

We must now set up the relationship between states and experience. We begin by asserting that to every observable p there corresponds an operator P . In order to ascertain the correct operator which is to be associated with a given observable we must rely on trial. One of the principal concerns for the new discipline is to develop operators for all physical observables. The table which follows gives an assignment which is sufficient for the solution of many physical problems.

p	P
q_k	q_k
p_k	$\frac{h}{2\pi i} \cdot \frac{\partial}{\partial q_k}$
m_k	$\frac{h}{2\pi i} \cdot \frac{\partial}{\partial q_1} - q_m \frac{\partial}{\partial q_1} \dots$
$H(p, q)$	$H \left[\frac{h}{2\pi i} \cdot \frac{\partial}{\partial q}, q \right]$

By q_k is meant any Cartesian co-ordinate q ; by p_k is meant the associated momentum.

When it is necessary to express a momentum operator or a combination of momentum operators in some other system of co-ordinates (e.g. polar co-ordinates) it is best to write the combination first in its Cartesian form and then to transform the differential expression in the usual manner. m_k denotes the observable "angular momentum". One remark should be made in connexion with the observable "energy". $H(p, q)$, the energy in its Hamiltonian form, is usually simply (kinetic + potential). This is not Lorentz-invariant, but it is easy to find a classical form which satisfies relativity requirements. The computation of the corresponding invariant operator, however, does not proceed by the rules outlined above.

In classical mechanics q_k and p_k are known as canonically conjugate variables. The corresponding operators obey

$$P_k Q_k - Q_k P_k = \frac{h}{2\pi i}.$$

These operators are said to be canonically conjugate operators, and the same term is applied to all operators with this property.

It is not true that every observable has its own unique operator. The form of the operator will in general also depend on the nature of the system in question. Thus the energy operator for a single point mass M not subject to forces is

$$-\frac{h^2}{8\pi^2 M} \cdot \nabla^2$$

If it has a potential energy of the form $V(x, y, z)$, this term has to be added (unmodified because the operator for a co-ordinate is simply the co-ordinate itself); the energy operator for two mass-points M_1 and M_2 is

$$-\frac{h^2}{8\pi^2} \left[\frac{1}{M_1} \cdot \nabla_1^2 + \frac{1}{M_2} \cdot \nabla_2^2 \right] + V(x_1, y_1, z_1, x_2, y_2, z_2)$$

where ∇_1^2 is the Laplacian involving the co-ordinates of the first mass-point, and ∇_2^2 contains those of the second.

The postulate of p-P correspondence is the first of our postulates, and will be denoted as (I).

Postulate (II) is that the only possible values which a measurement of the observable p can yield are the eigenvalues p_λ of the equation

$$P\psi_\lambda = p_\lambda \psi_\lambda, \text{ where } \psi_\lambda \text{ satisfies (A) and (B).}$$

If P is a differential operator, then this is a differential equation. Such an equation has solutions for any value of p_λ so that it alone cannot be instrumental in selecting a certain set. It can only do this if boundary conditions have to be satisfied.

The correspondence of this postulate to the facts can be decided only on the basis of experiment. The spectrum of eigenvalues generated by (II) has always been found to be the one observed.

The functions ψ_λ are the eigenfunctions of the operator P and define states as described previously. The states to which they refer are called eigenstates. We anticipate two mathematical properties of these ψ_λ , to be proved for special cases later on. These are

- (1) Orthogonality $\int \psi_\lambda^* \psi_\mu \cdot dT = 0, \text{ if } \lambda \neq \mu$
- (2) Completeness $\phi = \sum_\lambda a_\lambda \psi_\lambda$ where the a_λ are constants.

To be precise, it must be added that orthogonality as given above may break down when there is more than one eigenfunction ψ_λ belonging to the same eigenvalue p_λ (degeneracy). Even then it is possible to make these functions orthogonal by choosing linear combinations, so that the generality of the idea is not impaired.

A few examples of physically useful operator equations and their solutions are given below:

a) Let P in (II) represent the operator x. The resulting equation then represents the results of possible x-co-ordinate measurements performed on the system, which might be a charged particle, for example. We then have an algebraic equation (II) whose form is

$$x\psi_\lambda(x) = \beta_\lambda \psi_\lambda(x), \text{ where } \beta_\lambda \text{ is a number.}$$

The solution is obvious if we write (II) as $(x - \beta_\lambda)\psi_\lambda(x) = 0$. The first factor is different from zero if $x \neq \beta_\lambda$, so $\psi_\lambda(x)$ must be equal to 0 unless $x = \beta_\lambda$. At this point it may assume any value. The construct in question is Dirac's δ -function, and while it is an improper function, we can define it accurately as

$$\delta(q) = 0 \text{ if } q \neq 0, \delta(q) = \delta(-q), \text{ and } \int \delta(q) dq = 1$$

The last relation amounts to a definition of the value of the function at $q=0$. If we identify the argument q of δ with $(x-\beta_\lambda)$, we have what is equivalent to the eigenfunction of P, corresponding to the eigenvalue β_λ , viz.

$$\psi_\lambda = \delta(x - \beta_\lambda)$$

The eigenvalues β_λ are in no way restricted, ψ_λ satisfies (A) and (B), and so we conclude that any value of x is observable: the operator has a continuous spectrum.

b). The eigenvalues of the linear momentum operator P_x are given by the operator equation

$$\frac{\hbar}{2\pi i} \cdot \frac{d\psi_\lambda}{dx} = p_\lambda \psi_\lambda$$

This has the solution $\psi_\lambda = \text{constant} \cdot \exp\left\{\frac{2\pi i}{\hbar} \cdot p_\lambda x\right\}$

To determine the possible values of p_λ it is necessary to consider the fundamental range of x . Let this be the domain $(-L, L)$. Conditions (A) and (B) then merely require that

$$\int_{-L}^L \psi_\lambda^* \psi_\lambda \cdot dx = 2 \cdot |\text{constant}|^2 \cdot L$$

shall exist, and this is evidently true if both L and the constant are finite. Hence we again find that all values of p_λ may occur.

c). The eigenvalues of the angular momentum operator. These, for a classical rotation of a mass-point in the xy -plane, are given by the equation

$$\frac{\hbar}{2\pi i} \cdot x \cdot \frac{\partial}{\partial y} - y \cdot \frac{\partial}{\partial x} \cdot \psi_\lambda = m_\lambda \psi_\lambda$$

If we introduce polar co-ordinates, putting $x = r \cdot \cos\theta$, $y = r \cdot \sin\theta$, we have

$$\frac{d}{d\theta} = -r \cdot \sin\theta \cdot \frac{\partial}{\partial x} + r \cdot \cos\theta \cdot \frac{\partial}{\partial y} = x \cdot \frac{\partial}{\partial y} - y \cdot \frac{\partial}{\partial x}$$

The equation is then $\frac{\hbar}{2\pi i} \cdot \frac{d}{d\theta} \cdot \psi_\lambda = m_\lambda \psi_\lambda$

It has the solution $\psi_\lambda = \text{constant} \cdot \exp\left\{\frac{2\pi i}{\hbar} \cdot m_\lambda \theta\right\}$

Although any value of m_λ will cause ψ_λ to satisfy (A), for (B) we now require

$$\psi_\lambda(\theta) = \psi_\lambda(2\pi + \theta)$$

Hence $\exp\left\{\frac{2\pi i}{\hbar} \cdot m_\lambda \theta\right\} = \exp\left\{\frac{2\pi i}{\hbar} \cdot m_\lambda (\theta + 2\pi)\right\}$, so that $\exp\left\{\frac{4\pi^2 i}{\hbar} \cdot m_\lambda\right\} = 1$

This can only be true if $m_\lambda = \frac{\hbar \lambda}{2\pi}$, where λ is an integer. This is simply the Bohr condition for the angular momentum of an electron.

d). The eigenvalues of the energy operator. These are given by the equation

$$H\psi_\lambda = E_\lambda \psi_\lambda$$

Owing to the difference in the structure of this operator the equation takes on a variety of forms depending on the nature of the system to which H refers. Usually, however, it will be a second-order equation in which the Laplacian figures, and the type of equation is generally referred to as Schrödinger's Equation. The detailed solution of special forms of Schrödinger's equation will be entered into later.

At this juncture the remaining fundamental postulate will be given.

This is that, when a given system is in a state ϕ , the expected mean of a sequence of measurements on the observable p is

$$(III). \quad \bar{p} = \frac{\int \phi^* P \phi . dT}{\int \phi^* \phi . dT}, \text{ where } P \text{ is the operator corresponding to } p.$$

It is implied that measurements of p form a probability aggregate, the multiple repetition of the measurement involved in this idea occurring in either space or time; in the first case it (\bar{p}) is the mean of the numerical results of a great number of observations conducted simultaneously on similar systems in the same state, in the second it is the mean of a great number of experiments on the same system in the state in question. Most physical measurements explained by quantum mechanics are of the first type, so that \bar{p} is practically always understood in the sense of an average over the results of simultaneous observations. Postulate (III) may be called the "mean value relation", this terminology being suggested by its formal resemblance to the usual expression for an expected mean. (III) implies further conditions on state functions; we now have a criterion for examining ϕ -functions as to their suitability for any particular physical problem at hand. If we want to express a state with a finite energy, or a finite momentum, we must choose state-functions for which not only $\int \phi^* \phi . dT$, but also $\int \phi^* H \phi . dT$, etc. exist.

Postulate (III), which alone forms the connexion between observable quantities and state-functions, allows an interesting conclusion. If ϕ were multiplied by any constant, the value of \bar{p} would be unchanged. Hence ϕ is arbitrary to within a constant multiplier. We may therefore choose the constant to make condition (A) become:

$$\int \phi^* \phi . dT = 1$$

(III) then takes the form:

$$\bar{p} = \int \phi^* P \phi . dT$$

Henceforth every ϕ -function will be assumed to be normalised in this way. The condition (A) may then be combined with the orthogonality relation in the form:

$$\int \psi_\lambda^* \psi_\mu . dT = \delta_{\lambda\mu}$$

It will be observed that this procedure has not yet removed all arbitrariness from the choice of ϕ 's. We are still entitled to multiply ϕ by any constant factor $\exp(i\theta)$, so that state-functions differing from one another by a constant of absolute value unity are therefore empirically indistinguishable.

Some Useful Theorems

1). The formal "mean" of the operator P , viz. $\int \phi^* P \phi . dT$, is equivalent to an average over all possible eigenvalues of P with weights depending upon the function ϕ (i.e. the state of the system).

Proof:- Expand ϕ as $\phi = \sum_\lambda a_\lambda \psi_\lambda$, and choose ψ_λ to be an eigenfunction of the operator in question, so that $P\psi_\lambda = p_\lambda \psi_\lambda$.

$$\text{Then: } \int \phi^* P \phi . dT = \sum_\lambda a_\lambda^* a_\lambda p_\lambda$$

Then we have:
$$\int \phi^* \phi \cdot dT = \int \sum_{\lambda} a_{\lambda}^* \psi_{\lambda}^* \cdot P \cdot \sum_{\mu} a_{\mu} \psi_{\mu} \cdot dT = \sum_{\lambda} \sum_{\mu} \int \psi_{\lambda}^* P \psi_{\mu} \cdot dT \cdot a_{\lambda}^* a_{\mu}$$

Now, the integral reduces to p_{λ} when $\lambda = \mu$, otherwise it is zero, as the ψ are orthogonal. There thus remains a single sum instead of a double sum, and we have:

$$\bar{p} = \int \phi^* \phi \cdot dT = \sum_{\lambda} a_{\lambda}^* a_{\lambda} \cdot p_{\lambda}, \text{ thus proving the theorem.}$$

This result suggests an interesting consequence: can we regard the positive coefficients $a_{\lambda}^* a_{\lambda}$ as defining the frequencies with which the various possible values p_{λ} occur in our measurements? This is not an obvious consequence of the above and (III), for there are many different ways of choosing these coefficients, but only one set can define the true probability distribution.

2). The coefficients $|a_{\lambda}|^2$ define the probability distribution of the p_{λ} .

Proof:- Suppose the set of numbers w_{λ} defines the true probability distribution of the p_{λ} 's. Then we have:

$$\sum_{\lambda} w_{\lambda} = 1, \quad \sum_{\lambda} w_{\lambda} p_{\lambda} = \bar{p}, \quad \sum_{\lambda} w_{\lambda} p_{\lambda}^r = \bar{p}^r$$

On the other hand, we know that $\sum_{\lambda} a_{\lambda}^2 = 1$, as $1 = \int \phi^* \phi \cdot dT$ by the normalisation condition.

$$\begin{aligned} &= \sum_{\lambda} \sum_{\mu} \int a_{\lambda}^* \psi_{\lambda}^* a_{\mu} \psi_{\mu} \cdot dT \\ &= \sum_{\lambda} \sum_{\mu} a_{\lambda}^* a_{\mu} \delta_{\lambda\mu} \\ &= \sum_{\lambda} |a_{\lambda}|^2 \quad (i) \end{aligned}$$

Furthermore, $\bar{p}^r = \int \phi^* P^r \phi \cdot dT$, ~~and~~ $P^r \psi_{\lambda} = P^{(r-1)}(P \psi_{\lambda}) = P^{r-1} p_{\lambda} \psi_{\lambda}$

$$\begin{aligned} &= p_{\lambda}^2 P^{r-2} (P \psi_{\lambda}) \\ &= p_{\lambda}^2 P^{r-2} \psi_{\lambda}, \text{ etc.} \\ &= p_{\lambda}^r \psi_{\lambda} \end{aligned}$$

Hence by expanding ϕ in $\int \phi^* P^r \phi \cdot dT$ in terms of the ψ_{λ} and then making use of the above

we have that $\bar{p}^r = \sum_{\lambda} |a_{\lambda}|^2 p_{\lambda}^r$. (ii). On combining (i) and (ii), we have the system of equations:

$$\sum_{\lambda} (|a_{\lambda}|^2 - w_{\lambda}) p_{\lambda}^r = 0 \quad r = 1, 2, 3, \dots$$

The p_{λ} here are ~~not~~ given real numbers, all different. If these equations are to be true, either all the differences $|a_{\lambda}|^2 - w_{\lambda}$ all vanish, or $\Delta(p_{\lambda}^r)$ vanishes. This determinant certainly does not in the general case, hence: $w_{\lambda} = |a_{\lambda}|^2$

Corollary:- $w_{\lambda} = \left| \int \psi_{\lambda}^* \phi \cdot dT \right|^2$

This corollary follows by writing $\phi = \sum_i a_i \psi_i$, which defines the a_i (called coefficients of the development of ϕ in the system of ψ 's. We then have

$$\int \psi_\lambda^* \phi \cdot dT = \sum_i \int \psi_\lambda^* a_i \psi_i \cdot dT = \sum_i a_i \delta_{i\lambda} = a_\lambda$$

We are now in a position to calculate the relative frequency with which a given eigenvalue of any observable will be measured when a great number of measurements are made, provided we know the state ϕ of the system. For this purpose we either expand ϕ in terms of the eigenstates of the observable, pick out the coefficient belonging to the eigenvalue in question, and square its absolute value, or we perform the integration in the corollary (which gives the customary formula for the development coefficients of a function ϕ in an orthonormal system ψ_λ).

- 3) The probability (relative frequency) of finding the system, which is in the state $\phi(q_{1,2,\dots,s})$ at the point $(\beta_{1,2,\dots,s})$ is given by $\phi^*(\beta_{1,2,\dots,s})\phi(\beta_{1,2,\dots,s})$.

Proof: We have shown that the eigenstates belonging to the co-ordinate operator x have the form $\delta(x-\beta_\lambda)$, where the eigenvalues β_λ have a continuous distribution. By a generalisation of the argument leading to this we find that the eigenfunctions belonging to the operator $(q_{1,2,\dots,s})$ corresponding to the observable "simultaneous occupation of the points $q_{1,2,\dots,s}$ (s being the number of degrees of freedom), are simply products of δ -functions, viz.

$$\delta(q_1-\beta_{1\lambda}) \cdot \delta(q_2-\beta_{2\lambda}) \cdot \dots \cdot \delta(q_s-\beta_{s\lambda})$$

The eigenvalues $\beta_{i\lambda}$ form again a continuous distribution. Referring now to one particular λ , so that the index λ can be omitted, meaning by β_1 some fixed value of q_1 , and by β_2 some fixed value of q_2 , etc, the probability we want is (from (2) and its corollary),

$$\begin{aligned} & \left| \int \phi(q_{1,2,\dots,s}) \delta(q_1-\beta_1) \delta(q_2-\beta_2) \cdot \dots \cdot \delta(q_s-\beta_s) dq_1 dq_2 \dots dq_s \right|^2 \\ &= \phi^*(\beta_{1,2,\dots,s}) \phi(\beta_{1,2,\dots,s}), \text{ which proves the theorem.} \end{aligned}$$

In particular, if ϕ is the state of a single particle and therefore only a function of x,y,z , $\phi^*\phi$ is the probability that the particle be found at the point x,y,z of ordinary space. Hence it is permissible in this simple case to think of the square of the state function as the distribution of mass, or if the particle be an electron, the distribution of charge in space. This simple interpretation is not however applicable, and is, in view of (III), a very special idea which does not exhaust the meaning of a ϕ -function.

- 4) Certainty of measuring the eigenvalue p_k exists if and only if the state function $\phi = e^{i\mu} \psi_k$, where μ is a real constant.

This is almost obvious, for if certainty is to exist, there can only be one a_λ different from zero, namely a_k , and this must have an absolute value unity. The

above is the only ϕ -function which satisfies this requirement. This fact throws an interesting light upon eigenfunctions. They characterise states which yield certainty with regard to the outcome of at least one kind of measurement, that of the observable of which the state is an eigenstate. In the preceding discussion the latter have been denoted by ψ , while any state in general was denoted by ϕ . This usage of letters will be continued wherever feasible, but it should be remarked that there is no real distinction between states and eigenstates. Presumably every ϕ is an eigenstate of some operator, although it may be difficult (or perhaps impossible) to find it, or the operator may not be of physical interest. When we say that a state is an eigenstate of some operator, we mean that we know, and are interested in, this operator.

Suppose that a physical system is in an eigenstate belonging to one observable, say the energy; will it then be true, also, that some other observable, like the position of the system, will yield one definite value without spread on measurement? This is indeed possible, but only under a rather special condition.

- 5) If P and Q are permutable operators, the eigenstates belonging to P and Q are simultaneous eigenstates; that is, if the state of the system is such that a value p_i will be measured with certainty, then one value q_j will also be measured with certainty.

Proof:- Let the eigenstates of P be denoted by $\psi_{P\lambda}$, those of Q by $\psi_{Q\mu}$. We then know that the following equations are satisfied by hypothesis.

$$(a) P\psi_{P\lambda} = p_\lambda \psi_{P\lambda} \quad (b) Q\psi_{Q\mu} = q_\mu \psi_{Q\mu} \quad (c) PQ = QP$$

The state function is one of the $\psi_{P\lambda}$, say ψ_{Pi} . Then, because of (c) and (a),

$$QP\psi_{Pi} = PQ\psi_{Pi} = p_i Q\psi_{Pi}$$

If we write the last two members of this equation in the form

$$P(Q\psi_{Pi}) = p_i(Q\psi_{Pi})$$

it simply tells us that $Q\psi_{Pi}$ is an eigenfunction of the operator P, namely that belonging to the i 'th eigenstate. In other words, $Q\psi_{Pi} = \text{constant} \cdot \psi_{Pi}$. Comparing this with (b), we are forced to identify ψ_{Pi} with one of the $\psi_{Q\mu}$'s, say ψ_{Qj} , and the constant with q_j . Hence $\psi_{Pi} = \psi_{Qj}$. If we now apply the mean value relation to find the observed value of the q observable,

$$\bar{q} = \int \psi_{Pi}^* Q\psi_{Pi} \cdot dT = \int \psi_{Qj}^* Q\psi_{Qj} \cdot dT = q_j$$

i.e., the value q_j is measured with certainty. The opposite case, in which P and Q are not permutable, is also of interest. The above proof cannot then be conducted, and we infer that uncertainty of measurement for one observable will in general imply uncertainty of measurement for the other. In fact, the present formalism permits an answer to the more general question: for any given state ϕ , how is the precision in the measurement of one observable related to the precision in the measurements of the other? By precision we here mean a measure of the smallness of the range over which the various measurements are scattered. We shall here discuss this question as it regards two canonically conjugate observables, i.e. those whose corresponding operators are canonically conjugate.

For our present purpose we shall take the statistical dispersion of the measurements as a measure of the reciprocal precision, i.e. the uncertainty, so that the uncertainty in p is

$$(\Delta p)^2 = (p - \bar{p})^2$$

6). If P and Q are canonically conjugate operators, then $\overline{(\Delta p)^2} \cdot \overline{(\Delta q)^2} \geq \frac{h^2}{16\pi^2}$

Proof:- We shall need to assume that P and Q satisfy the relations

$$\begin{aligned} \int \phi^* P \phi . dT &= \int (P^* \phi^*) \phi . dT \\ \int \phi^* Q \phi . dT &= \int (Q^* \phi^*) \phi . dT \quad \text{if } \phi \text{ satisfies (A) and (B).} \end{aligned}$$

We first proceed to establish the inequality

$$\int u^* u . dT \cdot \int v^* v . dT \geq \frac{1}{4} \left[\int (u^* v + v^* u) . dT \right]^2$$

where u and v are any integrable functions of those co-ordinates on which the state-function for the system depends. This is done as follows: let λ be a real variable not depending on the co-ordinates. Then

$\int (\lambda u + v)(\lambda u^* + v^*) . dT$ is always positive or zero, the latter only for the trivial case where u is directly proportional to v , which we are here excluding.

Hence $\lambda^2 \int u^* u . dT + \lambda \int (u^* v + v^* u) . dT + \int v^* v . dT > 0$ for every real λ

Consider the l.h.s. as a quadratic form in λ . We know that it can have no real roots. But the roots of $a\lambda^2 + b\lambda + c$ are real unless $4ac > b^2$, whence we arrive immediately at the required inequality.

Now suppose that the system is in the state ϕ . If we then define

$$u = (P - \bar{p})\phi, \quad v = i(Q - \bar{q})\phi,$$

the inequality reads:

$$\begin{aligned} \int (P - \bar{p})^* \phi^* (P - \bar{p}) \phi . dT \cdot \int (Q - \bar{q})^* \phi^* (Q - \bar{q}) \phi . dT \\ \geq \frac{1}{4} \left[i \int (P - \bar{p})^* \phi^* (Q - \bar{q}) \phi . dT - i \int (Q - \bar{q})^* \phi^* (P - \bar{p}) \phi . dT \right]^2 \end{aligned}$$

Using now the first two relations, and cancelling terms on the right,

$$\int \phi^* (P - \bar{p})^2 \phi . dT \cdot \int \phi^* (Q - \bar{q})^2 \phi . dT \geq - \frac{1}{4} \left[\int \phi^* (PQ - QP) \phi . dT \right]^2$$

Now $\int \phi^* (P - \bar{p})^2 \phi . dT$ is equivalent to $\sum_{\lambda} a_{\lambda}^* a_{\lambda} (p_{\lambda} - \bar{p})^2$ (Th.1), and as Th.2 tells us that the $a_{\lambda}^* a_{\lambda}$ are probabilities, it is precisely what we mean by the uncertainty in p .

We have therefore arrived at the statement:

$$\overline{(\Delta p)^2} \cdot \overline{(\Delta q)^2} \geq -\frac{1}{4} \left[\int \psi^* (PQ - QP) \psi \cdot d\tau \right]^2$$

This is a theorem of great importance, more general indeed than the one we set out to derive. It contains (5) as a special case, for if P and Q are permutable, the r.h.s. vanishes, and it is possible for the two uncertainties to be zero.

If P and Q are operators belonging to canonically conjugate observables, like position and momentum, we have

$$PQ\psi - QP\psi = \frac{\hbar}{2\pi i} \cdot \frac{\partial(q\psi)}{\partial q} - q \cdot \frac{\hbar}{2\pi i} \cdot \frac{\partial\psi}{\partial q} = \frac{\hbar}{2\pi i} \cdot \psi$$

The relation derived above then becomes

$$\overline{(\Delta p)^2} \cdot \overline{(\Delta q)^2} \geq \frac{\hbar^2}{16\pi^2}, \text{ as was required.}$$

Hermitean Operators

An operator P is Hermitean if $\int u^*(Pv) \cdot d\tau = \int (P^*u^*)v \cdot d\tau$. A real operator is thus Hermitean if its position inside the integral is immaterial. The importance of such operators lies in their having real eigenvalues. To prove this, let u be one of the eigenfunctions of P, so that $u = \psi_k$, $P\psi_k = p_k \psi_k$, $P^*\psi_k^* = p_k^* \psi_k^*$.

We then have, expanding v in terms of the ψ 's (u and v satisfying the same boundary conditions by hypothesis),

$$\begin{aligned} \int u^*(Pv) \cdot d\tau &= \int \psi_k^* P \sum_{\lambda} a_{\lambda} \psi_{\lambda} \cdot d\tau = a_{k^*} p_k \\ \int (P^*u^*)v \cdot d\tau &= \int (P^*\psi_k^*) \sum_{\lambda} a_{\lambda} \psi_{\lambda} \cdot d\tau = a_{k^*} p_k^* \end{aligned}$$

If the two results are to be equal, i.e. if P be Hermitean, then the eigenvalues p_{λ} must all be real, for k may be any of the λ 's. Since eigenvalues must represent the results of measurements, which must be real, we are safe in physics if we admit only Hermitean operators.

The position and momentum operators introduced already are simply demonstrable to be Hermitean. Regarding the energy operator, we have, for a system with f degrees of freedom that

$$H = \sum_{\lambda=1}^f H_{\lambda} + V(q_{1,2}, \dots, f)$$

$$\text{where } H_{\lambda} = -\frac{\hbar^2}{2\pi m_{\lambda}} \cdot \frac{\partial^2}{\partial q_{\lambda}^2}$$

$$\text{Then } \int u^* H v \cdot d\tau = \sum_{\lambda} \int u^* H_{\lambda} v \cdot d\tau + \int u^* V v \cdot d\tau$$

Writing each term of the sum explicitly, $\int u^* H_{\lambda} v \cdot d\tau = -\frac{\hbar^2}{8\pi^2 m_{\lambda}} \int \dots \int u^* \frac{\partial^2 v}{\partial q_{\lambda}^2} dq_1 dq_2 \dots dq_f$

On performing a partial integration over q_λ this becomes:

$$-\frac{\hbar^2}{8\pi^2 m_\lambda} \int \dots \int \left[u^* \cdot \frac{\partial v}{\partial q_\lambda} \right]_{-\infty}^{+\infty} dq_1 dq_2 \dots dq_{\lambda-1} dq_{\lambda+1} \dots dq_f + \frac{\hbar^2}{8\pi^2 m_\lambda} \int \dots \int \frac{\partial u^*}{\partial q_\lambda} \cdot \frac{\partial v}{\partial q_\lambda} \cdot dT$$

The first integral vanishes. If we perform another similar partial integration with respect to q_λ in the second, the remaining term takes the form:

$$-\frac{\hbar^2}{8\pi^2 m_\lambda} \int \dots \int \frac{\partial^2 u^*}{\partial q_\lambda^2} \cdot v \cdot dq_1 dq_2 \dots dq_f$$

This is $(H_\lambda^* u^*) v \cdot dT$, and so we have demonstrated the Hermitean character of each term in the sum. As V is real and contains no differentiations, it is certainly Hermitean.

$$\text{Therefore } \int U^* H v \cdot dT = \int (H^* u^*) v \cdot dT = \int (H u^*) v \cdot dT$$

The class of operators under discussion has another interesting property. The eigenfunctions belonging to different eigenvalues of a Hermitean operator are orthogonal. Let p_1 and p_k be two different eigenvalues of the operator P in question. Then:

$$P \psi_1 = p_1 \psi_1$$

$$\text{Hence } \int \psi_k^* P \psi_1 \cdot dT = p_1 \int \psi_k^* \psi_1 \cdot dT.$$

Also, $P \psi_k = p_k \psi_k$ and $P^* \psi_k = p_k \psi_k^*$, since p_k is real, hence:

$$\int \psi_1 P^* \psi_k^* \cdot dT = p_k \int \psi_k^* \psi_1 \cdot dT$$

But P being Hermitean, $\int \psi_1 P^* \psi_k^* \cdot dT = \int \psi_k^* P \psi_1 \cdot dT = p_1 \int \psi_k^* \psi_1 \cdot dT$, so that:

$$p_1 \int \psi_k^* \psi_1 \cdot dT = p_k \int \psi_k^* \psi_1 \cdot dT$$

Since by hypothesis, $p_1 \neq p_k$, it is evident that $\int \psi_k^* \psi_1 \cdot dT = 0$

The proof breaks down if two eigenvalues are equal, i.e. in the case of degeneracy. That is not, however, a serious matter, for it is then possible to take linear combinations of the non-orthogonal functions, thereby producing orthogonal ones. These new orthogonal functions are eigenfunctions in as proper a sense as the old, for they satisfy the operator equation with the eigenvalue to which the old ones referred, the equation being linear. The orthogonalisation can be carried out conveniently by Schmidt's orthogonalisation process. The results of this process are not unique, and there are in general many different orthogonal linear combinations, all of which can be regarded as proper eigenfunctions belonging to a degenerate eigenvalue of P .

Special Forms of Schrödinger's Equation.

Since the energy of a physical system is a matter of great concern in many problems, we shall do well to consider in some detail the various common types of differential equation which the energy operator generates, all of which are of the standard form

$$H \psi_\lambda = E_\lambda \psi_\lambda$$

a). The simplest example is that of a freely-moving mass point. Let its constant potential energy in the region in which it moves be V . The point has three degrees of freedom, hence the Schrödinger equation becomes:

$$-\frac{\hbar^2}{8\pi^2 M} \nabla^2 \psi + V\psi = E\psi, \text{ omitting the subscripts on } \psi \text{ and } E.$$

Writing $k^2 = (E-V) \cdot \frac{8\pi^2 M}{\hbar^2}$, we see that k corresponds to $\frac{2\pi}{h}$ times the classical momentum

of the particle. We then have $\nabla^2 \psi + k^2 \psi = 0$, which can be solved at once by writing ψ as a product $X(x) \cdot Y(y) \cdot Z(z)$. Substituting for ψ and dividing through by it, we get:

$$\frac{X''}{X} + \frac{Y''}{Y} + \frac{Z''}{Z} + k^2 = 0$$

Since, except for the constant, each of the terms in this equation depends upon one independent variable which does not occur in the others, each term must be separately constant. Hence, denoting these constants by

$$-k_1^2, -k_2^2, -k_3^2 \text{ respectively, we obtain the solutions:}$$

$$X = A_1 \exp(ik_1 x) + B_1 \exp(-ik_1 x), Y = A_2 \exp(ik_2 y) + B_2 \exp(-ik_2 y), \text{ etc.}$$

The A's and B's are any constants compatible with the boundary conditions. The next question is: what are the boundary conditions? Let us first assume that the fundamental range of (x, y, z) covers all space. It is then necessary to satisfy condition (A). Condition (B) is automatically satisfied.

It will however at once be seen that it is impossible to satisfy (A) without making all A's and B's zero. This corresponds to an eigenfunction which vanishes everywhere. According to Th. 3, the particle would have no probability of being detected anywhere. We shall return to this situation later.

Another possibility is that the particle is, classically speaking, somewhere in the range $(0, l_1), (0, l_2), (0, l_3)$, and that it has a zero probability of being found at the boundaries of this parallelepiped. In view of Th. 3, $\psi^* \psi$ must then vanish at the boundaries, which is only possible if ψ vanishes there. If we impose this condition we find that

$k_1 = n\pi/l_1$, where n is an integer which will not be the same for all three k_i . This fixes the energies of the problem, and they are:-

$$E_{(n_1, n_2, n_3)} - V = \frac{\hbar^2}{8\pi^2 M} (k_1^2 + k_2^2 + k_3^2) = \frac{\hbar^2}{8M} \left[\left(\frac{n_1}{l_1}\right)^2 + \left(\frac{n_2}{l_2}\right)^2 + \left(\frac{n_3}{l_3}\right)^2 \right]$$

To every choice of the three integers (n_1, n_2, n_3) there corresponds an eigenfunction

$$\psi_{n_1, 2, 3} = c \cdot \sin \frac{n_1 \pi}{l_1} x \cdot \sin \frac{n_2 \pi}{l_2} y \cdot \sin \frac{n_3 \pi}{l_3} z$$

The constant c is determined by integrating $\psi^* \psi$ over all space and setting the result equal to unity. The result is

$$c = (8/l_1 l_2 l_3)^{\frac{1}{2}} = (8/v)^{\frac{1}{2}}, \text{ where } v \text{ is the volume of the}$$

parallelepiped. We see that the energy values form a denumerable set; they are discrete. To every triple set of numbers there corresponds an "energy level" of the system. To some sets of numbers there corresponds the same energy value. The states thus described are "degenerate".

If we increase l_1, l_2, l_3 , that is, the volume of the space in which the particle is confined, two things happen: the separation between the possible energies decreases, and the constant c becomes smaller. In the limit of an infinite volume, we achieve a continuous spectrum as nearly as we please, and c approaches zero as nearly as we please. This case is evidently that discussed before, but now we see the true nature of the problem. The energy spectrum becomes continuous, but the state-function becomes incapable of normalisation. This problem is a very general one in quantum mechanics.

How, then are we to deal with continuous ranges of eigenfunctions? The first possibility is to make (A) less stringent or replace it by a different condition in the case of continuous eigenvalues. We can, for instance, normalise the eigenfunction belonging to an infinitesimal range of eigenvalues instead of the one belonging to a single eigenvalue. This is the usual procedure and is the only one applicable in the numerical solution of physical problems. The other, logically perhaps more satisfactory, is to treat the problem as though the system were confined. The energy eigenvalues will then be discrete. The continuous spectrum will appear as the result of a limiting process in which the range of the fundamental variables approaches infinity. This does no violence to the physical situation since we can approximate to the actual condition as closely as desired, and, furthermore, no physical system moves, strictly speaking, in an infinite space.

As a final observation on the problem of the free mass point we add the following remarks: the Ψ -function is a triply sinusoidal relation. If we wish to complete the time picture of a wave we are at liberty to multiply by $\exp(i\omega t)$ for as we have seen, multiplication of the state-function by a constant (in space) of absolute value unity does not alter the state. If the wave is supposed to move entirely in the x direction $k_1 = k = 2\pi p/h$, which gives de Broglie's relation as $k\lambda = 2\pi$.

b) As a second example we consider the problem of finding the allowed energies of a system classically described as a rotator. The Hamiltonian of such a system, e.g. a dumb-bell having a moment of inertia I and rotating with angular momentum m_z about the z -axis, has the form

$$H = \frac{m_z^2}{2I}, \text{ since there is no potential energy. If in this}$$

expression we replace m_z by its operator $\frac{h}{2\pi i} \left[x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right] = \frac{h}{2\pi i} \frac{\partial}{\partial \theta}$, where θ is the longitudinal angle about the z -axis, the energy operator is seen to be $H = \frac{h^2}{8\pi^2 I} \frac{\partial^2}{\partial \theta^2}$

We are ascribing to the system only one degree of freedom, as is seen from the fact that the energy operator affects only one variable, θ . Schrödinger's equation is now

$$\frac{d^2 \Psi}{d\theta^2} + k^2 \Psi = 0, \text{ with } k^2 = \frac{8\pi^2 I}{h^2} \cdot E$$

It has the solution $\Psi = A e^{ik\theta} + B e^{-ik\theta}$. Condition (B) demands that $k = n$, an integer, so that the possible energies are

$$E = \frac{n^2 h^2}{8\pi^2 I}$$

This prediction is not verified. The system which should exhibit these energy levels, the diatomic molecule, does not do so. Instead it has energies arranged according to the law obtained by replacing n^2 with $n(n+1)$ in the above. The reason from which this stems is that we have not assigned to the system a sufficient number of degrees of freedom. We shall therefore now work out the problem by ascribing it two instead of one.

We shall now first write the complete Hamiltonian in Cartesian form, so that Schrödinger's equation becomes:

$$\nabla^2 \psi + \frac{8\pi^2 M}{h^2} \cdot E \psi = 0$$

On transforming to spherical co-ordinates and then putting $r = a = \text{constant}$, there results:

$$\frac{1}{a^2 \sin^2 \theta} \frac{\partial^2 \psi}{\partial \phi^2} + \frac{1}{a^2} \frac{\partial^2 \psi}{\partial \theta^2} + \cot \theta \cdot \frac{1}{a^2} \frac{\partial \psi}{\partial \theta} + \frac{8\pi^2 M}{h^2} \cdot E \psi = 0$$

We now repeat essentially the procedure used before. Assuming ψ to be a product of two functions $\bar{\theta}(\theta)$ and $\bar{\phi}(\phi)$, we substitute in the above and divide by ψ . Then, on multiplying by $a^2 \sin^2 \theta$, the equation assumes the form:

$$\frac{\bar{\phi}''}{\bar{\phi}} + \sin^2 \theta \cdot \frac{\bar{\theta}''}{\bar{\theta}} + \sin(\theta) \cos(\theta) \cdot \frac{\bar{\theta}'}{\bar{\theta}} + \bar{E} \cdot \sin^2 \theta = 0$$

In this equation \bar{E} is defined as $\frac{8\pi^2 M a^2}{h^2} \cdot E$

For an acceptable function $\bar{\phi}$, the first term must be representable as $-m^2$, where m is real and constant. Single-valuedness requires that m also be an integer. If this is then substituted into the equation for $\bar{\theta}$, we have the equation for the "associated Legendre function". The eigenvalues of this give $\bar{E} = n(n+1)$, and these eigenvalues turn out to be independent of the value of m . Because of the definition of \bar{E} , this means that:

$$E_n = \frac{n(n+1)h^2}{8\pi^2 I}, \text{ since } I = Ma^2.$$

The complete eigenfunctions are: $\psi = \text{const} \cdot P_n^m(\cos \theta) (A e^{im\phi} + B e^{-im\phi})$

The definition of $P_n^m(\cos \theta)$ shows that the value of m cannot be greater than n , nor can it be negative.

These examples illustrate the detailed manner of solving Schrödinger's Equation. It should be remarked that in many cases of interest exact solutions are not available and recourse must be made to approximate methods of solution. These approximate methods, usually comprised under the general heading of "perturbation theory" justify a closer study because of the insight which they afford into the workings of physical disturbances in an eigenstate. In a discussion of this matter a deeper knowledge of the basic properties ~~and~~ of solutions to and eigenvalues of Schrödinger's equation will be needed, and these are studied next. In particular, we justify the assumption of "completeness" of the eigenfunctions.

General Properties of Solutions to Schrödinger's Equation

Schrödinger's equation belongs to the class of Sturm-Liouville equations, usually written in the form: $L(u) + \lambda\beta u = 0$, where L is a differential operator defined by: $L(u) = (pu')' - qu$, primes denoting differentiation in x where $u = u(x)$. The functions p , q , and β are also functions of x . It will be supposed that $\beta(x)$, which has the nature of a weighting function, satisfies the condition $\beta \geq 0$ in the entire fundamental domain of x , which domain will be taken to be finite. λ is a constant, the eigenvalue of the Sturm-Liouville equation. We shall first demonstrate that the common forms of Schrödinger's equation, or the component equations into which it can be separated, are of this form.

Using the abbreviation $k^2 = 8\pi^2 M/h^2$, we can write every one-dimensional Schrödinger equation in the form $\psi'' - k^2 V(x)\psi + k^2 E\psi = 0$.

Comparison with the Sturm-Liouville form gives:

$$p = 1, \quad q = k^2 V(x), \quad \beta = 1, \quad \text{and} \quad \lambda = k^2 E$$

As another example, the radial equation for the H atom is :

$$(r^2 R')' - (\underline{1}(\underline{1}+1) + k^2 r^2 V(r))R + k^2 r^2 ER = 0$$

Here we have $p = r^2$, $q = \underline{1}(\underline{1}+1) + k^2 r^2 V(r)$, $\beta = r^2$, $\lambda = k^2 E$.

These two equations illustrate that from one Schrödinger equation to another the fundamental ranges of the independent variables may and in general will be different. Nor will the boundary conditions in general agree. In any one case, however, we may express the relevant boundary conditions in the following form : if u and v are any two admissible ψ functions, then $vpu' \Big|_b = vpu' \Big|_a$ if a and b are the end-points of the range in any one problem. The theory can be developed on the basis of this property. For the sake of simplicity, all functions u and v will be assumed real.

The operator L in the Sturm-Liouville equation is Hermitian with respect to functions satisfying the boundary conditions given above. If we work out the integral $\int vL(u)dx$ over the range from a to b we obtain

$$\int v(pu')' dx - \int vqu dx = vpu' \Big|_b^a - \int v'pu' dx - \int vqu dx$$

The integrated part vanishes because of the boundary conditions. Performing another partial integration, we find that $\int vL(u)dx = -v'pu \Big|_a^b + \int u(pv')' dx - \int vqu dx$

$$= \int uL(v)dx$$

Since L is a real operator, $L = L^*$ and the above result proves its Hermitian nature. The above is a special case of a theorem known as Green's theorem.

The Stationary Character of the Eigenvalues

As will now be shown, the Sturm-Liouville equation itself represents the condition that the integral $\Delta(u) = \int (pu'^2 + qu^2) dx$ shall be a maximum or minimum, provided that $\int \beta u^2 dx = 1$. This last is the normalising condition usually applied, as will be apparent from consideration of the functions β given on the previous page. To demonstrate this, the following result from the calculus of variations is required: if u is an unknown function of x , the form of which is to be determined, and $I(u, u', x)$ is a given function, then the u -functions which make $\int I dx$ stationary, i.e. $\delta \int I dx$ zero, must satisfy Euler's equation, viz.:

$$\frac{\partial I}{\partial u} - \frac{d}{dx} \frac{\partial I}{\partial u'} = 0$$

But if there is an accessory condition $\int G(u) dx = \text{constant}$ to be obeyed, the I in the above has to be replaced by $I - \lambda G$, where λ is an undetermined constant multiplier. The condition that $\delta \Delta(u)$ be zero subject to $\int \beta u^2 dx = 0$ is given by putting $I = pu'^2 + qu^2 - \lambda \beta u^2$, and substitution of this in the Euler equation gives the Sturm-Liouville equation. The process of solving the Sturm-Liouville equation is thus tantamount to finding the stationary values of $\Delta(u)$. (This is unrigorous, since no proof of the sufficiency of the Euler equation as a condition for stationarity can be given, but it holds for all physical problems). $\Delta(u)$ has in general many stationary values.

The eigenfunctions of the Sturm-Liouville equation will now be ordered in the following way: suppose that $u = u_1(x)$ satisfies the equation and the normalising condition, and that it makes $\Delta(u)$ a minimum (as it usually will for Schrödinger's equation). We shall denote the corresponding eigenvalue by λ_1 . We shall now seek a minimum of $\Delta(u)$ subject not only to $\int \beta u^2 dx = 1$, but also to $\int \beta u u_1 dx = 0$. Let $u_2(x)$ be the function which produces this minimum. It is at once clear that this minimum cannot lie lower than the first, because the requirement on the admissible u -functions has been made more stringent. u_2 must satisfy two accessory conditions, u_1 only one.

It is found that u_2 is also subject to the Sturm-Liouville equation. We now have I as

$$I = pu_2'^2 + qu_2^2 - \lambda_2 \beta u_2^2 - \mu \beta u_1 u_2, \lambda_2 \text{ and } \mu \text{ being two undetermined constants introduced in virtue of the two accessory conditions. Putting this in the Euler equation, we now obtain: } 2qu_2 - 2\lambda_2 \beta u_2 - \mu \beta u_1 - 2(pu_2')' = 0.$$

This is equivalent to: $L(u_2) + \lambda_2 \beta u_2 + \frac{1}{2} \mu \beta u_1 = 0$.

To determine μ we multiply by u_1 and integrate over x . In doing this we observe that

$$\int u_1 L(u_2) dx = \int u_2 L(u_1) dx = -\lambda_1 \int \beta u_1 u_2 dx = 0.$$

The first step is justified by L being Hermitian, the second because u_1 satisfies the Sturm-Liouville equation, the third because of the normalisation condition.

The term in L thus vanishes and we are left with $\lambda_2 \int \beta u_1 u_2 dx = -\frac{1}{2} \mu \int \beta u_1^2 dx$. Here the left-hand side vanishes on account of the second normalisation condition, and, as we have $\int \beta u_1^2 dx = 1$, we must conclude that $\mu = 0$. The equation for u_2 thus reduces to the Sturm-Liouville equation. In other words, the function u_2 , which produces another minimum in $\Delta(u)$, is simply another solution of this equation. λ_2 , the multiplier appearing in the minimising problem, is the second eigenvalue of this equation.

This process can be continued. We next seek a function u_3 which will minimise $\Delta(u)$, but subject to the three conditions

$$\int \beta u_3^2 dx = 1, \quad \int \beta u_1 u_3 = 0, \quad \int \beta u_2 u_3 = 0.$$

The minimum thus obtained will lie at least as high as that due to u_2 , for the choice of admissible functions has been further restricted. If we write down Euler's equation, there will now appear three undetermined constants, λ_3 and two others which may be shown to be zero by multiplication by u_1 and integrating, and multiplication by u_2 and integrating. The resulting equation is again a Sturm-Liouville equation, with λ_3 as its eigenvalue. In this way we obtain an ordered sequence of eigenfunctions $u_{1,2,3,\text{etc.}}$ and corresponding eigenvalues $\lambda_{1,2,3,\text{etc.}}$ the arrangement being such that the minimum of $\Delta(u)$ corresponding to u_1 is the lowest, that due to u_2 the second lowest, and so on, except for the possibility of coincidence of these minima.

The last and most important step is to show that the minima of $\Delta(u)$ are identical with the successive λ 's. This is done by substitution:

$$\begin{aligned} \Delta(u_n) &= \int (pu_n'^2 + qu_n^2) dx = u_n \int_a^b pu_n' dx - \int (u_n (pu_n')' - u_n qu_n) dx \\ &= - \int L(u_n) dx = \lambda_n \int \beta u_n^2 dx \\ &= \lambda_n \end{aligned}$$

We have thus proved that the eigenvalues of the Sturm-Liouville equation are the successive minima of $\Delta(u)$ as u is subjected to increasingly restrictive conditions. These results have an immediate physical consequence. They permit the determination, in some cases, of an approximate value for the lowest energy state without solving the Schrödinger equation at all, and, to a poorer approximation, the determination of the eigenfunction corresponding to this state. We know that the lowest eigenvalue is the lowest possible value of Δ , with u subject only to the condition of normalisation. In some cases a good guess can be made at the general form of the eigenfunction u_1 , which is left undetermined by using several variable parameters. One can then easily calculate $\lambda_1^i = - \overline{u_1} L(\overline{u_1}) dx$, where $\overline{u_1}$ is the trial function which depends on the parameters. λ_1^i will then also depend on the parameters. If now we minimise λ_1^i with respect to these parameters, its lowest possible value will be λ_1 ,

and \bar{u}_1 corresponding to these parameters will be u_1 . The accuracy of this method depends, of course, on the adequacy of the initial selection of \bar{u}_1 . In general, the λ_1 determined in this way will be too high.

On comparing with the one-dimensional form of Schrödinger's equation, it is seen that, apart from multiplicative constants, the operator L is the same as -H. Hence,

$$\lambda_1 = \text{constant} \cdot \int u_1 H(u_1) dx$$

and the method just outlined amounts to minimisation of the energy.

Further, if p and q are both greater than zero, $\Delta(u_n) = \int (pu_n'^2 + qu_n''^2) dx > 0$, and therefore every λ_n is greater than zero. There can therefore be no energy states below the minimum of potential energy (this is of course trivial from a classical point of view). That this is so may be illustrated by the one-dimensional case, for which the Schrödinger equation is $\psi'' - k^2(V(x) - V_0)\psi + k^2(E - V_0)\psi = 0$, where V_0 is the minimum of potential energy. Here p and q are greater than zero, and so $\lambda_n = k^2(E_n - V_0)$ is greater than zero. Thus E_n is greater than V_0 . **

Distribution of High Energies

Although the sequence of eigenvalues of the Sturm-Liouville equation follows no uniform law, depending as it does on the form of p, q, and β , it is nevertheless true that the high eigenvalues of all Sturm-Liouville equations are distributed according to the same law, viz:

$$\lim_{n \rightarrow \infty} \lambda_n = \text{constant} \cdot n^2$$

To prove this, we transform the equation by the substitutions

$$u = p^n \beta^k z, \quad t = \int_a^x p^l \beta^m dx$$

and use t as independent variable. If we find the values of l, m, k, n which cause (1) the coefficient of $\underline{T}z$ to vanish, and (2) the coefficient of $\underline{T}^2 z$ to be the same as that of λz , we find that $k = n = -\frac{1}{4}$, $l = -\frac{1}{2}$, $m = \frac{1}{2}$. The Sturm-Liouville equation then reduces to $\underline{T}^2 z - r(t)z + \lambda z = 0$, where r(t) is a function of t which is in general bounded, but its maximum depends on p, q, and β . Let the upper bound of its absolute value be M, so that it is a good approximation to say that if λ is much greater than M, the eigenvalues are those of the equation $\underline{T}^2 z + \lambda z = 0$. This has the general solution $z = A \cos(\lambda^{\frac{1}{2}} t) + B \sin(\lambda^{\frac{1}{2}} t)$. Now, whatever the boundary condition on u required by the physical problem may be, z must vanish at the ends of the t range. This is easily verified by examining the boundary conditions and observing that $z = (p\beta)^{1/4} u$. The end-points of the t range are 0 and $t^0 = \int_a^b (\beta/p)^{\frac{1}{2}} dx$. This allows for z only the solution $B \sin(\lambda^{\frac{1}{2}} t)$, with λ given by $n^2 \pi^2 / t^{0^2}$. Therefore

$$\lim_{n \rightarrow \infty} \lambda_n = \frac{\pi^2 n^2}{\int_a^b (\beta/p)^{\frac{1}{2}} dx}$$

This result is immediately reminiscent of the free electron case, where the energy levels are distributed according to the law $\text{const.} \cdot n^2$. The spectrum is discrete if (a, b) is a finite range. However, if t^0 tends to infinity in the same manner as n it becomes continuous, for then $\lambda_{n+1} - \lambda_n$ is found to be zero. We encounter here all the characteristics of the free electron problem, and may therefore summarise by saying that the high energies of any physical system are distributed like those of a free electron.

The simple harmonic oscillator apparently violates this rule, the analytical reason for this behaviour being that $r(t)$ is not bounded. This, in turn, results from the form of V which goes to infinity like x^2 . Clearly such a case is over-idealised, and if V is given a finite upper limit our result holds for the oscillator as well.

In the next section the completeness of the solutions of Schrödinger's equation is investigated, and the above result is required in the form that λ_n goes to infinity with n .

Completeness of Solutions of Schrödinger's Equation

Completeness of a set of eigenfunctions u_i means the following: if $f(x)$ satisfies the same boundary conditions as the u_i , and we define

$$c_n = \int \beta f u_n dx \quad \text{and} \quad \delta_n = f - \sum_{i=1}^n c_i u_i$$

then completeness means that $\lim_{n \rightarrow \infty} \overline{\delta_n^2} = 0$, where we define $\overline{\delta_n^2} = \int \delta_n^2 \beta dx$

In words, this means that any function f subject to the same boundary conditions as the u_i can be "approximated in the mean" by a series in $c_i u_i$. This does not mean that

$$f = \sum c_i u_i$$

in the general case, because the left and right of this equation may not be equal at certain points although the completeness condition is satisfied. Nevertheless, the uniform convergence of $\sum c_i u_i$, and hence its equality with f , are always assumed in physical problems.

We now wish to establish this completeness property for the u_i of the Sturm-Liouville equation. We define

$$\left[\overline{\delta_n^2} \right]^{1/2} = a_n$$

The quantity δ_n/a_n is evidently normalised, in the sense that $\int (\delta_n/a_n)^2 \beta dx = 1$.

We also observe that δ_n/a_n is orthogonal to every u_i up to and including u_n , the order of the u_i being as described above. To show this we calculate:

$$\int (\delta_n/a_n) u_i \beta dx = \frac{1}{a_n} \cdot \left\{ \int f u_i \beta dx - \sum_{j=1}^n c_j \int u_j u_i \beta dx \right\}$$

$$= \frac{1}{a_n} \cdot (c_i - c_i) = 0 \quad \text{if } i \leq n$$

$$\text{and } = \frac{1}{a_n} \cdot (c_i - 0) = c_i/a_n \quad \text{if } i > n$$

These results follow from the fact that u_n is orthogonal to all preceding u_i , and from the definition of c_n .

We can say, therefore, that the function δ_n/a_n satisfies all the accessory conditions of u_{n+1} , except that of minimising Δ . Considering now the class of functions which satisfy these conditions, u_{n+1} produces the smallest $\Delta(u)$, namely λ_{n+1} . Therefore

$$\Delta\left[\frac{\delta_n}{a_n}\right] \geq \Delta(u_{n+1})$$

$$\text{or } \int \frac{1}{a_n^2} \left\{ p \cdot (\delta_n)'^2 + q(\delta_n)^2 \right\} dx \geq \lambda_{n+1}, \text{ putting in the form of } \Delta(u) \text{ and the value of } \Delta(u_{n+1})$$

Defining the abbreviations $g_i = pu_i'f' + qu_i f$

$h_{ij} = pu_i'u_j' + qu_i u_j$, and inserting the definition of δ_n ,

$$\frac{1}{a_n^2} \left\{ \int (pf'^2 + qf^2) dx - 2 \sum_{i=1}^n c_i \int g_i dx + \sum_{i=1}^n \sum_{j=1}^n \int h_{ij} dx \right\} \geq \lambda_{n+1} \quad (E)$$

But $\int g_i dx = pu_i'f' \Big|_a^b - \int [u_j (pu_i')' - u_j qu_i] dx = \int f \lambda_i u_j \beta u_i dx = c_i \lambda_i$, performing a partial integration, applying the boundary conditions and the S-L equation, and the definition of c_n in that order.

Similarly, $\int h_{ij} dx = u_j pu_i' \Big|_a^b - \int [u_j (pu_i')' - u_j qu_i] dx = \int u_j \lambda_i \beta u_i dx = \lambda_i \delta_{ij}$ (δ_{ij} here = the Kronecker δ)

Substituting these values in (E) above, we find immediately that

$$\frac{1}{a_n^2} \left\{ \int (pf'^2 + qf^2) dx - \sum_{i=1}^n c_i^2 \lambda_i \right\} \geq \lambda_{n+1}$$

Since f satisfies the same continuity and boundary conditions as the u_i , which are essentially the state functions for the physical problem, the integral in this equation exists. (If f violates this rule, an "approximation in the mean" by the u_i may be impossible). Let its value be A . Also, because the lowest eigenvalue is the minimum of Δ , the very existence of such an eigenvalue ensures that it will lie above some finite lower bound. If we reckon all λ 's from this lower bound, i.e. measure all energies taking the lowest state as the zero level, which is clearly permissible, they are all positive. Hence the summation in the above equation is certainly not negative and $A \geq \lambda_{n+1} \cdot a_n^2$. But we know that λ_{n+1} tends to infinity with $n+1$, and therefore, since a_n^2 was defined by

$$a_n^2 = \overline{\delta_n^2}$$

we have the result that Limit $\delta_n^2 = 0$ as n tends to infinity.

Matrix Mechanics

Let P be an operator generating a complete system of orthogonal functions by means of the equation

$$P\phi_i = p_i\phi_i$$

The number of different ϕ -functions will in general be infinite. If Q is some other operator, defined with respect to the same variables as P , we can form a doubly infinite array of numbers

$$Q_{ij} = \int \phi_i^* Q \phi_j \cdot dT$$

This fact in itself is trivial, but it is noteworthy that the rules for combining several operators on this definition are precisely those for combining matrices. In other words, if M and N are two operators, then

$$M_{ij} + N_{ij} = (M + N)_{ij}$$

$$\text{and } (MN)_{ij} = \sum_{\lambda} M_{i\lambda} N_{\lambda j}$$

This first is obvious, as $M_{ij} + N_{ij} = \int \phi_i^* M \phi_j \cdot dT + \int \phi_i^* N \phi_j \cdot dT = \int \phi_i^* (M + N) \phi_j \cdot dT = (M+N)_{ij}$

Further, $(MN)_{ij} = \int \phi_i^* (MN \phi_j) \cdot dT$. If we expand $N \phi_j$ as $\sum_{\lambda} a_{\lambda j} \phi_{\lambda}$, we must have

$$\int \phi_i^* N \phi_j \cdot dT = \sum_{\lambda} a_{\lambda j} \int \phi_i^* \phi_{\lambda} \cdot dT = a_{\lambda j} \delta_{i\lambda} = a_{ij}$$

By definition, $\int \phi_i^* N \phi_j \cdot dT = N_{ij}$, so that $N_{ij} = a_{ij}$

Thus $N \phi_j$ may be written $N \phi_j = \sum_{\lambda} N_{\lambda j} \phi_{\lambda}$. If this is substituted into $(MN)_{ij}$, we have:

$$(MN)_{ij} = \int \phi_i^* (M \sum_{\lambda} N_{\lambda j} \phi_{\lambda}) \cdot dT = \sum_{\lambda} N_{\lambda j} \int \phi_i^* M \phi_{\lambda} \cdot dT = \sum_{\lambda} M_{i\lambda} N_{\lambda j}$$

To form a matrix by this rule, two things are required; first, an operator such as Q ; second, a complete orthogonal system of functions ϕ . This system need not be the system of eigenfunctions of Q ; it may belong to some other operator P . If, however, the system of functions is that belonging to the operator itself, the matrix is diagonal. In forming the matrix elements of P we observe that:

$$P_{ij} = \int \phi_i^* P \phi_j \cdot dT = \int \phi_i^* p_j \phi_j \cdot dT = p_j \delta_{ij}$$

The diagonal elements of the matrix P_{ij} are simply the eigenvalues of the operator P . This theorem provides the link between Schrödinger's theory and the matrix theory of Heisenberg.

It has been observed that quantum mechanics deals exclusively with Hermitean operators. The properties of the matrix corresponding to a Hermitean operator are therefore of interest. If H is Hermitean, then

$$H_{ij} = \int (H^* \phi_i^*) \phi_j \cdot dT = \int \phi_j H^* \phi_i^* \cdot dT = H_{ji}^*$$

$$H_{ij} = \int \phi_i^* H \phi_j \cdot dT$$

A matrix whose elements satisfy this relation is also called Hermitean. A real Hermitean matrix is clearly symmetrical.

The Formal Structure of Matrix Mechanics

Atomic spectroscopy provides us with two sets of numbers: possible energies E_i and transition probabilities T_{ij} , these being the probabilities of transitions between the i 'th and j 'th energy states, measuring the intensity of the line of frequency $(E_j - E_i)/h$. The latter form a square array reminiscent of a matrix, the former a single sequence. If we feel tempted to represent the energies as a matrix also, we should plausibly look for a diagonal matrix with the energies strung along the principal diagonal. The problem of Heisenberg, Born, and Jordan was to devise a method of calculating these two matrices. The following scheme was developed with utter abandon of classical postulates and its success was striking.

For the present we shall merely state the directions for solving the problem without giving any reasons for them. Afterwards, we shall show that this heuristic method is related to, and indeed follows from, the operator theory given above.

Let it be desired to find the observable properties of a physical system which has a classical Hamiltonian

$$H(q_{1,2,\dots,k}, p_{1,2,\dots,k}), \text{ where } p\text{'s and } q\text{'s are as usual.}$$

We then seek a system of $2k$ matrices $Q_{1,2,\dots,k}, P_{1,2,\dots,k}$ which satisfy the following conditions

$$(1) \quad Q_m Q_n - Q_n Q_m = P_m P_n - P_n P_m = 0, \quad P_m Q_n - Q_n P_m = \frac{h}{2\pi i} \cdot \delta_{nm} \cdot I \quad (I = \text{unit matrix})$$

$$(2) \quad \text{and } H(Q_{1,2,\dots,k}, P_{1,2,\dots,k}) \text{ is diagonal}$$

If such a system can be found, the diagonal elements of H are the energies of the problem, and the squares of the elements of the Q matrices are the transition probabilities.

It should be noticed that the subscripts to the Q 's distinguish not elements but entire matrices whose elements would be written $(Q_\lambda)_{ij}$, etc. The first part of (1) states

simply that all the P and Q matrices commute among one another, while the second part indicates that a given co-ordinated matrix does not commute with the matrix assigned to its conjugate momentum. The above statement might well be expressed in terms of the Poisson bracket (PQ), and the formal similarity between the axioms of the present theory and the transformation theory of mechanics is in fact very thorough-going. H in (2) is a function of matrices and is therefore a matrix itself. It is constructed by expanding the classical H as a function of p 's and q 's and then replacing each of these by its associated matrix. We state without proof that conditions (1) and (2) define a set of matrices unique with respect to the diagonal elements of H and the squares of the elements of Q . (The elements of Q are in fact indeterminate to within a constant multiplier of absolute value one).

If we regard q_1 as x , q_2 as y , q_3 as z , then $|(Q_1)_{ij}|^2$ is the probability of a transition from the i 'th to the j 'th energy level resulting in the emission or absorption of light having its electric vector along x , etc., so that the theory is competent to inform us not only about the total probability of transition, which would clearly be

$$T_{ij} = (Q_1)_{ij}^2 + (Q_2)_{ij}^2 + (Q_3)_{ij}^2$$

but even about the state of polarisation of the light emitted or absorbed.

Before considering the details, we require two theorems on matrix transformations.

- 1). If H' is Hermitean, there exists a matrix S which makes $H = S^{-1}H'S$ diagonal. This will be stated without proof.
- 2). The above transformation applied to a matrix product has the same effect as its application to every matrix in the product individually, i.e., if A, C, B , are matrices,
- $$S^{-1}(ABC)S = S^{-1}AS.S^{-1}BS.S^{-1}CS. \quad \text{This is an identity.}$$

We now show how the directions for finding the matrices in question can be carried out. It is usually fairly simple to find a set will will satisfy conditions (1). These may be called Q'_i, P'_i . These will not however in general make $H(Q'_i, P'_i)$ diagonal. By means of the above transformation with some as yet undetermined S , however, $H(Q'_i, P'_i)$ can be made diagonal. Hence, if we write H' for $H(Q'_i, P'_i)$ and H for the diagonal matrix we are seeking, $S^{-1}H'S = H$. If it is known, then the energies are given by the diagonal elements of H . Not only that, we can then calculate likewise the elements of the correct Q and P matrices. Equation (2) above shows that

$$Q_i = S^{-1}Q'_i S$$

The solution of the problem thus reduces to a determination of the correct transformation matrix S .

To complete the discussion, we indicate briefly how this is done. The transformation, on multiplication on the left by S , takes the form:

$$H'S = SH$$

The elements of H' in this equation are known. In terms of them, the matrix equation above may be written as a sequence of ordinary equations:

$$\sum_{\lambda} H'_{i\lambda} S_{\lambda j} = \sum_{\lambda} S_{i\lambda} H_{\lambda j} = S_{ij} H_{jj} \quad \text{because } H \text{ is diagonal}$$

For any one value of j , we find therefore that

$$\sum_{\lambda} (H'_{i\lambda} - H\delta_{i\lambda}) S_{\lambda} = 0, \quad i = 1, 2, 3, \dots, \quad (\text{we have omitted here the suffix } j.)$$

Here H is simply a number. This is a set of linear equations to be solved for the unknown quantities S_{λ} . If it is to have solutions, the determinant of the coefficients of the S_{λ} must vanish. Let us suppose the maximum value of i to be n (which may be infinite), so that we have here n equations. The vanishing of the determinant is equivalent to an algebraic equation of the n th degree in H , and has therefore n solutions for H , of which some may coincide (degenerate case). By introducing any one of the values of H thus determined into the above, n values of the S_{λ} can be calculated. We thus obtain a total of n^2 S -values in all, and these form, as they should, the square array composing the transformation matrix.

Equivalence of Matrix and Wave Mechanics

Let us consider the matrices which are to satisfy (1) and (2) as being formed from the set of operators according to the rule

$$Q_{ij} = \int \phi_i^* Q \phi_j dT$$

The matrices Q' and P' will obey (1) if the operators from which they are formed obey these relations. Interpreting (1) as operator equations we see that they are satisfied by the operators:

$$Q'_m = q_m \quad P'_m = \frac{h}{2\pi i} \cdot \frac{\partial}{\partial q_m}$$

This is precisely the assignment which has already been made. The elements of the matrices Q'_m and P'_m are then formed by choosing any complete orthogonal system of functions, ϕ_i , and using the definition of Q_{ij} . These matrices will not make H diagonal. The condition under which

$$H(Q'_{1,2,\dots,k}, P'_{1,2,\dots,k}) = \phi_i^* H(q_{1,2,\dots,k}, \frac{h}{2\pi i} \cdot \frac{\partial}{\partial q_{1,2,\dots,k}}) \phi_j \cdot dT$$

is diagonal is simply that the functions ϕ satisfy the operator equation

$$H\psi_i = E_i \psi_i \quad (\phi = \psi_i)$$

This is none other than Schrödinger's. Thus the problem of making H a diagonal matrix is mathematically identical with the problem of solving Schrödinger's equation. The ϕ 's satisfying Schrödinger's equation ~~make the~~ furnish the diagonal elements

$$\int \psi_i^* H \psi_j \cdot dT = E_{ij}$$

The correct choice of matrices in Heisenberg's matrix mechanics is the exact equivalent of the proper selection of eigenfunctions in Schrödinger's theory. When, in the former theory, we determine the elements of S we are doing the same as when, in the operator theory, we pass from one system of state functions ϕ which are not solutions of Schrödinger's equation to the ψ 's which are. This becomes entirely apparent if we try to find the counterpart of the elements of S in the operator theory.

For this purpose we suppose that we have satisfied (1) by matrices with elements

$$\int \phi_i^* q \phi_j \cdot dT, \quad \int \phi_i^* \frac{h}{2\pi i} \cdot \frac{\partial}{\partial q} \phi_j \cdot dT,$$

which do not make H diagonal because the ϕ 's are not solutions of $H\psi = E\psi$. These elements compose the matrix H' , and $H'S = SH$ now reads

$$\sum_{\lambda} \int \phi_i^* H \phi_{\lambda} \cdot dT \cdot S_{\lambda j} = S_{ij} E_j$$

It is solved if we put $\sum_{\lambda} \phi_{\lambda} S_{\lambda j} = \psi_j$, for on substituting this into the above there results

$$E_j \int \phi_i^* \psi_j \cdot dT = E_j S_{ij}$$

whence $S_{ij} = \int \phi_i^* \psi_j \cdot dT$, and this same result is obtained if we multiply $\phi_{\lambda} S_{\lambda j} = \psi_j$ by

both sides with ϕ_i^* and integrate. By closer examination we can show that we have constructed the only solution.

S_{ij} thus emerges as the i 'th coefficient in the development of the j 'th energy eigenfunction in terms of the state-functions ϕ .

The results of matrix mechanics contribute nothing in the way of basic information, but the matrix point of view does simplify many problems. The two views are essentially equivalent. Instead of saying: a system is in an eigenstate with respect to the energy, we say that the energy is diagonal.

In many physical problems an exact solution of the operator equation is difficult to obtain. It may then become necessary to use approximate methods in order to find the answer. The details of one such method which has proved highly fruitful will now be discussed in connexion with Schrödinger's equation. Its application to any operator equation is immediate, but limitation to the energy equation is appropriate, partly for the sake of definiteness, partly because the method is rarely applied in practice to any other case.

Perturbation Theory

Let us suppose initially that the eigenfunction of the energy operator is non-degenerate for the state in question. To provide a specific physical basis for our discussion we think of a physical system, like an atom, subject to a "perturbation", that is, to forces which shift the energy levels slightly without changing appreciably their general arrangement. Mathematically, the effect of this perturbation is to introduce added terms into the Hamiltonian H of the system. These added terms may be constants, or functions of the q 's, or functions of the P 's and q 's. In the last case ~~th~~ they appear as differential operators in the Schrödinger equation, otherwise as ordinary functions. Whatever they are, we shall denote them by the symbol V , without implying, of course, that V is anything in the nature of a scalar potential. If we then suppose that the unperturbed equation

$$(H - E_k^0)\psi_k = 0$$

is solved, our problem is to find solutions of the perturbed equation

$$(H + V - E_k)\phi_k = 0$$

In particular we wish to know the ϕ_k , i.e. the eigenfunctions of the perturbed problem, and the E_k 's, which we suppose to be slightly, but not greatly, different from the E_k^0 of the unperturbed. V is considered to be a "small" operator, and by this we mean that the matrix elements of V , formed with the use of the complete system ψ_k are small compared with the diagonal elements of H in this system, i.e. with the E_k^0 . We then approximate the ϕ 's by the series

$$\phi_k = \sum_{\lambda} a_{k\lambda} \psi_{\lambda} + \sum_{\lambda} b_{k\lambda} \psi_{\lambda} + \sum_{\lambda} c_{k\lambda} \psi_{\lambda} + \dots$$

and the energies by

$$E_k = E_k^{(0)} + E_k^{(1)} + E_k^{(2)} + \dots$$

and determine the various a 's, b 's, and c 's. To effect the calculation, we assume that

$$a \quad b \quad c \quad E_k^{(0)} \quad E_k^{(1)} \quad E_k^{(2)} \quad \text{by arrangement}$$

unless b and $E_k^{(1)}$ vanish, in which case we simply suppose a , c , etc. We shall not be interested in the higher approximations. If the results on application of the method are inconsistent with these conditions, our scheme fails and the results are not significant. To avoid circumlocution, we shall speak of the a 's and $E_k^{(0)}$'s as quantities of the zeroth order, the b 's and $E_k^{(1)}$'s as of the first order, etc. The matrix elements of V are assumed to be of the first order.

Upon substitution into the perturbed equation we get

$$\begin{aligned} & \sum_{\lambda} (H - E_k^{(0)}) a_{k\lambda} \psi_{\lambda} \\ & + \sum_{\lambda} (H - E_k^{(0)}) b_{k\lambda} \psi_{\lambda} + \sum_{\lambda} (V - E_k^{(1)}) a_{k\lambda} \psi_{\lambda} \\ & + \sum_{\lambda} (H - E_k^{(0)}) c_{k\lambda} \psi_{\lambda} + \sum_{\lambda} (V - E_k^{(1)}) b_{k\lambda} \psi_{\lambda} + - E_k^{(2)} \sum_{\lambda} a_{k\lambda} \psi_{\lambda} + \dots = 0 \end{aligned}$$

The terms are arranged in rows according to their order of smallness. Let us multiply this equation by ψ_j^* and integrate over all co-ordinates of the problem. Then if

$$V_{jk} = \int \psi_j^* V \psi_k \cdot dT$$

$$\begin{aligned} & (E_j^{(0)} - E_k^{(0)}) a_{kj} \\ & + (E_j^{(0)} - E_k^{(0)}) b_{kj} + \left(\sum_{\lambda} a_{k\lambda} V_{j\lambda} - E_k^{(1)} a_{kj} \right) \\ & + (E_j^{(0)} - E_k^{(0)}) c_{kj} + \sum_{\lambda} b_{k\lambda} V_{j\lambda} - E_k^{(1)} b_{kj} - E_k^{(2)} a_{kj} + \dots = 0 \end{aligned}$$

In the zeroth approximation we can neglect all but the first row of this equation, which is to be satisfied for any value of j . Putting the first row equal to 0 for $j \neq k$, we find that $a_{kj} = 0$ because the energy difference is then finite. If $j = k$, the equation is identically satisfied and does not yield a value for a_{kk} . We can determine it, however, if we recall that ϕ_k is to be normalised. To this approximation, the series for ϕ_k tells us that $\phi_k = a_{kk} \psi_k$; hence, because ψ_k is already normalised, $a_{kk} = 1$ and our first result is therefore that $a_{kj} = \delta_{kj}$. In accordance with the usual process of making successive approximations we substitute this answer back and solve again, this time retaining the next order of terms. The first row now vanishes, of course, and if $j \neq k$, we get, using $a_{kj} = \delta_{kj}$,

$$b_{kj} = \frac{V_{kj}}{E_k^{(0)} - E_j^{(0)}}, \quad j \neq k$$

If $j = k$,

$$E_k^{(1)} = V_{kk}$$

This constitutes a first-order approximation. To get the second, we substitute back and solve again, retaining all the terms written down this time. Now the first two rows vanish and to make the third row zero we proceed as before. In $j=k$, we get

$$E_k^{(2)} = \sum_{\lambda} b_{k\lambda} V_{k\lambda} - b_{kk} E_k^{(1)}$$

This in virtue of the above, reduces to $E_k^{(2)} = \sum_{\lambda} \frac{V_{k\lambda} V_{\lambda k}}{E_k^{(0)} - E_{\lambda}^{(0)}}$

The prime attached to the summation sign is to indicate that the term for which the two indices are equal is to be omitted. On putting $j \neq k$ we get an approximation for c_{jk} in which, however, we are not interested as we are limiting ourselves to two approximations. Higher approximations may be obtained by continuing this process. We are particularly interested in the energy changes given above. To the first approx., the increments of the energies due to the perturbation are simply the diagonal elements of the "perturbation matrix". To the second, they involve all unperturbed frequencies $E_k^{(0)} - E_\lambda^{(0)} = h\nu_{k\lambda}$, where $\nu_{k\lambda}$ is the emission frequency of the transition from $E_k^{(0)}$ to $E_\lambda^{(0)}$. It is interesting to compare these results with those used by astronomers to calculate, for instance, the effect of the moon's perturbation on the earth's orbit. The formulae are essentially identical.

The results will be illustrated by applying them to the Stark effect. The simple Schrödinger equation is then that of an atom (with one electron, for simplicity). V is the additional potential energy of this electron in a constant electric field F . If this lies along the x -axis,

$$V = - eFx$$

From the above, the first approximation in the energy is

$$E_k^{(1)} = - eF \int \psi_k^* x \psi_k \cdot d\tau$$

This energy change represents what is known as the first-order Stark effect. Let us suppose that k denotes the lowest energy state and that the atom is hydrogen. (This state is non-degenerate). This state is then a function of r alone, and if a product of a function of r alone is integrated with x over all space, the result is zero. The ground state is thus unaffected by the field to this order of approximation; it exhibits no "first-order" Stark effect. The same is true for the normal states of all other atoms and in all other states which have a spherically-symmetric energy state-function.

It is not true, however, that $\int \psi_k^* x \psi_\lambda \cdot d\tau$ vanishes if $k \neq \lambda$. Hence the "second-order" Stark effect is given by

$$E_k^{(2)} = e^2 F^2 \sum_{\lambda} \frac{x_{k\lambda} x_{\lambda k}}{E_k^{(0)} - E_\lambda^{(0)}}$$

which can be calculated for any atom if the energy eigenfunctions and the energy levels are known. In classical physics, the increment in energy of an atom due to a static electric field in the absence of a dipole moment is usually written in the form

$$E_k^{(2)} = - \frac{p_k}{2} \cdot F^2$$

The quantity p_k is called the polarisability of the k 'th state. Comparing this expression with that given above, we find that the polarisability is given by

$$p_k = 2 \cdot e^2 \sum_{\lambda} \frac{x_{k\lambda} x_{\lambda k}}{E_\lambda^{(0)} - E_k^{(0)}}$$

This has been tested experimentally for many instances and found to be correct.

We now consider the theory in the case where the eigenfunctions are degenerate.

The above theory is inapplicable when some of the energies $E_\lambda^{(0)}$, over which the sum is to be extended, coincide with $E_k^{(0)}$, i.e. if the state ψ_k for which the perturbation is to be computed is degenerate, since in this case some of the denominators vanish and we can no longer be sure that $E_k^{(0)} \gg E_k^{(2)}$ is satisfied. It is somewhat fortunate that practice has shown a first approximation to be sufficient for most purposes, so that the method need only be worked out this far.

Let us suppose that the state, the perturbations of which we wish to investigate, possess an s -fold degeneracy, so that there are s eigenfunctions for the eigenvalue E_k^0 . We label these $\psi_{k1}, \psi_{k2}, \dots, \psi_{ks}$. They are considered to be orthogonalised by the Schmidt process, so that not only

$$\int \psi_{k\lambda}^* \psi_{k1} \cdot dT = 0, \text{ but also } \int \psi_{k\lambda}^* \psi_{k\mu} \cdot dT = 0 \text{ if } \lambda \neq \mu$$

The index l here stands for any other state belonging to some other energy $E_l^{(0)}$. The solutions of the "unperturbed" equations $(H - E_k^0)\psi_{k\lambda} = 0, \lambda = 1, 2, \dots, s$

$$(H - E_l^0)\psi_l = 0 \quad \text{are known. The states } \psi_l$$

may or may not be degenerate.

Solutions of the equation $(H - E_{kh} + V)\phi_{kh} = 0$ are required. It is now necessary to add another subscript h to E_k , for the energies will in general not be degenerate after the perturbation. Let us suppose that:

$$\phi_{kh} = \sum_{\lambda=1}^s \alpha_{h\lambda} \psi_{k\lambda} + \sum_{l \neq k} a_l \psi_l + \sum_{\lambda=1}^s \beta_{h\lambda} \psi_{k\lambda} + \sum_{l \neq k} b_l \psi_l + \dots \text{ and } E_{kh} = E_k^0 + E_{kh}^{(1)} + \dots$$

The α 's and a 's are considered to be of the zeroth order, the β 's and b 's of the first. It is merely for convenience and clarity that the sums over the degenerate k 'th state have been separated from the others. Substitution of these sums in the equation now yields:

$$\begin{aligned} (H - E_k^0) \left\{ \sum_{\lambda=1}^s \alpha_{h\lambda} \psi_{k\lambda} + \sum_{l \neq k} a_l \psi_l \right\} \\ + (H - E_k^0) \left\{ \sum_{\lambda=1}^s \beta_{h\lambda} \psi_{k\lambda} + \sum_{l \neq k} b_l \psi_l \right\} \\ + (V - E_{kh}^{(1)}) \left\{ \sum_{\lambda=1}^s \alpha_{h\lambda} \psi_{k\lambda} + \sum_{l \neq k} a_l \psi_l \right\} + \dots = 0 \end{aligned}$$

Terms in the first row are of zeroth order, those in the second and third of the first. If we multiply on the left by ψ_m , where $m \neq k$, and integrate, neglecting the second and third lines gives every $a_l = 0$. Thus we may, to the approximation here sought, neglect all the a 's.

If we now multiply by $\psi_{k\mu}^*$, any one of the s degenerate functions, and integrate. The terms in $(H - E_k^0)$ vanish, and there remains:

$$\sum_{\lambda=1}^s \alpha_{h\lambda} (V_{\mu\lambda} - E_{kh}^{(1)} \delta_{\mu\lambda}) = 0 \quad (V_{\mu\lambda} = \int \psi_{k\mu}^* V \psi_{k\lambda} \cdot dT)$$

This equation must hold for every value of μ from 1 to s . It is therefore a system of s linear equations from which the coefficients $\alpha_{h\lambda}$ are to be determined. We can simplify

the writing if we omit the fixed indices k (designating the group of originally degenerate levels) and h (numbering the members of this group). These equations are then:

$$(V_{11} - E^{(1)})\alpha_1 + V_{12}\alpha_2 + V_{13}\alpha_3 + \dots + V_{1s}\alpha_s = 0$$

$$V_{21}\alpha_1 + (V_{22} - E^{(1)})\alpha_2 + V_{23}\alpha_3 + \dots + V_{2s}\alpha_s = 0$$

$$\vdots$$

$$V_{s1}\alpha_1 + V_{s2}\alpha_2 + V_{s3}\alpha_3 + \dots + (V_{ss} - E^{(1)})\alpha_s = 0$$

These can be solved for the α 's only if the determinant $\Delta(V, E^{(1)})$ vanishes. The expansion of this determinant produces an equation of the s 'th degree in $E^{(1)}$, and this has at most s different roots. Hence there are at most s different values of $E^{(1)}$ for which the set given above can be satisfied. They are the new energies which we set out to find. If some of the roots coincide, some of these coincide: the perturbation has not removed the degeneracy completely. $\Delta = 0$ is often called the "secular" equation for the determination of the perturbed energies. If V is zero, it has the s equal solutions $E^{(1)} = 0$; as the $V_{\mu\lambda}$ increase in magnitude the energies move apart. Figuratively speaking, the perturbation forces the s initially coincident energies apart. For each $E_k^{(1)}$ there is one set of α 's, s in number. Altogether, this method allows the calculation of s^2 coefficients α , which is of course just the right number. The remaining coefficients, the b 's and β 's need not be calculated. Since the a 's are zero, the full zeroth approx. has been obtained.

A study of the Zeeman effect provides an illustration of the method discussed. Let us consider again an atom with one electron in a central field. The eigenfunctions are

$$\psi_{k\lambda} = c_1 R(r) \cdot c_2 P_{\frac{1}{2}}^m(\cos\theta) \cdot c_3 e^{im\phi}$$

They possess a $(2\frac{1}{2} + 1)$ -fold degeneracy corresponding to the fact that the energy does not depend on m , and there are $(2\frac{1}{2} + 1)$ possible values of m for a given integer $\frac{1}{2}$. Hence λ runs from 1 to $s = 2\frac{1}{2} + 1$, and the various degenerate eigenfunctions can be numbered in such a way that $\lambda = 1$ corresponds to $m = -\frac{1}{2}$, $\lambda = 2$ to $m = -\frac{1}{2} + 1$, etc. Now it can be shown that, if a uniform magnetic field \underline{H} along the z -axis is present, the added term in the Hamiltonian is

$$V = \frac{he}{4\pi i M c} \cdot \underline{H} \cdot \frac{\partial}{\partial \phi}, \text{ where } M \text{ is the electronic mass.}$$

In forming the matrix of V between the degenerate states every integral will contain the part:

$$c_1^2 c_2^2 \int [R(r) P_{\frac{1}{2}}^m(\cos\theta)]^2 r^2 \cdot \sin\theta \cdot dr d\theta, \text{ which is unity since each}$$

of the three factors of ψ is normalised individually. The remainder of each integral will be:

$$\frac{heH}{4\pi i M c} \cdot \frac{1}{2\pi} \int_0^{2\pi} e^{-im_1\phi} \cdot \frac{\partial}{\partial \phi} e^{im_2\phi} \cdot d\phi = \frac{heH}{4\pi i M c} \cdot m_2 \cdot \frac{1}{2\pi} \int_0^{2\pi} e^{i(m_2 - m_1)\phi} d\phi$$

Here m_1 and m_2 are the two m values for the states between which the matrix element is taken, and $1/(2\pi)^{\frac{1}{2}}$ is the normalisation factor c_3 . This integral is zero unless $m_2 = m_1$.

We see that all the non-diagonal elements of V vanish and that the diagonal ones are:

$V_{11} = -A\bar{l}$, $V_{22} = -A(\bar{l} - 1)$, $V_{33} = -A(\bar{l} - 2)$, ... $V_{ss} = A\bar{l}$, where A is $he\bar{H}/4\pi Mc$. The solutions of the secular equation are clearly

$$E^{(1)} = mA = m \cdot \frac{he\bar{H}}{4\pi Mc}, \text{ where } m \text{ can take any integral value}$$

from $-\bar{l}$ to $+\bar{l}$. Thus every degenerate state with azimuthal quantum number \bar{l} will, under the influence of the field, split up into $2\bar{l} + 1$ sublevels, all equally spaced and grouped symmetrically about the unperturbed level $E^{(1)} = 0$, the energy difference being $he\bar{H}/4\pi Mc$. What effect will this have on the appearance of spectral lines emitted in the presence of a field?

A spectral line which is due to the transition of an atom from the energy state $\bar{l}=2$ to $\bar{l}=1$, the former having 5 m-values and the latter 3, might be expected to split up into ~~fifteen lines~~ seven lines corresponding to the fifteen different possible transitions. This overlooks, however, the fact that the transition probabilities are given by the squares of the matrix elements of the co-ordinates. If we calculate these we find that they vanish between two states whose m-values differ by more than one. Transitions which violate this rule have a zero probability, and hence do not take place. In this way, the number of possible transitions is limited to nine, with three distinguishable energy gaps, giving three lines, one of the frequency of the original, one less, and one greater. Hence the line should appear as a symmetrical triplet, and the frequency separation between its components should be

$$\Delta w = \frac{e\bar{H}}{2Mc}$$

This quantity is indeed the Larmor frequency.

This theory, giving the "normal" Zeeman effect is in perfect agreement with experiment. In general, the theory is not as simple as this however, for the electron has magnetic properties not considered in the expression for V . It possesses indeed a magnetic moment (spin) which we have not taken into account. The theory, then, only describes the Zeeman effect in its simplest aspects.

1. Show that if the two Hermitean operators A and B satisfy the commutation relation $AB - BA = iC$, the following relation will hold:

$$(\Delta A)^2 (\Delta B)^2 \geq \frac{1}{4} |C|^2$$

2. Find the uncertainty relation for the operators Q and F(P) if Q and P satisfy the commutation relation $QP - PQ = i\hbar/2\pi$.
3. Use the Heisenberg relations to find the ground state of a harmonic oscillator approximately.
4. Inasmuch as the Schrödinger equation is a first-order differential equation with respect to time, $\Psi(t)$ is uniquely determined by the value of $\Psi(0)$. Write this connexion in the form:

$$\Psi(t) = S(t)\Psi(0) \quad \text{where } S(t) \text{ is some operator.}$$

- a) Show that the operator $S(t)$ satisfies the equation

$$\frac{i\hbar}{2\pi} \frac{d}{dt} S(t) = HS(t)$$

and is unitary operator.

- b) Show that in the case where H does not depend on time, $S(t)$ is of the form

$$S(t) = \exp(-2\pi i H t / \hbar)$$

5. The average value of an operator L at time t follows from the expression:

$$\bar{L}(t) = \int \Psi^*(t) L \Psi(t) . dT$$

- a) Show that the time dependence of the operator $L' = S^{-1}(t) L S(t)$, with $S(t)$ determined as before, satisfies the equation

$$\int \Psi^*(0) L' \Psi(0) . dT = \bar{L}(t)$$

- b) Prove the following:

$$i\hbar \dot{L}' = L' H' - H' L' \quad \text{where } H' = S^{-1} H S$$

- c) Show that if the operators L and M satisfy the commutation relation

$$LM - ML = iN$$

the corresponding time-dependent operators satisfy the equation

$$L'M' - M'L' = iN'$$

6. Determine the time dependence of the co-ordinate operator x for

- a) a free particle
b) an oscillator