

Physics 195a
Course Notes
Ideas of Quantum Mechanics
021024 F. Porter

1 Introduction

This note summarizes and examines the foundations of quantum mechanics, including the mathematical background.

2 General Review of the Ideas of Quantum Mechanics

2.1 States

We have in mind that there is a “system”, which is describable in terms of possible “states”. A system could be something simple, such as a single electron, or complex, such as a table.

Suppose we have a system consisting of N spinless particles. We use the term “particle” to denote any object for which any internal structure is unimportant. Classically, we may describe the state of this system by specifying, at some time t the generalized coordinates and momenta:

$$\{q_i(t), p_i(t), i = 1, 2, \dots, N\}, \quad (1)$$

where the spatial dimensionality of the q_i and p_i is implicit. The time evolution of this system is given by Hamilton’s equations:

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad (2)$$

$$-\dot{p}_i = \frac{\partial H}{\partial q_i} \quad (3)$$

In quantum mechanics, it is not possible to give such a complete specification to arbitrary precision. For example, the limit to how well we may specify the position and momentum of a particle in one dimension is limited by the “uncertainty principle”: $\Delta x \Delta p \geq 1/2$, where Δ indicates a range of possible values. We’ll investigate this relation more explicitly later, but for

now it should just be a reminder of your elementary quantum mechanics understanding. We must be content with selecting a suitable set of quantities which can be simultaneously specified to describe the state. We refer to this set as a “Complete Set of Commuting Observables” (CSCO).

Specifying a CSCO corresponds to specifying the eigenvalues of an appropriate complete set of commuting Hermitian operators, for the state in question. Measurements (eigenvalues of Hermitian operators) of other quantities cannot be predicted with certainty, only probabilities of outcomes can be given. The evolution in time of the system is described by a “wave equation”, for example, the Schrödinger equation.

2.2 Probability Amplitudes

The quantum mechanical state of a system is described in terms of waves, called **probability amplitudes**, or just “amplitudes” for short. Note that probabilities themselves are always non-negative, so it is more difficult to imagine the probabilities themselves as wavelike. Instead, the probabilities are obtained by squaring the amplitudes:

$$\text{Probability} \sim |\psi|^2, \quad (4)$$

where ψ stands for the amplitude. More explicitly, the probability of observing state variable (*e.g.*, position) \mathbf{x} in volume element $d^3(\mathbf{x})$ around \mathbf{x} is equal to:

$$|\psi(\mathbf{x})|^2 d^3(\mathbf{x}). \quad (5)$$

A quantum mechanical probability is analogous to the intensity of a classical wave.

The quantum mechanical wave evolves in time according to a time evolution operator, e^{-iHt} involving the Hamiltonian, H . Hence, if

$$e^{-iHt}\psi_0(\mathbf{x}) = \psi(\mathbf{x}, t), \quad (6)$$

where $\psi_0(\mathbf{x})$ is the wave function at $t = 0$ in terms of coordinate position \mathbf{x} , then differentiation gives:

$$i\frac{d\psi(\mathbf{x}, t)}{dt} = H\psi(\mathbf{x}, t). \quad (7)$$

We recognize this as the Schrödinger equation. Thus, the temporal frequency of the wave is determined by the energy structure. For a particle of energy

E , the frequency is $\omega = E$ (or $\nu = E/2\pi$). This hypothesis is also applied in relativistic situations, for example, for a photon.

The spatial behavior of a wave is given by the **deBroglie hypothesis**: A particle is described as a quantum mechanical wave with wavelength:

$$\lambda \equiv \frac{\lambda}{2\pi} = \frac{1}{p}, \quad (8)$$

or with wavenumber $k = 1/\lambda = p$. This relation is assumed to also hold relativistically.

We may make a brief aside on the subject of “dispersion relations”. As in classical electrodynamics, the relation between ω and k for a wave is called a dispersion relation. In the case of the quantum mechanics of a free particle of mass m , the dispersion relation is

$$\omega = k^2/2m \quad (9)$$

for a non-relativistic particle (when we do not include the rest mass in the energy, hence $E = p^2/2m$), and

$$\omega^2 = k^2 + m^2 \quad (10)$$

for a relativistic particle.

2.3 Wave Equations

In quantum mechanics, the dynamics is determined by the wave equation. The form of the wave equation is given by the dispersion relation. By analogy with light, and ignoring issues of mathematical rigor, let us build physical waves describing a particle of mass m from superpositions of plane waves. Note well that we are assuming that the wave equation is linear, so linear combinations of solutions are also solutions. Our plane wave building blocks are:

$$\psi(\mathbf{x}, t) = Ae^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}, \quad (11)$$

where $\mathbf{k}=\mathbf{p}$ and $\omega = E$. We presume that this forms a “complete set” of functions, that is, any physical state may be expanded as a linear superposition of elements of this set. Let us suppose we have a free particle. We search for a differential equation which is satisfied by all plane waves which could

describe the particle. We shall postulate this to be our wave equation. As we saw above, the dispersion equation for a (possibly relativistic) free particle is

$$E^2 - p^2 = m^2, \quad (12)$$

or

$$\omega^2 - k^2 = m^2. \quad (13)$$

Considering the following derivatives of the plane waves

$$\frac{\partial^2 \psi(\mathbf{x}, t)}{\partial t^2} = -\omega^2 \psi(\mathbf{x}, t), \quad (14)$$

$$\nabla^2 \psi(\mathbf{x}, t) = -k^2 \psi(\mathbf{x}, t), \quad (15)$$

we have

$$\frac{\partial^2 \psi}{\partial t^2} - \nabla^2 \psi = -m^2 \psi. \quad (16)$$

We postulate this to be the desired wave equation. It is known as the **Klein-Gordon equation**. It describes the motion, in quantum mechanics, of a free particle of mass m .

However, free particles quickly become boring; we really want to be able to discuss interactions, *e.g.*, interacting particles. Demanding relativistic invariance leads us into quantum field theory. However, we often don't require full invariance, and typically make two very useful simplifying assumptions in non-relativistic quantum mechanics:

- The creation and destruction of particles is assumed not to occur. The number of each particle type is constant. However, there is occasional need to make exceptions to this assumption; the most notable is for the photon.
- All particles (again, except for photons!) are assumed to be non-relativistic. Typically this means we stop at order v^2 in the energy, but sometimes we carry out calculations to higher order.

We have already seen that these assumptions are reasonable for ordinary atomic systems.

In non-relativistic quantum mechanics (Schrödinger theory) the wave function $\psi(\mathbf{x}, t)$ really has the precise meaning:

The probability that a measurement of the particle's position at time t will yield \mathbf{x} in $d^3(\mathbf{x})$ around \mathbf{x} is:

$$|\psi(\mathbf{x}, t)|^2 d^3(\mathbf{x}). \quad (17)$$

$|\psi(\mathbf{x}, t)|^2$ is a **probability density**.

Note that $\psi(\mathbf{x}, t)$ and $\psi'(\mathbf{x}, t) = e^{i\theta}\psi(\mathbf{x}, t)$, where θ is a real number, describe the same physical situation, since the probability density is unchanged [where we make the inherent assumption that it is probabilities we can measure, not probability amplitudes].

Let us take the non-relativistic limit of our free particle wave equation:

$$\begin{aligned} E &= \sqrt{m^2 + p^2} \\ &= m + \frac{p^2}{2m} + p\mathcal{O}\left[\left(\frac{p}{m}\right)^3\right]. \end{aligned} \quad (18)$$

Hence,

$$\begin{aligned} \psi(\mathbf{x}, t) &= Ae^{i(\mathbf{p}\cdot\mathbf{x}-Et)} \\ &\approx Ae^{i(\mathbf{p}\cdot\mathbf{x}-\frac{p^2}{2m}t)}e^{-imt} \\ &\approx \psi_S(\mathbf{x}, t)e^{-imt}, \end{aligned} \quad (19)$$

where

$$\psi_S(\mathbf{x}, t) \equiv e^{i(\mathbf{p}\cdot\mathbf{x}-\frac{p^2}{2m}t)}. \quad (20)$$

But $|\psi(\mathbf{x}, t)|^2 = |\psi_S(\mathbf{x}, t)|^2$, so we may drop the overall e^{-imt} phase factor and look for a linear differential equation satisfied by our non-relativistic plane wave solutions (dropping the S subscript now). We have

$$\frac{\partial\psi}{\partial t} = -i\frac{p^2}{2m}\psi, \quad (21)$$

$$\nabla^2\psi = -p^2\psi. \quad (22)$$

Thus,

$$i\frac{\partial\psi}{\partial t} = -\frac{1}{2m}\nabla^2\psi. \quad (23)$$

This is the **Schrödinger equation** for a free particle of mass m . Note the correspondence with the dispersion relation, Eq. 9.

In the non-relativistic case, it is easy to generalize to situations where the particle is not free: Introduce a potential function, $V(\mathbf{x}, t)$ to describe interactions. Our hypothesis is that the time dependence of the wave is determined by $\omega = E$. With $V \neq 0$, this gives

$$E = T + V = \frac{p^2}{2m} + V(\mathbf{x}, t). \quad (24)$$

Thus,

$$\begin{aligned} i\frac{\partial\psi}{\partial t} &= i\frac{\partial}{\partial t} \exp[i(\mathbf{p} \cdot \mathbf{x} - Et)] \\ &= E\psi \\ &= H\psi, \end{aligned} \quad (25)$$

where $H = T + V$ is the Hamiltonian operator. We may sometimes need to make the distinction between an operator and its spectral values (eigenvalues) more explicit. When this arises, we will use notation of the form \hat{H} or H_{op} to denote the operator. However, we usually rely on context, and omit such notational guides.

We also have that

$$-\frac{1}{2m}\nabla^2\psi(\mathbf{x}, t) = \frac{p^2}{2m}\psi(\mathbf{x}, t) = (E - V)\psi(\mathbf{x}, t). \quad (26)$$

Putting this together with Eq. 25, we find:

$$H\psi(\mathbf{x}, t) = \left\{ \begin{array}{l} E\psi(\mathbf{x}, t) \\ i\partial_t\psi(\mathbf{x}, t) \end{array} \right\} = -\frac{1}{2m}\nabla^2\psi(\mathbf{x}, t) + V(\mathbf{x}, t)\psi(\mathbf{x}, t). \quad (27)$$

The upper form, in which E is the energy eigenvalue for a static potential, is referred to as the **time-independent** Schrödinger equation. The lower form is referred to as the **time-dependent** Schrödinger equation.

3 Mathematical Considerations

Let us take a step back now, and set up a more rigorous mathematical framework in which to implement the notions we have been discussing. It is a highly reassuring feature of quantum mechanics that we are able to do so.

We have decided to describe particles by “waves”, giving “probability amplitudes”, where absolute squares lead to measurable physical probabilities.

Waves are conveniently described by complex-valued functions of whatever generalized coordinates are involved. An essential feature is that waves interfere, hence our state space must allow for the possibility of superposition of waves.

- The requirement of superposability suggests that the state space of permissible wave functions should be a vector space. This gives us the property that linear combinations of physical amplitudes lead to new physically allowed amplitudes.
- To deal with the probability interpretation, we briefly consider the definition of probability:

Def: Probability: If S is a (sample) space, and $P(E)$ is a real additive set function defined on sets E in S , then P is referred to as a **probability function** if:

1. If E is a subset (**event**) in S , then $P(E) \geq 0$.
2. $P(S) = 1$.
3. If $E, F \subseteq S$, and $E \cap F = \emptyset$, then $P(E \cup F) = P(E) + P(F)$.
4. If S is an infinite sample space, we require that:

$$P(E_1 \cup E_2 \cup \dots) = P(E_1) + P(E_2) + \dots \quad (28)$$

for any sequence of disjoint events E_1, E_2, \dots in S .

[For those with the mathematical background, we remark that a shorter definition for probability is: A probability function is a measure on S such that $P(S) = 1$. We note that P is defined on all subsets E of S , hence is defined on a σ -ring.]

Thus, the requirement of a probability interpretation means that any allowable wave function $\psi(s)$ defined on sample space S must be normalizable and square-integrable such that:

$$\int_S \psi^*(s)\psi(s)\mu(ds) = 1. \quad (29)$$

The integral here is in the Lebesgue-summable sense, and μ is the appropriate measure function on subsets of S . A “measure function” is simply a prescription for measuring the “sizes” of sets, implemented in a mathematically rigorous manner.

The mathematical considerations here are both critical to the foundation of quantum mechanics and potentially unfamiliar to the reader, so we will digress briefly in order to develop an intuitive understanding of the need for them.

Apparently, it is important to know how to measure the sizes of sets in our probability sample space. This is implemented abstractly in measure theory. We will not develop this here; a couple of examples should provide the intuition that is sufficient for present purposes. For a first example, suppose our sample space is the set of real numbers, R^1 . In this case, the appropriate way to measure sizes of sets is a suitable generalization of our ordinary notion that the size of the interval (a, b) is just $b - a$. This generalization is called the **Lebesgue measure** on R^1 . It has the property that a denumerable set of discrete points is measurable, with measure zero.

We remark that the Riemann integral is not sufficiently general for our purposes. A function $f(x)$ is Riemann-integrable on $[a, b]$ if and only if:

- $f(x)$ is bounded.
- The set of points of discontinuity of f has Lebesgue measure zero.

For example, the integral

$$\int_0^1 f(x) dx, \tag{30}$$

where

$$f(x) = \begin{cases} 0 & \text{if } x \text{ is rational,} \\ 1 & \text{if } x \text{ is irrational,} \end{cases} \tag{31}$$

is not defined. The function is discontinuous at every point, hence the measure of the points of discontinuity is non-zero:

$$\mu(\{\text{points of discontinuity}\}) = \mu(\{(0, 1)\}) = 1. \tag{32}$$

This is perhaps a pathological example. A more obvious example is that the Riemann sum doesn't allow us to sum over state variables with possibly discrete spectra, *e.g.*, quantized energy levels. We could handle such situations in an ad hoc manner, but to build a rigorous foundation we resort to the **Lebesgue integral**.

The idea of the Lebesgue integral is simple and elegant. Rather than divide the “ x -axis” up into intervals, as in the Riemann integral, we divide the “ y -axis”. That is, we partition the y -axis into intervals Δy_i , $i = 1, 2, \dots$

Choose a point y_i in each interval. Consider the sets $f^{-1}(\Delta y_i)$. Multiply the measure of each such set by the corresponding y_i . Then sum the products. The Lebesgue integral is the value of this sum in the limit where all of the Δy_i intervals vanish.

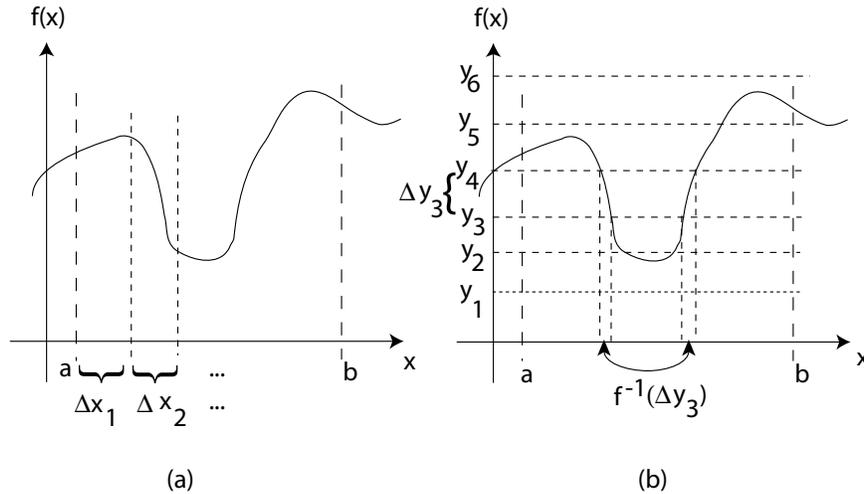


Figure 1: (a) The Riemann integral is a limit of slices in x . (b) The Lebesgue integral is a limit of slices in y .

$$I_L \equiv \lim_{\Delta y_i \rightarrow 0} \sum_i y_i \mu[f^{-1}(\Delta y_i)]. \quad (33)$$

For this to work, $f^{-1}(\Delta y_i)$ must be measurable sets, that is, $f(x)$ must be a “measurable function”:

Def: A real function $f(x)$, defined on S is said to be **measurable** if, for every real number u , the set $S_u = \{x : f(x) < u, x \in S\}$ is measurable.

For example, consider the function of Eqn. 31. Take, in the limit, $y_1 = 0$ and $y_2 = 1$. Then

$$\begin{aligned} f^{-1}(\Delta y_1) &= \text{rational numbers} \\ f^{-1}(\Delta y_2) &= \text{irrational numbers}, \end{aligned} \quad (34)$$

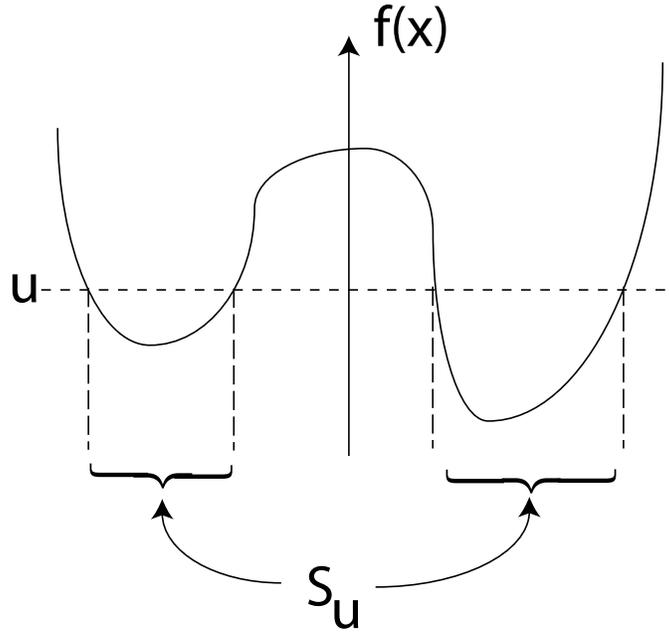


Figure 2: The set S_u .

and

$$\begin{aligned}
 1 = \mu([0, 1]) &= \mu(\{\text{rationals on } [0, 1]\}) + \mu(\{\text{irrationals on } [0, 1]\}) \\
 &= 0 + \mu(\{\text{irrationals on } [0, 1]\}).
 \end{aligned}
 \tag{35}$$

Hence,

$$\int_0^1 f(x) \mu(dx) = 1
 \tag{36}$$

is the Lebesgue integral.

The choice of measure (of the size of a set) may depend on the physical circumstance. We have used the Lebesgue measure on R^1 , appropriate to continuous state variables. Another important measure is the Dirac measure (on $S = R^1$): Let $x_0 \in R^1$. The Dirac measure associated with point x_0 is defined by:

$$\mu(E) = \begin{cases} 1 & \text{if } x_0 \in E \\ 0 & \text{if } x_0 \notin E. \end{cases}
 \tag{37}$$

Note that this is the appropriate measure to use for “discrete” state variables:

$$\int_S f(x) \mu(dx) = f(x_0). \quad (38)$$

We are still trying to build a suitable function space for our quantum mechanical wave functions. What about “pathological” functions, *e.g.*, with many discontinuities. We will build into our space the concept that two functions that differ only in ways which will not affect observable probabilities are not to be considered as distinct. We proceed as follows:

Def: A property, $Q(x)$, which depends on location x in space S , is said to hold **almost everywhere** if the set of points for which Q does not hold has measure zero.

Def: Two functions $f_1(x), f_2(x)$, defined on S (that is, assume finite values at every point of S) are said to be **equivalent** if $f_1(x) = f_2(x)$ almost everywhere: $f_1 \sim f_2$.

If equivalent functions $f_1(x)$ and $f_2(x)$ are integrable in the Lebesgue sense (“summable”) on a set E , then

$$\int_E f_1(x) \mu(dx) = \int_E f_2(x) \mu(dx). \quad (39)$$

Thus, if we decompose the set of summable functions into classes of equivalent functions, the integral can be regarded as a functional defined on the space F , of these classes.

3.1 The Space L^2

For our quantum mechanical wave functions we are of course interested in complex functions. A complex function $f(x) = f_1(x) + if_2(x)$, where f_1 and f_2 are real functions, is said to be summable on E if f_1 and f_2 are summable:

$$\int_E f(x) \mu(dx) = \int_E f_1(x) \mu(dx) + i \int_E f_2(x) \mu(dx). \quad (40)$$

Theorem: A complex function $f(x)$ is summable if and only if its absolute value,

$$|f(x)| = \sqrt{f_1(x)^2 + f_2(x)^2}, \quad (41)$$

is summable.

Proof: Suppose $f(x) = f_1(x) + if_2(x)$ is summable on E . Then f_1 and f_2 are summable on E . By virtue of our definition of the integral, Eqn. 33, $|f_1|$ and $|f_2|$ are therefore also summable. Hence,

$$\begin{aligned} \int_E |f(x)|\mu(dx) &= \int_E |f_1(x) + if_2(x)|\mu(dx) \\ &\leq \int_E |f_1(x)| + |f_2(x)|\mu(dx), \text{ by the triangle inequality} \\ &< \infty. \end{aligned} \tag{42}$$

Conversely, suppose $|f|$ is summable on E . Then, again referring to the definition in Eqn. 33,

$$\begin{aligned} \left| \int_E f(x)\mu(dx) \right| &\leq \int_E |f(x)|\mu(dx) \\ &< \infty. \end{aligned} \tag{43}$$

Even for complex functions, the integral defines a linear functional. Let L denote the space of complex functions $f(x)$ such that $|f(x)|^2$ is summable on S :

$$\int_S |f(x)|^2\mu(dx) < \infty, \quad \text{for } f(x) \in L. \tag{44}$$

Theorem: The space L is a linear space (or “vector” space).

Proof: The principal step of the proof is as follows: Suppose $f(x) \in L$ and $g(x) \in L$. Then

$$\begin{aligned} |f(x) + g(x)|^2 &= 2|f(x)|^2 + 2|g(x)|^2 - |f(x) - g(x)|^2 \\ &\leq 2|f(x)|^2 + 2|g(x)|^2. \end{aligned} \tag{45}$$

But $2|f(x)|^2 + 2|g(x)|^2$ is summable, and hence $|f(x) + g(x)|^2$ is summable.

The space L is our candidate space for physical quantum mechanical wave functions. However, there is a problem with it: There are distinct elements of L , differing on sets of measure zero, which correspond to the same physics. Let us tidy this ugliness up. Consider Z the subset of L consisting of functions $f(x)$ such that

$$\int_S |f(x)|^2\mu(dx) = 0. \tag{46}$$

Note that Z is a linear subspace of L , since if $f \in Z$, then $kf \in Z$ for all complex constants k , and if $f, g \in Z$, then $f + g \in Z$ since $|f + g|^2 \leq 2|f|^2 + 2|g|^2$. Thus, we can define the “factor space”,

$$L^2 = L/Z. \quad (47)$$

That is, two functions $f_a(x), f_b(x)$ in L determine the same class in L^2 if and only if the difference $f_a - f_b$ vanishes almost everywhere, *i.e.*,

$$\int_S |f_a(x) - f_b(x)|^2 \mu(dx) = 0. \quad (48)$$

We say that the space L^2 consists of functions $f(x)$ such that $|f(x)|^2$ is summable on S , with the understanding that equivalent functions are not considered distinct. In other words, L^2 is a space of equivalence classes.

Finally, we add to this space the notion of a scalar product. We start by noting that the product of two elements of L^2 is summable:

Theorem: If $f, g \in L^2$, then f^*g is summable on S .

Proof: Write

$$f^*g = \frac{1}{4} (|f + g|^2 - |f - g|^2 + i|f - ig|^2 - i|f + ig|^2) \quad (49)$$

Each term on the right is summable, and hence the product $f^*g \in L^2$.

Theorem: L^2 is a Hilbert space, with scalar product defined by:

$$\langle f|g \rangle \equiv \int_S f(x)^* g(x) \mu(dx). \quad (50)$$

Proof: The proof starts by showing that L^2 is a pre-Hilbert space, that is, a linear space upon which a scalar product has been properly defined. This consists in showing that:

1. $\langle f|f \rangle = 0$ if and only if $f = 0$. The fact that L^2 is a space of equivalence classes is crucial here.
2. $\langle f|g \rangle = \langle g|f \rangle^*$.
3. $\langle f|cg \rangle = c\langle f|g \rangle$.
4. $\langle f|g_1 + g_2 \rangle = \langle f|g_1 \rangle + \langle f|g_2 \rangle$.

Once it has been demonstrated that L^2 is a pre-Hilbert space, it remains to show that L^2 is complete. That is, it must be shown that every Cauchy sequence of vectors in L^2 converges to a vector in L^2 .

A fundamental postulate of quantum mechanics is:

**To every physical system S there corresponds
a separable Hilbert space, H_S .**

- L^2 appears to be a suitable space for our probability amplitudes since it is a linear space (hence we have superposition), and its elements are normalizable (square-summable, hence can make a probability interpretation).
- The addition of the scalar product permits us to make projections in our vector space. Note tht L itself was not sufficient for this construction, since $\langle f|f \rangle = 0$ is not equivalent to $f = 0$ in L . It should be understood that using L^2 is all right, since functions which differ only on sets of measure zero will yield the same probabilistic, and hence physical, results.
- The availability of the scalar product in particular leads to the possibility of constructing a (orthonormal) “basis”.
- Completeness means that we have included a sufficiently large set of vectors that we won't encounter difficulties when we consider certain sequences of vectors. We can construct a complete orthonormal basis $\{|e_\alpha\rangle\}$ on a Hilbert space such that every vector $|x\rangle \in H$ can be expanded:

$$|x\rangle = \sum_{\alpha} |e_\alpha\rangle \langle e_\alpha|x\rangle. \quad (51)$$

- Abstractly, a separable space is a topological space T which contains a denumerable (countable) set of points $\{t_1, t_2, \dots\}$ which is dense in T . The point of the postulate that the Hilbert space corresponding to any physical system be separable is that there is then a denumerable dense set of vectors. We may find a complete denumerable basis in which to expand our vectors.

To complete the connection of this postulate with our space L^2 we have the theorem:

Theorem: The space $L^2(a, b)$ (where it is permissible for $a = -\infty$, $b = \infty$) with Lebesgue measure is separable.

Proof: To prove this theorem, we first prove that there exists a complete denumerable orthonormal basis in L^2 . For example, on $L^2(0, 2\pi)$, the set of functions:

$$\frac{e^{ikx}}{\sqrt{2\pi}}, \quad k = 0, \pm 1, \pm 2, \dots, \quad (52)$$

forms a complete orthonormal system. Then we show that from this basis we can construct a countable dense set of vectors in L^2 .

It may be noted that non-separable Hilbert spaces do exist. However, we have so far not found a need to consider them for quantum mechanics.

On $L^2(-\infty, \infty)$, with measure $\mu(dx) = e^{-x^2} dx$, the Hermite polynomials:

$$H_n(x) = \frac{e^{x^2}}{\sqrt{2^n n! \sqrt{\pi}}} \frac{d^n}{dx^n} e^{-x^2}, \quad n = 0, 1, \dots \quad (53)$$

form a complete orthonormal system. Alternatively, with measure $\mu(dx) = dx$, the functions $e^{-x^2}/2H_n(x)$ form a complete orthonormal system.

4 Observables

An observable Q is a physical quantity. In quantum mechanics, we deal with the probability $p(Q, \Delta)$ that a measurement will yield a value of Q in a subset Δ of the set of real numbers. A fundamental postulate of quantum mechanics is that:

Every observable corresponds to a self-adjoint operator defined in H_S .

The term “defined in H_S ” means, for operator Q , that $x \in D_Q \subset H_S$, and $Qx \in R_Q \subset H_S$, where D_Q is the domain of the operator, and R_Q is its range.

Self-adjoint operators are evidently an important class of operator – the key point is that a self-adjoint operator is also a Hermitian operator, and hence has a real eigenvalue spectrum. This is the physical reason why they are of interest. Let us look at some of the mathematical aspects.

Def: (Adjoint) Let L be a linear operator, defined in H_S with domain D_L , such that D_L is dense in H_S (that is, $\bar{D}_L = H_S$, where \bar{D}_L is the closure of D_L). The **adjoint**, L^\dagger , of L is defined by:¹

$$\langle L^\dagger u | v \rangle = \langle u | Lv \rangle, \quad \forall v \in D_L. \quad (54)$$

In other words, u is a vector in H_S such that there exists a $w \in H_S$ satisfying $\langle u | Lv \rangle = \langle w | v \rangle$. If this holds, then we say $w = L^\dagger u$; the adjoint operator maps u to w . The requirement that D_L be dense in H_S is necessary in order for L^\dagger to be uniquely defined. To see this, suppose that it is not unique, *i.e.*, suppose there exist two vectors w_a, w_b such that

$$\langle u | Lv \rangle = \langle w_a | v \rangle = \langle w_b | v \rangle, \quad \forall v \in D_L. \quad (55)$$

In this case, $\langle (w_a - w_b) | v \rangle = 0$. But $w_a - w_b$ is thus orthogonal to every vector in a dense set, and therefore $w_a - w_b = 0$. This last point could use some further proof; we'll depend on its evident plausibility here.

Def: Self-adjoint: If $L^\dagger = L$ (which means: $D_{L^\dagger} = D_L$, and $L^\dagger u = Lu$ for all $u \in D_L$), then L is said to be **self-adjoint**.

Note the distinction between a self-adjoint operator, and a Hermitian operator, defined according to:

Def: Hermitian: A linear operator L , with $D_L \subset H_S$, is called **Hermitian** if

$$\langle Lu | v \rangle = \langle u | Lv \rangle, \quad \forall u, v \in D_L. \quad (56)$$

For example, in the case of a finite dimensional vector space, L is a square matrix, and we have:

$$\langle u | Lv \rangle = u^\dagger Lv \quad (57)$$

$$\langle Lu | v \rangle = (Lu)^\dagger v \quad (58)$$

$$\begin{aligned} &= u^\dagger L^\dagger v \\ &= u^\dagger Lv \quad \text{if } L^\dagger = L. \end{aligned} \quad (59)$$

¹There are a variety of notations used to denote the adjoint of an operator, most notably L^\dagger , L^+ , and L^* . We'll adopt the "dagger" notation here, as it is consistent with the familiar "complex-conjugate-transpose" notation for matrices. The "asterisk" notation is common also, but we avoid it here on the grounds of potential confusion with simple complex conjugation.

In this case, a self-adjoint operator is also a Hermitian operator:

$$\langle L^\dagger u|v\rangle = \langle Lu|v\rangle = \langle u|Lv\rangle, \quad \forall u, v \in D_L = D_{L^\dagger}. \quad (60)$$

However, a Hermitian operator is not necessarily a self-adjoint operator, if the space is infinite dimensional. The issue is one of domain. It can happen, in an infinite dimensional Hilbert space, that a Hermitian operator, L , has $D_L \subset D_{L^\dagger}$, as a proper subset. In this case, L is not self-adjoint.

Consider an example to illustrate this inequivalence. A differential equation (where L is a differential operator, and we write $Lu = a$) is not completely specified until we give certain boundary conditions which the solution must satisfy. Thus, for a function u to belong to D_L , not only must the expression Lu be defined, but u must also satisfy the boundary conditions. If the boundary conditions are too restrictive, we might have $D_L \subset D_{L^\dagger}$ but $D_L \neq D_{L^\dagger}$, so that a Hermitian operator may not be self-adjoint.

To illustrate with a specific example, let $L = p$ be the momentum operator in one dimension:

$$p = \frac{1}{i} \frac{d}{dx}, \quad x \in [a, b]. \quad (61)$$

The boundary conditions are to be specified, but the domain of this operator is otherwise the set of continuous functions on $[a, b]$. This set of functions is dense in our Hilbert space $L^2(a, b)$. We look at the scalar product of pv with u , where u and v are continuous functions:

$$\langle u|pv\rangle = \int_a^b u^*(x) \frac{1}{i} \frac{d}{dx} v(x) dx \quad (62)$$

$$\begin{aligned} &= \frac{1}{i} u^*(x)v(x) \Big|_a^b - \int_a^b \left[\frac{1}{i} \frac{d}{dx} u^*(x) \right] v(x) dx \\ &= \frac{1}{i} [u^*(b)v(b) - u^*(a)v(a)] + \langle pu|v\rangle \end{aligned} \quad (63)$$

$$= \langle p^\dagger u|v\rangle \quad (64)$$

The $u^*(b)v(b) - u^*(a)v(a)$ boundary term portion of Eqn. 63 must vanish for all $v \in D_p$ in order for p to be a Hermitian operator; hence we shall assume this condition. There is, however, more than one way to achieve this, even with the dense requirement. For example, we could impose the boundary condition $v(a) = v(b) = 0$, so that $D_p = \{v|v \text{ is continuous, and } v(a) = v(b) = 0\}$. In this case, u need not satisfy any constraints at a or b , and

$D_{p^\dagger} = \{u|u \text{ is continuous}\} \neq D_p$, since

$$\langle p^\dagger u|v \rangle = \langle u|pv \rangle = \left\langle \frac{1}{i} \frac{d}{dx} u|v \right\rangle, \quad (65)$$

for all continuous functions u . We have $p^\dagger = \frac{1}{i} \frac{d}{dx}$ with $D_{p^\dagger} = \{u|u \text{ continuous on } [a, b]\}$. So, p is Hermitian, but not self-adjoint, since D_p is a proper subset of D_{p^\dagger} .

On the other hand, if we had chosen the extension of the above p with boundary condition $v(a) = v(b)$, then we would find a restriction of the above p^\dagger , with $D_{p^\dagger} = \{u|u(a) = u(b), u \text{ continuous on } [a, b]\}$. With this definition p is a self-adjoint operator.

5 The Uncertainty Principle

The famous ‘‘uncertainty principle’’ is discussed in every introductory quantum mechanics course. We revisit it briefly here. First, the reader is reminded of the important **Schwarz inequality**:

Theorem: For any vectors ϕ, ψ in our Hilbert space,

$$|\langle \phi|\psi \rangle| \leq \sqrt{\langle \phi|\phi \rangle \langle \psi|\psi \rangle}. \quad (66)$$

Equality holds if and only if ϕ and ψ are linearly dependent: $\phi = c\psi$, where c is a complex number.

Proof: One way to prove the Schwarz inequality is to consider the non-negative definite scalar product:

$$\langle \phi + re^{i\theta}\psi | \phi + re^{i\theta}\psi \rangle \geq 0. \quad (67)$$

Expanding the left hand side results a quadratic expression for r . Considering the possible solutions for r yields a constraint on the discriminant. The resulting inequality is the Schwarz inequality.

Suppose now that we have two self-adjoint operators A and B , and a state vector in the domains of both. The average (mean) value (expectation value) of observable A if the system is in state ψ is:

$$\langle A \rangle = \langle \psi | A \psi \rangle, \quad (68)$$

where we assume that ψ is normalized. Likewise, the mean of observable B is $\langle B \rangle = \langle \psi | B \psi \rangle$. We are presently interested in learning something about the spreads of the distributions of observations of A and B . Thus, it is convenient to subtract out the means by defining “shifted” operators:

$$A_S \equiv A - \langle A \rangle \quad (69)$$

$$B_S \equiv B - \langle B \rangle. \quad (70)$$

The domains of the shifted operators are the same as the domains of the unshifted operators. We immediately have that $\langle A_S \rangle = \langle B_S \rangle = 0$.

Define the commutator of A_S and B_S :

$$[A_S, B_S] \equiv A_S B_S - B_S A_S = [A, B]. \quad (71)$$

It should be noted that the product of two operators is defined by their operation on a state vector: $AB|\psi\rangle$ means first apply operator B to ψ , then apply A to the result. The obvious questions of domain need to be dealt with, of course. Thus, let us further require $\psi \in D_{AB}, \psi \in D_{BA}$, and consider:

$$\begin{aligned} |\langle [A, B] \rangle| &= |\langle \psi | A_S B_S \psi \rangle - \langle \psi | B_S A_S \psi \rangle| & (72) \\ &\leq |\langle \psi | A_S B_S \psi \rangle| + |\langle \psi | B_S A_S \psi \rangle| & \text{(triangle inequality)} \\ &\leq |\langle A_S \psi | B_S \psi \rangle| + |\langle B_S \psi | A_S \psi \rangle| & \text{(self-adjointness)} \\ &\leq 2|\langle A_S \psi | B_S \psi \rangle| \\ &\leq 2\sqrt{\langle A_S \psi | A_S \psi \rangle \langle B_S \psi | B_S \psi \rangle} & \text{(Schwarz inequality)} \\ &\leq 2\sqrt{\langle \psi | A_S^2 \psi \rangle \langle \psi | B_S^2 \psi \rangle} & \text{(self-adjointness)}. \end{aligned} \quad (73)$$

The **variance** of a distribution is a measure of its spread. For an observable Q , the variance for a system in state ψ is defined by:

$$\sigma_Q^2 \equiv \langle \psi | (Q - \langle Q \rangle)^2 \psi \rangle = \langle Q^2 \rangle - \langle Q \rangle^2. \quad (74)$$

The square root of the variance, σ_Q is called the **standard deviation**. We see that, for example, $\sigma_A^2 = \langle \psi | A_S^2 \psi \rangle$. Thus, we may rewrite Eqn. 73 in the form:

$$\sigma_A \sigma_B \geq \frac{1}{2} \sqrt{|\langle [A, B] \rangle|}. \quad (75)$$

This is a precise statement of the celebrated “uncertainty principle”. We shall often use the convenient notation $\Delta A \equiv \sigma_A$. The physical interpretation is

that, if we have two non-commuting observables, the product of the variances of the probability distributions for these two observables is bounded below. This is typically interpreted further with statements such as “the ability to measure both variables simultaneously is limited. A measurement of one observable disturbs the system in a way that affects the result of a second measurement of the other observable.” While there is some justification for such statements, one must be careful not to carry them too far – in case of confusion, come back to what the principle actually says! For example, “the ability to measure” carries a connotation that there may be an issue of experimental resolution involved. While experimental resolution generally needs to be folded into the analysis of an actual experiment, it has nothing to do with the present point.

5.1 Example: Angular Momentum

The angular momentum operator for a particle is $\mathbf{L} = \mathbf{x} \times \mathbf{p}$, where \mathbf{x} is the position operator, and \mathbf{p} is the momentum operator. This may be expressed in components as:

$$L_i = \epsilon_{ijk} x_j p_k. \quad (76)$$

The summation convention is used here: a sum is implied over repeated indices, in this case, there is a sum over $j, k = 1, 2, 3$. The quantity ϵ_{ijk} is known as the **antisymmetric symbol**:

$$\epsilon_{ijk} = \begin{cases} +1 & i, j, k = \text{cyclic permutation of } 1, 2, 3, \\ -1 & i, j, k = \text{anti-cyclic permutation of } 1, 2, 3, \\ 0 & \text{any two indices the same.} \end{cases} \quad (77)$$

We remark that $\mathbf{L} = \mathbf{x} \times \mathbf{p} = -\mathbf{p} \times \mathbf{x}$ are both acceptable, since only commuting components of \mathbf{x} and \mathbf{p} are paired.

We know that

$$[p_m, x_n] = -i\delta_{mn}. \quad (78)$$

We are interested in the commutation relations of the angular momentum operators:

$$[L_\alpha, L_\beta] = \epsilon_{\alpha j k} \epsilon_{\beta m n} [x_j p_k, x_m p_n] \quad (79)$$

$$= i(\epsilon_{\alpha j m} \epsilon_{\beta j n} - \epsilon_{\alpha j n} \epsilon_{\beta j m}) x_m p_n, \quad (80)$$

where the algebra between Eqns. 79 and 80 is left as an exercise for the reader. The reader is also encouraged to demonstrate that

$$E_{\alpha\beta, mn} \equiv \epsilon_{\alpha jm} \epsilon_{\beta jn} - \epsilon_{\alpha jn} \epsilon_{\beta jm} = \epsilon_{\alpha j\beta} \epsilon_{mjn}. \quad (81)$$

With the identity of Eqn. 81, we obtain

$$[L_\alpha, L_\beta] = i\epsilon_{\alpha\beta\gamma} L_\gamma. \quad (82)$$

Thus, the uncertainty relation between components of angular momentum is:

$$\Delta L_\alpha \Delta L_\beta \geq \frac{1}{2} |\langle [L_\alpha, L_\beta] \rangle| = \frac{1}{2} |\epsilon_{\alpha\beta\gamma} \langle L_\gamma \rangle|. \quad (83)$$

Let us illustrate this with an explicit example. We first anticipate the generalization of angular momentum to include spin, with the same commutation relations, and consider the simplest system with non-zero angular momentum, spin-1/2. We'll follow common convention, and pick our basis to be eigenvectors of J_3 (using now J to indicate angular momentum, leaving L to stand for “orbital” angular momentum). We again anticipate the quantization of spin, where the eigenvalues of J_3 are $\pm 1/2$ for a spin-1/2 system. In this basis, our angular momentum operator is:

$$\mathbf{J} = \frac{1}{2} \boldsymbol{\sigma}, \quad (84)$$

where $\boldsymbol{\sigma}$ are the **Pauli matrices**:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (85)$$

These are Hermitian matrices, hence correspond to observables.

Suppose, in this basis, we have the state

$$\psi = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad (86)$$

which is a superposition of $J_3 = \pm 1/2$ eigenstates. We may compute expectation values of angular momentum for this state:

$$\langle J_1 \rangle = \frac{1}{4} (1, 1) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{2} \quad (87)$$

$$\langle J_2 \rangle = \frac{1}{4} (1, 1) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0 \quad (88)$$

$$\langle J_3 \rangle = \frac{1}{4} (1, 1) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0. \quad (89)$$

To obtain the second moments, we notice that $\sigma_i^2 = 1$, $i = 1, 2, 3$. Thus,

$$\langle J_i^2 \rangle = \frac{1}{4}, \quad i = 1, 2, 3, \quad (90)$$

hence

$$(\Delta J_1)^2 = \langle J_1^2 \rangle - \langle J_1 \rangle^2 = \frac{1}{4} - \left(\frac{1}{2}\right)^2 = 0, \quad (91)$$

$$(\Delta J_2)^2 = \frac{1}{4} - 0 = \frac{1}{4}, \quad (92)$$

$$(\Delta J_3)^2 = \frac{1}{4} - 0 = \frac{1}{4}. \quad (93)$$

Let us check the uncertainty relation involving J_1 and J_2 :

$$\Delta J_1 \Delta J_2 = 0 \cdot \frac{1}{2} = 0 \geq \frac{1}{2} |\langle J_3 \rangle| = 0. \quad (94)$$

So this relation is satisfied. Physically, it may readily be seen that our state is actually an eigenstate of J_1 with eigenvalue $1/2$. It is a superposition of $J_2 = \pm 1/2$ eigenstates. Even though our lower bound on the product of uncertainties is zero, and is achieved, we cannot measure J_1 and J_2 simultaneously with arbitrary precision. As soon as we know $J_1 = 1/2$, a measurement of J_2 will yield $\pm 1/2$ with equal probability. Alternatively, if we first measure J_2 , obtaining a value of either $1/2$ or $-1/2$, a subsequent measurement of J_1 will yield $\pm 1/2$ with equal probability. The measurement of J_2 has disturbed the state.

It should perhaps be remarked that the term “precision” here is in the frequency sense: Imagine that you can prepare the identical state many times and repeat the measurements. The measurements will yield different results among the samplings, with expectation values as we have calculated, in the limit of averaging over an infinite number of samplings.

Finally, let us also look at:

$$\Delta J_2 \Delta J_3 = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \geq \frac{1}{2} |\langle J_1 \rangle| = \frac{1}{4}. \quad (95)$$

Again, the uncertainty principle is satisfied.

6 Exercises

1. Show that L^2 is complete.
2. Complete the proof that the space $L^2(a, b)$ is separable.
3. Show that if $x \in H$, where H is a separable Hilbert space, is orthogonal to every vector in a dense set, then $x = 0$.
4. Complete the proof of the Schwarz inequality.
5. Complete the derivation of Eqns. 80, 81, and 82.
6. Time Reversal in Quantum Mechanics:

We wish to define an operation of time reversal, denoted by T , in quantum mechanics. We demand that T be a “physically acceptable” transformation, *i.e.*, that transformed states are also elements of the Hilbert space of acceptable wave functions, and that it be consistent with the commutation relations between observables. We also demand that T have the appropriate classical correspondence with the classical time reversal operation.

Consider a system of structureless (“fundamental”) particles and let $\vec{X} = (X_1, X_2, X_3)$ and $\vec{P} = (P_1, P_2, P_3)$ be the position and momentum operators (observables) corresponding to one of the particles in the system. The commutation relations are, of course:

$$\begin{aligned} [P_m, X_n] &= -i\delta_{mn}, \\ [P_m, P_n] &= 0, \\ [X_m, X_n] &= 0. \end{aligned}$$

The time reversal operation $T : t \rightarrow t' = -t$, operating on a state vector gives (in Schrödinger picture – you may consider how to make the equivalent statement in the Heisenberg picture):

$$T|\psi(t)\rangle = |\psi'(t')\rangle.$$

The time reversal of any operator, Q , representing an observable is then:

$$Q' = TQT^{-1}$$

- (a) By considering the commutation relations above, and the obvious classical correspondence for these operators, show that

$$TiT^{-1} = -i.$$

Thus, we conclude that T must contain the complex conjugation operator K :

$$KzK^{-1} = z^*,$$

for any complex number z , we require that T on any state yields another state in the Hilbert space. We can argue that (for you to think about) we can write: $T = UK$, where U is a unitary transformation. If we operate twice on a state with T , then we should restore the original state, up to a phase:

$$T^2 = \eta 1,$$

where η is a pure phase factor (modulus = 1).

- (b) Prove that $\eta = \pm 1$. Hence, $T^2 = \pm 1$. Which phase applies in any given physical situation depends on the nature of U , and will turn out to have something to do with spin, as we shall examine in the future.
7. Let us consider the action of Gallilean transformations on a quantum mechanical wave function. We restrict ourselves here to the “proper” Gallilean Transformations: (i) translations; (ii) velocity boosts; (iii) rotations. We shall consider a transformation to be acting on the state (not on the observer). Thus, a translation by \mathbf{x}_0 on a state localized at \mathbf{x}_1 produces a new state, localized at $\mathbf{x}_1 + \mathbf{x}_0$. In “configuration space”, we have a wave function of the form $\psi(\mathbf{x}, t)$. A translation $T(\mathbf{x}_0)$ by \mathbf{x}_0 of this state yields a new state (please don’t confuse this translation operator with the time reversal operator of the previous problem, also denoted by T , but without an argument):

$$\psi'(\mathbf{x}) = T(\mathbf{x}_0)\psi(\mathbf{x}, t) = \psi(\mathbf{x} - \mathbf{x}_0, t). \quad (96)$$

Note that we might have attempted a definition of this transformation with an additional introduction of some overall phase factor. However, it is our interest to define such operators as simply as possible, consistent with what should give a valid classical correspondence. Whether we have succeeded in preserving the appropriate classical limit must be checked, of course.

Consider a free particle of mass m . The momentum space wave function is

$$\hat{\psi}(\mathbf{p}, t) = \hat{f}(\mathbf{p}) \exp\left(-\frac{itp^2}{2m}\right), \quad (97)$$

where $p = |\mathbf{p}|$. The configuration space wave function is related by the (inverse) Fourier transform:

$$\psi(\mathbf{x}, t) = \frac{1}{(2\pi)^{3/2}} \int_{(\infty)} d^3(\mathbf{p}) e^{i\mathbf{x}\cdot\mathbf{p}} \hat{\psi}(\mathbf{p}, t). \quad (98)$$

Obtain simple transformation laws, on both the momentum and configuration space wave functions, for each of the following proper Galilean transformations:

- (a) Translation by \mathbf{x}_0 : $T(\mathbf{x}_0)$ (note that we have already seen the result in configuration space).
- (b) Translation by time t_0 : $M(t_0)$.
- (c) Velocity boost by \mathbf{v}_0 : $V(\mathbf{v}_0)$. (Hint: first find

$$\hat{\psi}'(\mathbf{p}, 0) = \hat{f}'(\mathbf{p}) = V(\mathbf{v}_0)\hat{f}(\mathbf{p}), \quad (99)$$

then

$$\hat{\psi}'(\mathbf{p}, t) = \hat{f}'(\mathbf{p})e^{-itp^2/2m}, \quad (100)$$

etc.)

- (d) Rotation about the origin given by 3×3 matrix R : $U(R)$.

Make sure your answers make sense to you in terms of classical correspondence.

8. Consider the (real) vector space of real continuous functions with continuous first derivatives in the closed interval $[0, 1]$. Which of the following defines a scalar product?

$$(a) \langle f|g \rangle = \int_0^1 f'(x)g'(x)dx + f(0)g(0)$$

$$(b) \langle f|g \rangle = \int_0^1 f'(x)g'(x)dx$$

9. Consider the following equation in E_∞ (infinite-dimensional Euclidean space – let the scalar product be $\langle x|y \rangle \equiv \sum_{n=1}^\infty x_n^* y_n$):

$$Cx = a,$$

where the operator C is defined by (in some basis):

$$C(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$$

Is C :

- (a) A bounded operator [*i.e.*, does there exist a non-negative real number α such that, for every $x \in E_\infty$, we have $|Cx| \leq \alpha|x|$ (“ $|x|$ ” denotes the norm: $\sqrt{\langle x|x \rangle}$)?]
- (b) A linear operator?
- (c) A hermitian operator (*i.e.*, does $\langle x|Cy \rangle = \langle Cx|y \rangle$)?
- (d) Does $Cx = 0$ have a non-trivial solution? Does $Cx = a$ always have a solution?

Now answer the same questions for the operator defined by:

$$G(\alpha_1, \alpha_2, \dots) = (\alpha_1, \alpha_2/2, \alpha_3/3, \dots). \quad (101)$$

Note that we require a vector to be normalizable if it is to belong to E_∞ – *i.e.*, the scalar product of a vector with itself must exist.

10. Let $f \in L^2(-\pi, \pi)$ be a summable complex function on the real interval $[-\pi, \pi]$ (with Lebesgue measure).
- (a) Define the scalar product by:

$$\langle f|g \rangle = \int_{-\pi}^{\pi} f^*(x)g(x)dx, \quad (102)$$

for $f, g \in L^2(-\pi, \pi)$. Starting with the intuitive, but non-trivial, assumption that there is no vector in $L^2(-\pi, \pi)$ other than the

trivial vector ($f \sim 0$) which is orthogonal to all of the functions $\sin(nx)$, $\cos(nx)$, $n = 0, 1, 2, \dots$, show that any vector f may be expanded as:

$$f(x) = \sum_{n=0}^{\infty} (a_n \cos nx + b_n \sin nx), \quad (103)$$

where

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx \quad (104)$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx \quad (n > 0) \quad (105)$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx. \quad (106)$$

[You may consult a text such as Fano's *Mathematical Methods of Quantum Mechanics* for a full proof of the completeness of such functions.]

(b) Consider the function:

$$f(x) = \begin{cases} -1 & x < 0, \\ 0 & x = 0, \\ +1 & x > 0. \end{cases} \quad (107)$$

Determine the coefficients $a_n, b_n, n = 0, 1, 2, \dots$ for this function for the expansion of part (a).

(c) We wish to investigate the partial sums in this expansion:

$$f_N(x) = \sum_{n=0}^N (a_n \cos nx + b_n \sin nx). \quad (108)$$

Find the position, x_N of the first maximum of f_N (for $x > 0$). Evaluate the limit of $f_N(x_N)$ as $N \rightarrow \infty$. Give a numerical answer. In so doing, you are finding the maximum value of the series expansion in the limit of an infinite number of terms. [You may find the following identity useful:

$$\sum_{n=1}^N \cos(2n-1)x = \frac{1}{2} \frac{\sin 2Nx}{\sin x}.] \quad (109)$$

(d) Obviously, the maximum value of $f(x)$, defined in part (b), is 1. If the value you found for the series expansion is different from 1, comment on the possible reconciliation of this difference with the theorem you demonstrated in part (a).

11. Show that, with a suitable measure, any summation over discrete indices may be written as a Lebesgue integral:

$$\sum_{n=1}^{\infty} f(x_n) = \int_{\{x\}} f(x) \mu(dx). \quad (110)$$

12. Resonances II: Quantum mechanical resonances – Earlier we investigated some features of a classical oscillator with a “resonant” behavior under a driving force. Let us begin now to develop a quantum mechanical analogue, of relevance also to scattering and particle decays. For concreteness, consider an atom with two energy levels, $E_0 < E_1$, where the transition $E_0 \rightarrow E_1$ may be effected by photon absorption, and the decay $E_1 \rightarrow E_0$ via photon emission. Because the level E_1 has a finite lifetime – we denote the mean lifetime of the E_1 state by τ – it does not have a precisely defined energy. In other words, it has a finite width, which (assuming that E_0 is the ground state) can be measured by measuring precisely the distribution of photon energies in the $E_1 \rightarrow E_0$ decay. Call the mean of this distribution ω_0 .

(a) Assume that the amplitude for the atom to be in state E_1 is given by the damped oscillatory form:

$$\psi(t) = \psi_0 e^{-i\omega_0 t - \frac{t}{2\tau}}$$

Show that the mean lifetime is given by τ , as desired.

(b) Note that our amplitude above satisfies a “Schrödinger equation”:

$$i \frac{d\psi(t)}{dt} = \left(\omega_0 - \frac{i}{2\tau} \right) \psi(t)$$

Suppose we add a sinusoidal “driving force” $F e^{-i\omega t}$ on the right hand side, to describe the situation where we illuminate the atom with monochromatic light of frequency ω . Solve the resulting inhomogeneous equation for its steady state solution.

- (c) Convince yourself (*e.g.*, by “conservation of probability”) that the intensity of the radiation emitted by the atom in this steady-state situation is just $|\psi(t)|^2$. Thus, the incident radiation is “scattered” by our atom, with the amount of scattering proportional to the emitted radiation intensity in the steady state. Give an expression for the amount of radiation scattered (per unit time, per unit amplitude of the incident radiation), as a function of ω . For convenience, normalize your expression to the amount of scattering at $\omega = \omega_0$. Determine the full-width at half maximum (FWHM) of this function of ω , and relate to the lifetime τ .

Note that the “Breit-Wigner” function is just the Cauchy distribution in probability.

13. Time Reversal in Quantum Mechanics, Part II

We earlier showed that the time reversal operator, T , could be written in the form:

$$T = UK,$$

where K is the complex conjugation operator and U is a unitary operator. We also found that

$$T^2 = \pm 1.$$

Consider a spinless, structureless particle. All kinematic operators for such a particle may be written in terms of the \vec{X} and \vec{P} operators, where

$$\begin{aligned} [P_j, X_k] &= -i\delta_{jk} \\ T\vec{X}T^{-1} &= \vec{X} \\ T\vec{P}T^{-1} &= -\vec{P} \end{aligned}$$

(where the latter two equations follow simply from classical correspondence).

If we work in a basis consisting of the eigenvectors of \vec{X} , the eigenvalues are simply the real position vectors, and hence:

$$U\vec{X}U^{-1} = \vec{X}.$$

In this basis, the matrix elements of \vec{P} may be evaluated:

$$\vec{P} = -i\vec{\nabla} :$$

$$\begin{aligned}
\langle \vec{x}_1 | \vec{P} | \vec{x}_2 \rangle &= \int_{(\infty)} \delta^{(3)}(\vec{x} - \vec{x}_1) (-i \vec{\nabla}_x) \delta^{(3)}(\vec{x} - \vec{x}_2) d^{(3)}\vec{x} \\
&= -i \vec{\nabla}_{x_1} \delta^{(3)}(\vec{x}_1 - \vec{x}_2).
\end{aligned}$$

Thus, these matrix elements are pure imaginary, and

$$K \vec{P} K^{-1} = -\vec{P},$$

which implies finally

$$U \vec{P} U^{-1} = \vec{P}.$$

We conclude that for our spinless, structureless particle:

$$U = 1 e^{i\theta},$$

where the phase θ may be chosen to be zero if we wish. In any event, we have:

$$T = e^{i\theta} K,$$

and

$$T^2 = e^{i\theta} K e^{i\theta} K = 1.$$

- (a) Show that, for a spin 1/2 particle, we may in the Pauli representation (that is, an angular momentum basis for our spin-1/2 system such that the angular momentum operators are given by one-half the Pauli matrices) write:

$$T = \sigma_2 K,$$

and hence show that:

$$T^2 = -1.$$

Note that the point here is to consider the classical correspondence for the action of time reversal on angular momentum.

By considering a direct product space made up of many spin-0 and spin 1/2 states (or by other equivalent arguments), this result may be generalized: If the total spin is 1/2-integral, then $T^2 = -1$; otherwise $T^2 = +1$.

- (b) Show the following useful general property of an antiunitary operator such as T :

Let

$$|\psi'\rangle = T|\psi\rangle$$

$$|\phi'\rangle = T|\phi\rangle.$$

Then

$$\langle\psi'|\phi'\rangle = \langle\phi|\psi\rangle.$$

This, of course, should agree nicely with your intuition about what time reversal should do to this kind of scalar product.

- (c) Show that, if $|\psi\rangle$ is a state vector in an “odd” system ($T^2 = -1$), then $T|\psi\rangle$ is orthogonal to $|\psi\rangle$.
14. Suppose we have a particle of mass m in a one-dimensional potential $V = \frac{1}{2}kx^2$ (and the motion is in one dimension). What is the minimum energy that this system can have, consistent with the uncertainty principle? [The uncertainty relation is a handy tool for making estimates of such things as ground state energies.]