Rick's Formulation of Quantum Mechanics: QM1

Hilbert Space: Pure States, Eigenstates, Evolution & Interpretation

Last Update: 15th June 2008

1. Hilbert Space

Quantum mechanics holds that the state of a physical system can be represented by a vector in Hilbert space, *H*. Hilbert space is a linear vector space over the complex numbers. So any vector can be multiplied by a complex number to obtain another vector, and vectors can be added together. Writing vectors in the Dirac ket notation, $|\psi\rangle$, this means that for any vectors $|\psi\rangle$ and $|\phi\rangle$ and any complex numbers a and b, it is possible to combine them to form another vector $a|\psi\rangle + b|\phi\rangle$. There is a unique zero vector, which can be obtained by multiplying any vector by the number zero, $|\text{zero}\rangle = 0|\psi\rangle$, and which leaves any vector unchanged when added to it: $|\psi\rangle + |\text{zero}\rangle = |\psi\rangle$.

NB: We shall often use the notation $|0\rangle$ to represent either the ground state of a system or perhaps a particular spin state of a spinning particle. In neither case should this state be confused with the zero vector, $|\text{zero}\rangle$. A physical system cannot be presented by the zero vector, since a physical system is always normalisable, i.e. it has non-zero norm (see below).

In addition to the above laws of combination, which are essentially the definition of a linear vector space, a Hilbert space also has a scalar product. This is a mapping from pairs of vectors to the positive indefinite real numbers: $H \otimes H \to \Re_{\geq 0}$.

A Hilbert space can be spanned by a set of orthonormal vectors $\{|\phi_i\rangle\}$, which means that any vector can be expressed as $|\psi\rangle = \sum_i a_i |\phi_i\rangle$ for some complex numbers a_i.

Orthonormality of the basis vectors means that under the scalar product they give δ_{ij}.

A conjugate Hilbert space is defined, in which every vector $|\psi\rangle = \sum_i a_i |\phi_i\rangle$ has a corresponding conjugate vector (a 'bra' corresponding to a 'ket') which is written $\langle\psi| = \sum_i a_i^* \langle\phi_i|$. The scalar product is then completely defined by $\langle\phi_i|\phi_j\rangle = \delta_{ij}$.

This gives, $\langle\psi|\xi\rangle = \sum_i a_i^* b_i$ and hence, $\langle\xi|\psi\rangle = \langle\psi|\xi\rangle^*$ where, $|\psi\rangle = \sum_i a_i |\phi_i\rangle$ and $|\xi\rangle = \sum_i b_i |\phi_i\rangle$. Hence also $\langle\psi|\psi\rangle \equiv \|\psi\|^2 = \sum_i a_i^* a_i$, where $\|\psi\|$ is called the 'norm' (or 'length' or 'magnitude') of the vector $|\psi\rangle$. Hence, the norm of any non-zero vector is positive definite, the zero vector alone having zero norm.

In contrast, the scalar product of two different vectors is, in general, complex. When real it can be negative.

In the particular basis $\{|\phi_i\rangle\}$ we may think of $|\psi\rangle = \sum_i a_i |\phi_i\rangle$ as being represented by a column matrix consisting of the numbers a_i, and the conjugate vector $\langle\psi| = \sum_i a_i^* \langle\phi_i|$ as represented by a row matrix containing the complex conjugate numbers a_i^*.

The scalar product can thus be regarded as the ordinary matrix multiplication of the vectors as represented by their row/column matrices, i.e.,

$$\langle\psi|\xi\rangle \equiv \begin{pmatrix} a_1^* & a_2^* & a_3^* \ldots\ldots \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \text{etc.} \end{pmatrix} = \sum_i a_i^* b_i$$

A "linear operator" on Hilbert space is a linear mapping which takes any vector to another vector, $\hat{Q} : H \to H$. Thus the result of acting the operator on any vector $|\psi\rangle$ is $\hat{Q}|\psi\rangle \in H$. Linearity means that for any pair of vectors $\hat{Q}[a|\psi\rangle + b|\xi\rangle] \equiv a\hat{Q}|\psi\rangle + b\hat{Q}|\xi\rangle$. Henceforward we shall be lazy and speak of "operators" when we exclusively mean "linear operators". Note that linearity of the action of the operator on the Hilbert space does not imply algebraic linearity of the operator, e.g., linearity of $\hat{Q}$ implies linearity of $\hat{Q}^2, \hat{Q}^3$, etc.

Any operator can be fully defined by its operation on a basis set of vectors, i.e. by the square matrix of complex numbers $Q_{ij} = \langle\phi_i|\hat{Q}|\phi_j\rangle$. Note that we can also write $\hat{Q}|\phi_j\rangle = \sum_i Q_{ij}|\phi_i\rangle$, due to the assumed orthonormality of the basis $\{|\phi_i\rangle\}$. Note the order of the indices on Q_{ij}. Hence, for an arbitrary vector, $|\psi\rangle$,

$$\hat{Q}|\psi\rangle = \hat{Q}\sum_i a_i|\phi_i\rangle = \sum_i a_i \hat{Q}|\phi_i\rangle = \sum_i a_i \sum_j Q_{ji}|\phi_j\rangle = \sum_j a'_j|\phi_j\rangle$$

where $a'_j = \sum_i Q_{ji} a_i$. This means that if we consider the numbers Q_{ji} as a square matrix, (Q), then the action of the operator is to transform the components of a vector as per ordinary matrix multiplication, i.e. $\bar{a}' = (Q)\bar{a}$. Note that all the complex numbers a_i, a_i^* and Q_{ji} are with respect to the specified basis $\{|\phi_i\rangle\}$.

Conversely, given any square matrix of complex numbers, Q_{ji}, we can use it to define a linear operator by defining the operator's action as being $\hat{Q}|\phi_i\rangle = \sum_j Q_{ji}|\phi_j\rangle$. One means of specifying a matrix is to form the product of a column and a row vector,

$$(Q) \equiv \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \text{etc.} \end{pmatrix} \begin{pmatrix} a_1^* & a_2^* & a_3^* ..... \end{pmatrix}$$

so that $Q_{ji} = b_j a_i^*$. We can consider the column and row matrices as being the components of vectors, $|\xi\rangle = \sum_j b_j |\phi_j\rangle$ and $|\psi\rangle = \sum_j a_j |\phi_j\rangle$, so that we can write,

$$Q_{ji} \equiv \langle\phi_j|\hat{Q}|\phi_i\rangle = \langle\phi_j|\xi\rangle\langle\psi|\phi_i\rangle$$

From this there is an obvious notation for operators as $\hat{Q} = |\xi\rangle\langle\psi|$.

Note that it follows from the above definitions that any operator can be written in terms of its components wrt basis $\{|\phi_i\rangle\}$ as,

$$\hat{Q} \equiv \sum_{i,j} Q_{ij} |\phi_i\rangle\langle\phi_j|$$

2. Special Operators

The operator $\hat{P}_i = |\phi_i\rangle\langle\phi_i|$ projects out the i'th component of any vector, i.e.,

$$\hat{P}_i|\psi\rangle = |\phi_i\rangle\langle\phi_i|\sum_j a_j|\phi_j\rangle = \sum_j a_j|\phi_i\rangle\langle\phi_i|\phi_j\rangle = \sum_j a_j|\phi_i\rangle\delta_{ij} = a_i|\phi_i\rangle$$

So summing over all such projection operators gives the unit operator: $\sum_i \hat{P}_i = I$, where I is such that $I|\psi\rangle = |\psi\rangle$ for all vectors.

More generally, if $|\psi\rangle$ is a normalised state, i.e. it has a norm of unity, then $|\psi\rangle\langle\psi|$ is a projection onto the direction of $|\psi\rangle$. Hence, the result of applying this operator to any vector is that the component of this vector projected onto $|\psi\rangle$ results. This is just another way of saying that $(|\psi\rangle\langle\psi|)|\xi\rangle = |\psi\rangle\langle\psi|\xi\rangle \equiv (\langle\psi|\xi\rangle)|\psi\rangle$.

Rick's Formulation of Quantum Mechanics: QM1

Any projection operator, $\hat{P}$, is idempotent, i.e., it obeys $\hat{P}^2 = \hat{P}$. Physically this is because projecting again does nothing further. Mathematically, $|\psi\rangle\langle\psi\|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi|$.

For any $\hat{Q}$ the "adjoint" operator is defined via the adjoint matrix of elements. Thus the adjoint of $\hat{Q} \equiv |\phi_i\rangle Q_{ij} \langle\phi_j|$ is written, $\hat{Q}^+ \equiv |\phi_i\rangle Q^*_{ji} \langle\phi_j|$. For example, $\left(|\psi\rangle\langle\xi|\right)^+ = |\xi\rangle\langle\psi|$.

An operator is said to be Hermetian, or self-adjoint, if it equals its adjoint, $\hat{H}^+ = \hat{H}$, and this is equivalent to its matrix of components being Hermetian, i.e. $\left(H^+\right)_{ij} = H^*_{ji} = H_{ij}$, or simply $(H)^+ = (H)$. It is simple to demonstrate that being self-adjoint wrt one basis $\{|\phi_i\rangle\}$ implies being self-adjoint in any other basis, so that being self-adjoint is a property of the operator, not of the basis.

The inverse of an operator reverses its action. Thus, if $\hat{Q}|\psi\rangle = |\xi\rangle$ then $\hat{Q}^{-1}|\xi\rangle = |\psi\rangle$ for any vector $|\psi\rangle$. Not all operators have an inverse. For example, a projection operator clearly does not have an inverse. For example, $\hat{P}_1|\psi\rangle = a_1|\phi_1\rangle$ looses all information about the other components of $|\psi\rangle$, i.e. a_2, a_3,…, and hence clearly cannot be inverted. The components of an inverse operator are the inverse matrix of the components of the operator, i.e. $\hat{Q}^{-1} = |\phi_i\rangle\left((Q)^{-1}\right)_{ij}\langle\phi_j|$. The non-existence of an inverse operator is therefore equivalent to the non-existence of the inverse matrix, i.e. that the determinant $\|(Q)\|$ is zero. For example, the matrix for $\hat{P}_1$ is $\begin{matrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \text{etc.} \end{matrix}$, which clearly has zero determinant.

A Unitary operator is one whose adjoint equals its inverse, $\hat{U}^+ = \hat{U}^{-1}$. The same expression holds for the matrix of components. Unitary operators have the property that they preserve the norm of a vector. Thus, if $|\xi\rangle = \hat{U}|\psi\rangle$ then $\langle\xi|\xi\rangle = \langle\psi|\psi\rangle$. This follows simply from $\langle\xi|\xi\rangle = \langle\psi|\hat{U}^+\hat{U}|\psi\rangle = \langle\psi|\hat{U}^{-1}\hat{U}|\psi\rangle = \langle\psi|I|\psi\rangle = \langle\psi|\psi\rangle$, because $\left(\hat{Q}|\xi\rangle\right)^+ = \langle\xi|\hat{Q}^+$.

An operator which transforms from one orthonormal basis to another must be Unitary. Proof: put $|\psi_i\rangle = \hat{Q}|\phi_i\rangle$ where both $|\psi_i\rangle$ and $|\phi_i\rangle$ form orthonormal bases. It follows that, $\langle\psi_j|\psi_i\rangle = \delta_{ji} = \langle\phi_j|\hat{Q}^+\hat{Q}|\phi_i\rangle = \langle\phi_j|\hat{Q}^+\sum_k|\phi_k\rangle\langle\phi_k|\hat{Q}|\phi_i\rangle = \sum_k\left(Q^+\right)_{jk}(Q)_{ki}$, so that the matrix of components of Q^+ is seen to be the inverse of that of Q, hence Q is Unitary.

3. Eigenvectors and Eigenvalues

The eigenvalues, q_i, and eigenvectors, $|q_i\rangle$, of an operator $\hat{Q}$ are defined via $\hat{Q}|q_i\rangle = q_i|q_i\rangle$, excluding the trivial zero vector for which any q_i would obey the equation.. The eigenvalues can be found by solving the "secular equation", which is the requirement that, for a non-zero eigenvector, the determinant of $(Q) - q_i(I)$ must be zero, where (I) is the unit matrix. Thus, for example, $\hat{P}_1$ has one eigenvalue equal to 1, and the rest are 0. Generally, projection operators can only have eigenvalues 0 or 1. This is easily seen from $\hat{P}^2|\psi\rangle = \hat{P}|\psi\rangle = \lambda|\psi\rangle = \lambda^2|\psi\rangle$, and hence $\lambda = \lambda^2$.

Eigenvectors are defined only up to an arbitrary multiplicative constant. This is reduced to an arbitrary phase factor ($e^{i\theta}$) by requiring that states be normalised to unity, $\langle q_i|q_i\rangle = 1$.

The set of all eigenvalues of a given operator is called its spectrum.

It is simple to show that Hermetian operators have real eigenvalues and that their eigenvectors are mutually orthogonal. Proof: put $\hat{H}|\psi_i\rangle = \lambda_i|\psi_i\rangle$. Then, $\langle\psi_j|H^+ = \langle\psi_j|H = \lambda_j^*\langle\psi_j|$. Hence, $\langle\psi_j|H|\psi_i\rangle = \lambda_i\langle\psi_j|\psi_i\rangle = \lambda_j^*\langle\psi_j|\psi_i\rangle$. In the case i = j the norm $\langle\psi_i|\psi_i\rangle$ cannot be zero and hence we have $\lambda_i = \lambda_i^*$, i.e., the eigenvalues are real. However, if the eigenvalues are different, and knowing now that they are real, the above expression can hold only if $\langle\psi_j|\psi_i\rangle = 0$, i.e. the eigenvectors are orthogonal. A subtlety occurs for degenerate eigenvalues, i.e., when distinct eigenvectors have the same eigenvalue. In such as case, any linear combination of the degenerate eigenvectors is also an eigenvector with the same eigenvalue. But this "degenerate subspace" can be spanned by a mutually orthogonal set of eigenvectors, so the result is the essentially same. We merely assume that some (arbitrary) orthonormal basis of this subspace has been chosen[1].

It is also simple to show that Unitary operators have eigenvalues which are complex numbers of modulus unity, i.e. pure phase factors of the form $e^{i\theta}$ for real θ, and whose eigenvectors are mutually orthogonal. Proof: put $\hat{U}|\psi_i\rangle = \lambda_i|\psi_i\rangle$. Then, $\langle\psi_j|U^+ = \langle\psi_j|U^{-1} = \lambda_j^*\langle\psi_j|$. Hence, $\langle\psi_j| = \lambda_j^*\langle\psi_j|U$, and thus, $\langle\psi_j|\psi_i\rangle = \lambda_j^*\langle\psi_j|U|\psi_i\rangle = \lambda_j^*\lambda_i\langle\psi_j|\psi_i\rangle$. Hence, in the case i = j the norm $\langle\psi_i|\psi_i\rangle$ cannot be zero and hence we have $\lambda_i^*\lambda_i = 1$, i.e. the eigenvalues are of unit modulus as claimed.

[1] A more virulent form of this subtlety occurs for zero eigenvalues. Any vector would obey the eigenvalue equation when $q_i = 0$, including a 'true' eigenvector with a non-zero eigenvalue. It is unwelcome to have two eigenvalues for a single eigenvector. So, in such a case, only the subspace orthogonal to all the eigenvectors with non-zero eigenvalues is regarded as home to eigenvectors with zero eigenvalues.

Where the eigenvalues are different we thus have $\lambda_j^* \lambda_i = e^{i(\theta_i - \theta_j)}$, and the exponent is non-zero (generally not a multiple of 2π), so this cannot equal unity. Hence to satisfy the above expression we must have $\langle \psi_j | \psi_i \rangle = 0$, i.e. the eigenvectors are orthogonal. As for Hermetian matrices there is a subtlety for degenerate eigenvalues, but this is salvaged in the same manner by adopting a mutually orthogonal set of eigenvectors which span such degenerate subspaces.

It follows from the above observations that any Hermetian or Unitary operator can be used to define a basis, i.e., its eigenvectors form an orthonormal basis, assuming they have been normalised.

Spectral Representation: Any operator can be written $\hat{Q} \equiv \sum_{i,j} Q_{ij} | \phi_i \rangle \langle \phi_j |$ in terms of an arbitrary orthonormal basis $\{ | \phi_i \rangle \}$. If $\hat{Q}$ is an operator of a type whose eigenvectors form an orthonormal basis, e.g. Hermetian or Unitary, then we can use these eigenvectors as the basis, i.e. use $\{ | \phi_i \rangle \} = \{ | q_i \rangle \}$. In that case $Q_{ij} = \langle q_i | \hat{Q} | q_j \rangle = q_i \delta_{ij}$ and hence the matrix of components is diagonal in this basis, $\hat{Q} \equiv \sum_i q_i | q_i \rangle \langle q_j |$. This is called the "spectral representation" but is only valid if the operator's eigenvectors form an orthonormal basis.

4. Physical Interpretation

Vectors in Hilbert space represent the state of a physical system. Quantum mechanics holds that, so long as the right Hilbert space is chosen, the vector represents the physical system fully. Hermetian operators on this Hilbert space represent physical observables. The physical action of a measuring device which is constructed so as to measure an observable Q is represented ***in part*** by the action of the corresponding operator $\hat{Q}$ on the vector $| \psi \rangle$ which represents the initial state of the system. Thus, the action of taking a measurement is regarded as having the potential to alter the state of the system, since $\hat{Q} | \psi \rangle$ will in general be a different vector from $| \psi \rangle$. However, the physical interaction which comprises a measurement also involves a second component which is not specified by the operator $\hat{Q}$. More of this later.

The numerical value which results from a measurement has to be one of the eigenvalues of the operator representing the observable, i.e. it has to be one of the discrete set of numbers q_1, q_2, q_3, etc. This is where the discreteness, which is the hallmark of quantum mechanics, originates. The requirement that physical observables be represented by Hermetian operators means that all the possible outcomes of a measurement are *real* numbers.

If the system happens to be in a state which is one of the eigenvectors of the observable being measured, say $| q_i \rangle$, then the outcome of the measurement (i.e. the value assigned to

the observable) will definitely[2] be q_i. Thus, the algebraic relation $\hat{Q}|q_i\rangle = q_i|q_i\rangle$ can be interpreted thus: the LHS represents performing the measurement; the RHS represents getting out the resulting measured value. The Hilbert space operator, $\hat{Q}$, represents the measuring equipment's interaction with the system, and the number, q_i, which pops out is the reading on its dial or screen.

In general, an arbitrary state will be a combination of eigenvectors of the observable, Q, being measured, i.e., $|\psi\rangle = \sum_i a_i|q_i\rangle$. (The existence of such an expression is ensured by the orthonormality of the eigenvectors of an Hermetian operator, see above). Performing the measurement would thus appear to give $\hat{Q}|\psi\rangle = \hat{Q}\sum_i a_i|q_i\rangle = \sum_i a_i\hat{Q}|q_i\rangle = \sum_i a_i q_i|q_i\rangle$
The trouble with this is that no single outcome is forthcoming from the algebra, and yet a physical measurement will yield a definite result – and this result must be one of the eigenvalues q_1, q_2, q_3, etc. This is the famous "measurement problem" (or one aspect of it).

The solution to this problem is known as the Born Rule. This holds that, upon measurement, the vector representing the state of the system undergoes an irreversible change so that it is left in one of the eigenstates, $|q_i\rangle$, of the measuring device (as opposed to a sum over several such eigenstates). Thus, the action of a measuring device is not, in general, fully accounted for by the action of the corresponding operator on the state vector, but there is also a second 'step' which selects and projects out just one of the contributing eigenstates. This can be written symbolically as,

$$\text{Measurement: } \Re_Q|\psi\rangle = \Re_Q\sum_i a_i|q_i\rangle \to |q_k\rangle \text{ with probability } |a_k|^2$$

The symbol $\Re$ does not mean anything which can be interpreted as part of the Hilbert space algebra. Indeed, not only is its mathematical nature unknown, but its physical nature is equally obscure. It is sometimes called "reduction of the wavepacket". Exactly of what this "reduction of the wavepacket" consists has not been resolved. Its nature is part of the measurement problem.

The Born Rule holds that the probability of the result of the measurement being q_k is $|a_k|^2$. Thus, if the initial state were $0.9|q_1\rangle + 0.4359|q_2\rangle$, then the measured value would be found to be q_1 on 81% of occasions, and q_2 on 19% of occasions, but never found to be q_3, q_4, etc. However, the outcome of an individual measurement cannot be predicted with certainty. Thus the Born Rule (or the $\Re$-process) is the origin of the infamous indeterminacy of quantum mechanics.

[2] Subject to experimental uncertainties, of course. We say "definitely" to emphasise that there is no fundamental, quantum mechanical, obstacle to obtaining a result which is 100% predictable and reproducible.

It would seem that the $\Re$-process is also the sole cause of the entropy of the universe increasing (more of this in Part 6 of these QM notes), and therefore also responsible for the arrow of time.

From the Born interpretation springs the requirement that the vectors representing physical states be normalised to unity, since the interpretation of the $|a_k|^2$ as probabilities means that we need to ensure that $\sum_i |a_i|^2 = \langle\psi|\psi\rangle = 1$.

Since each possible measurement outcome, q_i, has probability $p_i = |a_i|^2$, it follows that the expectation value of Q prior to measurement, if the state is $|\psi\rangle$, is,

$$\langle Q\rangle = \sum_i p_i q_i = \sum_i |a_i|^2 q_i = \langle\psi|\hat{Q}|\psi\rangle$$

Thus, if we transformed to a basis in which $|\psi\rangle$ were one of the basis vectors, then the expectation value of Q for this state would be the corresponding on-diagonal component of the matrix of components of the operator Q.

Note that the expectation value of an observable, i.e. the average outcome of many measurements, is well defined within the Hilbert space algebra. It is only the outcome of individual measurements which is not.

Note also that it is only the $\Re$-process, the "reduction of the wavepacket", that makes quantum mechanics indeterminate. The time evolution of pure quantum states is perfectly determinate (i.e., unitary evolution, see below).

The Wavefunction: Personally, I use the term 'wavefunction' as synonymous with 'state vector'. Historically, it meant the solution to the Schrodinger equation (see below), and hence a function of the position coordinate, $\bar{r}$. In Dirac (Hilbert space) notation we may define a state representing a particle being at position $\bar{r}$ as $|\bar{r}\rangle$. Thus, the wavefunction corresponding to an arbitrary state vector $|\psi\rangle$ is defined as the scalar product $\langle\bar{r}|\psi\rangle$, which is thus a function of position. For example, the state $|\psi\rangle$ may be a state of definite energy, i.e. an eigenstate of the Hamiltonian operator (see below). In this case $\psi(\bar{r}) \equiv \langle\bar{r}|\psi\rangle$ is the representation of this energy eigenstate in configuration space, i.e. as a 'wavefunction'. We may imagine the state $|\psi\rangle$ expanded in terms of the position states, i.e. $|\psi\rangle = \sum_i a_i|\bar{r}_i\rangle$, where we have discretised space and each $|\bar{r}_i\rangle$ may be interpreted as covering some finite volume dV. Hence, a measurement of position carried out on a particle in state $|\psi\rangle$ will yield the result $\bar{r}_i$ with probability $|a_i|^2 = |\langle\bar{r}_i|\psi\rangle|^2 \equiv |\psi(\bar{r})|^2$. Thus,

the Born Rule also yields the interpretation of the wavefunction, i.e., that its absolute square is the probability density of the particle.

The Trace of an operator, $\hat{Q}$, is defined as $\mathrm{Tr}\left(\hat{Q}\right)=\sum_i \langle \phi_i | \hat{Q} | \phi_i \rangle$, i.e. the trace of an operator is the trace of the matrix of its components. It is easily seen that the trace of an operator does not depend upon the basis chosen. Consider bases $\{|\phi_i\rangle\}$ and $\{|\psi_i\rangle\}$ connected by the transformation: $|\psi_i\rangle = U|\phi_i\rangle$, for all i. Then,

$$\begin{aligned}\sum_i \langle \phi_i | \hat{Q} | \phi_i \rangle &= \sum_i \langle \psi_i | U^+ \hat{Q} U | \psi_i \rangle = \sum_{i,k,n} \langle \psi_i | U^+ | \psi_k \rangle \langle \psi_k | \hat{Q} | \psi_n \rangle \langle \psi_n | U | \psi_i \rangle \\ &= \sum_{i,k,n} \langle \psi_n | U | \psi_i \rangle \langle \psi_i | U^+ | \psi_k \rangle \langle \psi_k | \hat{Q} | \psi_n \rangle \\ &= \sum_{k,n} \langle \psi_n | UU^+ | \psi_k \rangle \langle \psi_k | \hat{Q} | \psi_n \rangle = \sum_{k,n} \langle \psi_n \| \psi_k \rangle \langle \psi_k | \hat{Q} | \psi_n \rangle \\ &= \sum_{k,n} \delta_{nk} \langle \psi_k | \hat{Q} | \psi_n \rangle = \sum_k \langle \psi_k | \hat{Q} | \psi_k \rangle\end{aligned}$$

thus establishing that the trace is the same in all orthonormal bases. Note that we have used the fact the orthonormal bases are connected by Unitary transformations, as proved above.

The trace can be used to express the expectation value in an alternative way. For this we represent the state of the system, not by the vector $|\psi\rangle$ directly, but rather by the corresponding projection operator, $\hat{\rho} = |\psi\rangle\langle\psi|$. The expectation value of observable Q is then,

$$\mathrm{Tr}\left(\hat{\rho}\hat{Q}\right) = \sum_i \langle \phi_i \| \psi \rangle \langle \psi | \hat{Q} | \phi_i \rangle = \sum_i \langle \psi | \hat{Q} | \phi_i \rangle \langle \phi_i \| \psi \rangle = \langle \psi | \hat{Q} | \psi \rangle$$

as claimed. The operator $\hat{\rho} = |\psi\rangle\langle\psi|$ is called the "density matrix" for the pure state $|\psi\rangle$. Note that we have used $\sum_i |\phi_i\rangle\langle\phi_i| = I$, the unit operator, because of the completeness of the basis.

Which Operators Are Observables?

Only Hermetian operators can represent observables. For a long time it was thought that all Hermetian operators of a given, physically realisable, Hilbert space represented observables. The modern view is that this is no longer tenable, and that only a subset of Hermetian operators are observables. Here my knowledge expires, except that C* algebras enter the picture when this issue is taken further.

The wavefunction (or state vector) itself is *not* an observable. We will see in Parts 10 and 11 of these QM notes why the state vector cannot be observable. It is because this would yield far more information than the state can convey.

5. Dynamics: Evolution in Time (Schrödinger Picture)

There is a particular Hermetian operator, called the Hamiltonian, which represents the observable "total energy", and is denoted $\hat{H}$. Since energy is conjugate to time, the time evolution of the state of the system, $|\psi\rangle$, is given by the Schrödinger equation,

$$i\hbar\frac{\partial}{\partial t}|\psi\rangle = \hat{H}|\psi\rangle \tag{1.3.1}$$

This can be solved formally as, $|\psi(t_2)\rangle = \exp\left\{-\frac{i}{\hbar}\hat{H}(t_2 - t_1)\right\}|\psi(t_1)\rangle$. The operator $\hat{U}(t_2 - t_1) = \exp\left\{-\frac{i}{\hbar}\hat{H}(t_2 - t_1)\right\}$ is therefore the 'propagator', i.e. it propagates the state from one time to another. Note that because $\hat{H}$ is Hermetian it follows that $\hat{U}$ is Unitary.

The fact that the state evolves by the action of a unitary operator, $|\psi(t)\rangle = \hat{U}(t)|\psi(0)\rangle$, is necessary in order that a normalised state evolves into another normalised state. Reversing this reasoning, evolution according to Equ.(1.3.1) appears to be inevitable. The Schrödinger equation is thus a consequence of the probabilistic interpretation of the wave vector.

The eigenvalues of $\hat{H}$ are the possible energy levels of the system, E_1, E_2, etc. Thus the set of all energy levels is the spectrum of $\hat{H}$ (thus revealing the origin of this terminology in atomic spectroscopy).

The eigenvalues of $\hat{U}$ are $\exp\left\{-\frac{i}{\hbar}E_1(t_2 - t_1)\right\}, \exp\left\{-\frac{i}{\hbar}E_2(t_2 - t_1)\right\}$, etc.

If the system is in an eigenstate of $\hat{H}$, say $|E_i\rangle$, then the Schrödinger equation becomes,

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

The solution to which is no longer merely 'formal' but is actually,

$$|\psi(t_2)\rangle = \exp\left\{-\frac{i}{\hbar}E_i(t_2 - t_1)\right\}|\psi(t_1)\rangle$$

Thus, states of well defined total energy, i.e. eigenstates of the Hamiltonian, are stationary in the sense that they only change in time by a phase factor. They are represented by the same 'ray' in Hilbert space (i.e. the same 'direction') at all times.

6. Functions of an Operator

A function $f(\hat{Q})$ of an operator may be taken to be defined either via the power series for the function, e.g.,

$$\exp(\hat{Q}) \equiv 1 + \hat{Q} + \frac{\hat{Q}^2}{2} + \frac{\hat{Q}^3}{3!} + \ldots$$

or by a spectral representation. The latter is when the operator is expanded in terms of the projection operators defined by its eigenvectors, i.e., $\hat{Q} = \sum_i q_i |q_i\rangle\langle q_i|$ (see above). In this case it follows from the power series definition that,

$$f(\hat{Q}) = \sum_i f(q_i) |q_i\rangle\langle q_i|$$

This latter definition only applies if the operator can be expressed in a diagonal form, as per the spectral representation. The power series is the more general definition, but has the disadvantage that it will generally not be easy to evaluate, except when Q is reduced to diagonal form in which case it becomes the same as the spectral representation.

7. Different Pictures

The above description of the dynamics of a system uses the so-called Schrödinger picture, in which the state, i.e. the vector in Hilbert space, is taken to be time dependent but the operators representing the observables are time independent. This is the picture which I have personally always found the most natural. However, it is possible to adopt different 'pictures'. The Heisenberg picture attributes the time dependence to the observables, whilst the state remains unchanged. The expectation values must be the same, and so the two pictures are related by,

$$\langle \psi(t) | \hat{Q}_S | \psi(t) \rangle = \langle \psi(0) | \exp\left\{+\frac{i}{\hbar}\hat{H}t\right\} \hat{Q}_S \exp\left\{-\frac{i}{\hbar}\hat{H}t\right\} | \psi(0) \rangle \equiv \langle \psi(0) | \hat{Q}_H | \psi(0) \rangle$$

which shows that the time-dependent observable-operator in the Heisenberg picture must be given in terms of the Schrödinger observable-operator by,

$$\hat{Q}_H = \exp\left\{+\frac{i}{\hbar}\hat{H}t\right\} \hat{Q}_S \exp\left\{-\frac{i}{\hbar}\hat{H}t\right\}$$

whilst the system state remains unchanged, $|\psi(0)\rangle$, in the Heisenberg picture. The only remaining question is how the dynamics is formulated. This must now be an equation which specifies how $\hat{Q}_H$ changes in time. Actually this is essentially trivial in that the above equation has already given us the *solution* to the correct dynamical equation. By differentiating it wrt time, and assuming for simplicity that the operator in the Schrödinger picture has no time dependence, i.e. that the partial derivative $\frac{\partial \hat{Q}_S}{\partial t}$ is zero, then we easily derive the Heisenberg "equation of motion",

$$\frac{d\hat{Q}_H}{dt} = \frac{i}{\hbar}\left[\hat{H}, \hat{Q}\right]$$

where the square bracket is a commutator, i.e. $\left[\hat{A}, \hat{B}\right] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$. This brings out the importance of non-commuting operators to the dynamics of a quantum system, a feature which was not so readily apparent in the Schrödinger picture.

It also shows that any observable whose operator commutes with the Hamiltonian is a constant of the motion. Thus, the commutator is the direct counterpart of the Poisson bracket in classical Hamiltonian dynamics. Quantities which have zero Poisson bracket are constants of the motion in classical mechanics.

There are other possible 'pictures'. The picture habitually employed in the theory of particle scattering, e.g. quantum field theory, is the 'interaction picture'. This is a half-way house between the Schrödinger and Heisenberg pictures. The free-particle part of the Hamiltonian is treated in the Heisenberg picture, whilst the interaction part of the Hamiltonian is treated in the Schrödinger picture. The result is that the system state tracks only the time dependence due to the interaction, which is the focus of attention in such problems.

8. Whence Quantum Weirdness?

The origin of all quantum weirdness is superposition, i.e. the fact that the sum of two Hilbert states is another valid Hilbert state, together with the fact that such superpositions can be realised in practice. The 'wave' like nature of quantum systems follows from superposition. For example, 'wave' properties are generally manifest through interference. Interference experiments usually involve splitting a beam into two parts, e.g. by impingement on a double slitted screen or by passing through a half-silvered mirror. The two half-beams are then subject to different conditions, e.g. different path lengths or different magnetic fields. This results in different quantum states being subject to different phase factors. For example, suppose the incident beam was a superposition of two quantum states, $|1\rangle + |2\rangle$, and suppose that the differing conditions after the beam splitting results in state 1 acquiring a phase factor $e^{i\theta}$ and state 2 acquiring a phase factor

$-e^{i\theta}$. The final state after the two half beams have been recombined is the interference state $\frac{1}{2}\left(1+e^{i\theta}\right)|1\rangle+\frac{1}{2}\left(1-e^{i\theta}\right)|2\rangle$. Thus, if we imagine the phase angle θ as increasing as we look consider points further up a screen, the state will oscillate between $|1\rangle$ and $|2\rangle$, producing an interference pattern.

On the other hand, measurements always produce discreet outcomes. The individual electrons in an electron double-slit experiment always land at just one spot. Only when the positions of many electrons are recorded on the screen does the interference pattern emerge. Both particle and wave natures are manifest. The wave nature results from the superposition of Hilbert states. The particle nature results from the eigenvalue structure of measurements and the Born Rule (reduction of the wavepacket).

But the wave-particle duality is only the beginning of the quantum weirdness. Far more weirdness is yet to be revealed – but all of it is a consequence of the formulation already introduced above.