

QUANTUM THEORY AND ANALYSIS

John C. Baez 1989

1. States, Observables, and Symmetries

The basic ingredients of a theory of a physical system are mathematical objects describing 1) a set \mathbf{S} of **states** of the system, 2) the set \mathbf{O} of **observables** of a system, that is, real-valued properties of the system which can be experimentally determined, and 3) the group \mathbf{G} of **symmetries** of the system. The difference between classical physics and quantum theory arises from very different choices of mathematical objects used to describe states, observables, and symmetries. But in both classical and quantum physics certain mathematical structures appear that are associated with $(\mathbf{S}, \mathbf{O}, \mathbf{G})$. Three of the most important are:

1. Given a state $\psi \in \mathbf{S}$ and an observable $A \in \mathbf{O}$ there is a probability measure $\mu_{A,\psi}$ on \mathbb{R} which represents the probability distribution of values that the observable A may be measured to have when the system is in the state ψ . (We are not assuming that the observable A has a unique numerical value in a given state ψ ; repeated measurements of the same observable a system in the same state may have *different* values, and these values are distributed according to a probability distribution.)

A **probability measure** μ on a measure space X is a finite measure such that $\mu(X) = 1$. Measures on a topological space, e.g. \mathbb{R} , will always be assumed to be defined on the σ -algebra of Borel sets.

2. The group \mathbf{G} has an action U on the set \mathbf{S} of states.

An **action** U of a group G on a set S is a mapping $U: G \times S \rightarrow S$, usually written as $(g, x) \mapsto U(g)x$, such that: 1) $U(e)x = x$ where e is the identity in G ; 2) $U(gh)x = U(g)U(h)x$. These imply $U(g^{-1}) = U(g)^{-1}$.

The most important and simple example is where $\mathbf{G} = \mathbb{R}$ and U is **time evolution**; i.e. given a state $\psi \in \mathbf{S}$ and $t \in \mathbb{R}$, the state $U(t)\psi$ is obtained by simply waiting t units of time, letting the state evolve according to the physical laws the system obeys. The requirement $U(t+s) = U(t)U(s)$ amounts to saying that the physical laws don't change with time, i.e. there are no time-dependent external influences on the system.

3. The group \mathbf{G} is a **topological group**, that is, a group with a topology such multiplication is continuous from $\mathbf{G} \times \mathbf{G}$ to \mathbf{G} and the map $g \mapsto g^{-1}$ is continuous from \mathbf{G} to \mathbf{G} . A continuous homomorphism $t \mapsto g_t$ from \mathbb{R} to \mathbf{G} is called a **one-parameter subgroup**,

though strictly speaking its *range* is a subgroup of \mathbf{G} . Given a one-parameter subgroup let $U(t) = U(g_t)$. Then associated to the one-parameter subgroup there is an observable $A \in \mathbf{O}$ called the **associated conserved quantity**, which has the property that $\mu_{A,U(t)\psi}$ is independent of t . For example, in physics the conserved quantity associated with time evolution is called **energy** or the **Hamiltonian**. In systems which have symmetry under spatial translations, the associated conserved quantities are called **momenta** (in the x, y , and z directions, for example). In systems which have rotational symmetry the associated conserved quantities are called **angular momenta** (about the x, y and z axes, for example). In electromagnetism there is a one-parameter group of symmetries called **global gauge transformations** which corresponds to the conserved quantity **electric charge**.

This relation between states, observables, and symmetries is less obvious and more interesting than the first two; it often goes by the name of **Noether's Theorem**, after Emmy Noether, the student of Hilbert, since a version of it applicable to problems in variational calculus was proved in her 'Invariante Variationsprobleme' article in 1918, available in English in *Transport Theory and Stat. Phys.* **1**, 186-207 (1971).

We now will say what mathematical objects fill the roles of $(\mathbf{S}, \mathbf{O}, \mathbf{G})$ in classical mechanics and quantum theory. In the case of quantum theory we will immediately be led to the spectral theorem and Stone's theorem, which we will then prove.

2. States, Observables, and Symmetries in Classical Mechanics

In classical mechanics, the set of states \mathbf{S} is $\mathbb{R}^{2n} = \mathbb{R}^n \oplus \mathbb{R}^n$, called **phase space**, and the first summand \mathbb{R}^n is called **configuration space**. One uses coordinates $\{q_i\}_{i=1}^n$ for configuration space; these coordinates are called **positions**, and one uses coordinates $\{q_i, p_i\}_{i=1}^n$ for phase space; the p_i being called **momenta**.

The relation between these 'momenta' and the 'momenta' mentioned above is indirect. A typical example of phase space appears in the description of a system of n point particles in ordinary 3-dimensional space; each particle's state is described by its position $\vec{q}_i = (q_{1i}, q_{2i}, q_{3i})$ and its momentum, or mass times velocity vector, $\vec{p}_i = (p_{1i}, p_{2i}, p_{3i})$, and the whole assemblage is described by the following vector in the phase space $\mathbf{S} = \mathbb{R}^{3n} \oplus \mathbb{R}^{3n}$: $(\vec{q}_1, \dots, \vec{q}_n, \vec{p}_1, \dots, \vec{p}_n)$.

The set of observables \mathbf{O} is the space $C(\mathbb{R}^{2n})$ of continuous real-valued functions on phase space. Given a state $\psi \in \mathbb{R}^{2n}$ and an observable $A \in C(\mathbb{R}^{2n})$, the probability measure $\mu_{A,\psi}$ mentioned above is just $\delta_{A(\psi)}$, the Dirac delta measure at $A(\psi) \in \mathbb{R}$. In other words, if the state is ψ the value of any observable A is *always* measured to be $A(\psi)$; there's no 'uncertainty' in the value of A .

We will consider only the simplest case, in which the group \mathbf{G} of symmetries is just \mathbb{R} . Given a smooth function $H \in C^\infty(\mathbb{R}^{2n})$, we define the associated vector field v_H on

\mathbb{R}^{2n} by

$$v_H = \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial}{\partial p_i}.$$

A vector field v on \mathbb{R}^{2n} is **integrable** if there is an action U of \mathbb{R} on \mathbb{R}^{2n} such that

$$\frac{d}{dt}U(t)\psi = v(U(t)\psi).$$

If v_H is integrable and U is an action of \mathbb{R} on \mathbb{R}^{2n} such that $\frac{d}{dt}U(t)\psi = v_H(U(t)\psi)$, and we write in coordinates $U(t)\psi = (q_i, p_i) \in \mathbb{R}^{2n}$, then **Hamilton's equations** hold:

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \quad ; \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}.$$

In classical mechanics the action U of the group \mathbb{R} on states is required to be one arising from some $H \in C^\infty(\mathbb{R}^{2n})$ in this way, and H is the conserved quantity associated to U .

Exercise: Show that Hamilton's equations are indeed equivalent to

$$\frac{d}{dt}U(t)\psi = v_H(U(t)\psi).$$

Show that if $H \in C^\infty(\mathbb{R}^{2n})$ and v_H is integrable, the directional derivative $v_H H$ equals zero. Show that this implies that the measure $\mu_{H,U(t)x}$ is independent of t for any state $x \in \mathbb{R}^{2n}$. (This is a baby version of Noether's theorem.)

Exercise: Show that if $H = \frac{1}{2} \sum_{i=1}^n q_i^2 + p_i^2$ (the **harmonic oscillator Hamiltonian**) the vector field v_H is integrable by explicitly solving Hamilton's equations. Show that if $H = \sum_{i=1}^n \frac{1}{2}(q_i^2 + p_i^2) + \frac{1}{4}\lambda_i p_i^4$ (the **anharmonic oscillator Hamiltonian**) the vector field v_H is integrable if and only if all the **coupling constants** λ_i are non-negative.

The setup for classical mechanics as we have described it can be substantially generalized. The three most important directions of generalization are: 1) The group of symmetries \mathbf{G} can be an arbitrary **Lie group** (a group which is also a smooth manifold, such that the group operations are smooth maps). 2) The phase space \mathbf{S} can be an arbitrary **symplectic manifold** (a manifold with a non-degenerate closed 2-form). 3) The group \mathbf{G} and the phase space \mathbf{S} can be infinite-dimensional; this occurs in the study of classical field theory, i.e. the study of wave equations.

3. States, Observables and Symmetries in Quantum Theory

In quantum mechanics the set of states \mathbf{S} is the unit sphere in a Hilbert space \mathbf{H} ; i.e.

$$\mathbf{S} = \{\psi \in \mathbf{H}: \|\psi\| = 1\}.$$

As we shall see, for many purposes it's better to take the state space to be the quotient of this sphere by the equivalence relation: $\psi \sim \phi$ if $\psi = c\phi$ for some $c \in \mathbb{C}$, i.e. our present description is 'redundant', but this is unimportant for introductory purposes.

The set of observables \mathbf{O} is then the set of self-adjoint operators on \mathbf{H} . Recall that a **operator on \mathbf{H}** is a linear map A from a linear subspace of \mathbf{H} to \mathbf{H} . This subspace, the domain of A , is denoted $D(A)$, and we are *not* requiring even that it be dense in \mathbf{H} . If $D(A)$ is dense in \mathbf{H} we say that A is **densely defined**. An operator with domain equal to all of \mathbf{H} , satisfying

$$\exists c \geq 0 \forall \psi \in \mathbf{H} \quad \|A\psi\| \leq c\|\psi\|$$

is called a **bounded** operator, and the space of all bounded operators on \mathbf{H} is denoted by $\mathcal{L}(\mathbf{H})$. The **adjoint** of an operator A is supposed to have the property $\langle \phi, A\psi \rangle = \langle A^*\phi, \psi \rangle$, where we use $\langle \cdot, \cdot \rangle$ for the inner product in \mathbf{H} . (Also, we follow the physicists' convention that the inner product is linear in the *second* variable.) But for operators that are not bounded this requires some care in formulation. Thus we define the **adjoint** of a densely defined operator A to be the operator A^* with domain

$$D(A^*) = \{\phi \in \mathbf{H}: \exists \phi' \in \mathbf{H} \forall \psi \in D(A) \quad \langle \phi, A\psi \rangle = \langle \phi', \psi \rangle\}$$

and noting that if ϕ' with the above property exists it is *unique* (because A is densely defined), we define $A^*\phi$ for $\phi \in D(A^*)$ to be the unique ϕ' such that

$$\forall \psi \in D(A) \quad \langle \phi, A\psi \rangle = \langle \phi', \psi \rangle.$$

An operator A on a Hilbert space is said to be **self-adjoint** if it is densely defined and equal to its adjoint (which in particular means that $D(A) = D(A^*)$.)

The definition of adjoint may look complicated, but it is the most natural possible definition for densely defined operators. It foreshadows the fact that densely defined (unbounded) operators are altogether more subtle than bounded operators. Physicists typically ignore these nuances, and indeed the fastest way to tell if a physicist is a 'mathematical physicist' is to ask him to define 'adjoint'.

Given a state ψ in \mathbf{S} and a self-adjoint operator A on \mathbf{H} there is a probability measure $\mu_{A,\psi}$ on \mathbb{R} , called the 'spectral measure', which describes the probability distribution

of measured values for A when the system is in the state ψ . This measure arises as a consequence of the spectral theorem, which we shall soon state and prove.

A bounded operator T on a Hilbert space \mathbf{H} is **unitary** if $TT^* = T^*T = I$, or equivalently if T is onto and $\|T\psi\| = \|\psi\|$ for all $\psi \in \mathbf{H}$. Recall that an action U of a group G on a vector space V is a **representation** if $U(g)$ is linear for all $g \in G$; if in addition S is a Hilbert space and $U(g)$ is unitary for all $g \in G$ we say U is a **unitary representation**. If \mathbf{G} is a topological group we say that a representation of \mathbf{G} on a Hilbert space is **strongly continuous** if $g_\alpha \rightarrow g$ in \mathbf{G} implies $\|U(g_\alpha)\psi - U(g)\psi\| \rightarrow 0$ for all $\psi \in \mathbf{H}$. In quantum theory one requires that the group \mathbf{G} of symmetries be a topological group and that the action U of \mathbf{G} on the states \mathbf{S} come from a strongly continuous unitary representation of \mathbf{G} on \mathbf{H} . Stone's theorem then associates to any one-parameter subgroup g_t of \mathbf{G} a conserved quantity A , a self-adjoint operator on \mathbf{H} also called the 'generator' of $U(t) = U(g_t)$.

We recall here the three most important topologies on $\mathcal{L}(\mathbf{H})$. The **norm** topology is defined by: $T_\alpha \rightarrow T$ if $\|T_\alpha - T\| \rightarrow 0$. The **strong** topology is defined by: $T_\alpha \rightarrow T$ if $\|T_\alpha\psi - T\psi\| \rightarrow 0$ for all $\psi \in \mathbf{H}$. And the **weak** topology is defined by: $T_\alpha \rightarrow T$ if $\langle \phi, (T_\alpha - T)\psi \rangle \rightarrow 0$ for all $\psi, \phi \in \mathbf{H}$. A function from or to $\mathcal{L}(\mathbf{H})$ is said to be norm-continuous, strongly continuous, or weakly continuous depending on which topology is used for $\mathcal{L}(\mathbf{H})$.

Exercise - Let ℓ^2 denote the Hilbert space of complex sequences $\psi = (\psi_1, \psi_2, \dots)$ with finite norm $\|\psi\| = (\sum |\psi_i|^2)^{1/2}$, given the inner product $\langle \psi, \phi \rangle = \sum \bar{\psi}_i \phi_i$. Let T be the **left shift** operator, given by $T(\psi_1, \psi_2, \dots) = (\psi_2, \psi_3, \dots)$. Calculate $T^*(\psi_1, \psi_2, \dots)$ and explain why T^* is called the **right shift** operator. These shift operators illustrate some of the things that bounded linear operators can do in infinite dimensions that they cannot in finite dimensions. For example, show that T^* is an **isometry**, e.g. $\|T^*\psi\| = \|\psi\|$, but not unitary. Show that an isometry of a finite-dimensional Hilbert space is unitary. Show that T and T^* do not commute.

Exercise - Show that norm convergence implies strong convergence, which implies weak convergence, but that the converses do not hold. (For counterexamples, consider the sequences $\{T^n\}$ and $\{T^{*n}\}$, where T is the left shift.) Show that converses hold in a finite-dimensional Hilbert space. Show also that if U_α are *unitary* and $U_\alpha \rightarrow U$ weakly then $U_\alpha \rightarrow U$ strongly; thus a weakly continuous unitary representation is automatically strongly continuous.

Exercise - Show that for bounded operators S, T we have $T^{**} = T$, $\|T^*\| = \|T\|$, $\|T^*T\| = \|T\|^2$, $(cT)^* = \bar{c}T^*$, $(S+T)^* = S^* + T^*$, and $(ST)^* = T^*S^*$. Show that the map $T \mapsto T^*$ is norm-continuous and weakly continuous but not strongly continuous. (For the counterexample consider shift operators.)

4. The Spectral Theorem - Bounded Case

We begin by proving the spectral theorem for bounded self-adjoint operators and then deal with the unbounded case. First we define the **spectrum** of a bounded operator (not necessarily self-adjoint) $T \in \mathcal{L}(\mathbf{H})$ to be the set of $\lambda \in \mathbb{C}$ such that $T - \lambda I$ is not invertible, i.e. is not a bijection with a bounded inverse. The spectrum of T is denoted $\sigma(T)$. The complement of the spectrum of T is called the **resolvent set** of T and denoted $\rho(T)$. The exercises below show that the spectrum is a natural generalization of the set of eigenvalues of a matrix.

Exercise - Show that if $\mathbf{H} = \mathbb{C}^n$ the spectrum of $T \in \mathcal{L}(H)$ is just the set of $\lambda \in \mathbb{C}$ such that $\det(T - \lambda I) = 0$. Show *in this case* that the spectrum of T is the set of eigenvalues of T (for self-adjoint T this is easy; for general T one can use the Jordan canonical form - see any advanced book on linear algebra, e.g. Birkhoff and MacLane, *A Survey of Modern Algebra*.)

Exercise - Given a bounded operator T on \mathbf{H} define $\lambda \in \mathbb{C}$ to be an **approximate eigenvalue** of T if for any $\epsilon > 0$ there exists $\psi \in \mathbf{H}$ with $\|\psi\| = 1$ and $\|T\psi - \lambda\psi\| \leq \epsilon$. Show that if λ is an approximate eigenvalue for T then $\lambda \in \sigma(T)$. For a partial converse, show that if A is a bounded self-adjoint operator on a Hilbert space and $\lambda \in \sigma(A)$, then λ is an approximate eigenvalue of A . Hint for the second part: show that for any densely defined operator T on $L(\mathbf{H})$,

$$\text{Ran}(T)^\perp = \text{Ker}(T^*),$$

so that for bounded self-adjoint A , $(A - \lambda I)$ is invertible unless λ is an approximate eigenvalue of A .

Exercise - Show that $\lambda \in \sigma(T)$ implies $\bar{\lambda} \in \sigma(T^*)$. Show that if A is bounded and self-adjoint $\sigma(A^*) \subseteq \mathbb{R}$.

The Spectral Mapping Theorem (for polynomials). *If T is a bounded operator and P is a nonconstant polynomial then $\sigma(P(T)) = \{P(\lambda) : \lambda \in \sigma(T)\}$. (For short we simply write $\sigma(P(T)) = P(\sigma(T))$.)*

Proof - Let $\lambda \in \sigma(T)$. Since $x = \lambda$ is a root of $P(x) - P(\lambda)$, we can factor:

$$P(x) - P(\lambda) = (x - \lambda)Q(x),$$

so $P(T) - P(\lambda) = (T - \lambda I)Q(T)$. Since the operator $T - \lambda I$ is not invertible neither is $P(T) - P(\lambda)$, i.e. $P(\lambda) \in \sigma(P(T))$.

Conversely suppose $\mu \in \sigma(P(T))$. Factor $P(x) - \mu$ as $a(x - \lambda_1) \cdots (x - \lambda_n)$. Then $P(T) - \mu I = a(T - \lambda_1) \cdots (T - \lambda_n)$. If all the λ_i are *not* in $\sigma(T)$ then we have

$$(P(T) - \mu I)^{-1} = a^{-1}(T - \lambda_1)^{-1} \cdots (T - \lambda_n)^{-1},$$

so μ is *not* in $\sigma(P(T))$, a contradiction. Thus there must be some $\lambda_i \in \sigma(T)$, and

$$P(\lambda_i) - \mu = a(\lambda_i - \lambda_1) \cdots (\lambda_i - \lambda_n) = 0$$

so $\mu = P(\lambda_i)$. \square

Theorem. *The spectral radius of a bounded operator T is defined to be $r(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}$. If A is a bounded self-adjoint operator on \mathbf{H} then $r(A) = \|A\|$.*

Proof - If $\lambda \in \sigma(A)$ then λ is an approximate eigenvalue for A so there exists $\psi_n \in \mathbf{H}$ with $\|\psi_n\| = 1$ and $\|A\psi_n\| \rightarrow |\lambda|$. Thus $\|A\| \geq |\lambda|$, so $\|A\| \geq r(A)$.

Now choose ψ_n such that $\|\psi_n\| = 1$ and $\|A\psi_n\| \rightarrow \|A\|$. Let $\lambda = \|A\|$. If A is self-adjoint then

$$\begin{aligned} \|A^2\psi_n - \lambda^2\psi_n\|^2 &= \langle A^2\psi_n - \lambda^2\psi_n, A^2\psi_n - \lambda^2\psi_n \rangle \\ &= \|A^2\psi_n\|^2 - 2\lambda^2\|A\psi_n\|^2 + \lambda^4 \\ &\leq (\|A\| \|A\psi_n\|)^2 - 2\lambda^2\|A\psi_n\|^2 + \lambda^4 \\ &= -\lambda^2\|A\psi_n\|^2 + \lambda^4 \rightarrow 0. \end{aligned}$$

Thus λ^2 is an approximate eigenvalue for A^2 , so $\lambda^2 \in \sigma(A^2) = \sigma(A)^2$. It follows that $\|A\| \leq r(A)$. Thus $\|A\| = r(A)$. \square

The remarkable thing about this theorem is that it relates a topological concept, the norm, to a concept which is defined purely algebraically, the spectral radius. This interplay of topology and algebra is behind the beautiful theory of C^* -algebras, a kind of algebra over the complex numbers which has a norm and $*$ operator satisfying axioms like those which hold for $\mathcal{L}(\mathbf{H})$. C^* -algebras are important in the mathematics of quantum field theory, and in understanding the relation between classical and quantum physics.

Exercise - Fill out the following proof of a generalization of the theorem above: if T is a bounded operator then $r(T) \leq \|T\|$ and

$$r(T) = \limsup_{n \rightarrow \infty} \|T^n\|^{1/n}.$$

The idea is to use the Laurent series for $(T - \lambda I)^{-1}$, which is modeled after the series for $(z - \lambda)^{-1}$:

$$(T - \lambda I)^{-1} = - \sum_{n=0}^{\infty} \lambda^{-n-1} T^n.$$

In the region in \mathbb{C} where the series is norm-convergent, term-by-term multiplication is justified in showing that $(T - \lambda I)$ multiplied by $-\sum_{n=0}^{\infty} \lambda^{-n-1} T^n$, in either order, is the identity. Recall however Hadamard's theorem that the radius of convergence of $\sum a_n z^n$ is $R = 1/\limsup_{n \rightarrow \infty} |a_n|^{1/n}$, and that $|\sum a_n z^n| \rightarrow \infty$ for some sequence of z 's with $|z| \rightarrow R$. The same argument (fortified with the uniform boundedness theorem) shows that the radius of convergence of $\sum T^n z^n$ is $r = 1/\limsup \|T^n\|^{1/n}$, and that $\|\sum T^n z^n\| \rightarrow \infty$ for some sequence of z 's with $|z| \rightarrow R$. Taking $\lambda = 1/z$, it follows that $(T - \lambda)^{-1}$ exists for $\lambda \geq R = \limsup \|T^n\|^{1/n}$, and

$$\|(T - \lambda I)^{-1}\| \rightarrow \infty$$

for some sequence of λ 's with $|\lambda|$ converging to R . This immediately implies that $r(T) \leq R$, and to show that it also implies $r(T) = R$ it suffices to show that $\|(T - \lambda I)^{-1}\|$ is continuous on $\rho(T)$, since then $r(T) < R$ would imply have $\|(T - \lambda I)^{-1}\|$ bounded near $|\lambda| = R$, a contradiction. The continuity follows from a similar power series argument.

Exercise - Let T denote the left shift operator on ℓ^2 . Compute the spectral radius of T , show that the whole unit circle in \mathbb{C} consists of approximate eigenvalues of T , and show that 0 is an eigenvalue of T . Show that the whole unit circle in \mathbb{C} consists of approximate eigenvalues of T^* and that 0 is in the spectrum of T^* , but 0 is not an approximate eigenvalue of T^* .

Theorem. *If T is a bounded operator its spectrum is compact.*

Proof - In the previous theorem we showed $r(T) \leq \|T\|$, so $\sigma(T)$ is bounded and it suffices to show that $\sigma(T)$ is closed or equivalently that the resolvent set $\rho(T)$ is open. Suppose $\lambda_0 \in \rho(T)$ and let $R = (T - \lambda_0 I)^{-1}$. Then for λ sufficiently close to λ_0 the power series

$$\sum_{n=0}^{\infty} (\lambda - \lambda_0)^n R^{n+1}$$

is norm-convergent and

$$\begin{aligned} (T - \lambda I) \left(\sum_{n=0}^{\infty} (\lambda - \lambda_0)^n R^{n+1} \right) &= \\ \sum_{n=0}^{\infty} [(T - \lambda_0 I) (\lambda - \lambda_0)^n R^{n+1} + (\lambda_0 - \lambda) (\lambda - \lambda_0)^n R^{n+1}] &= \\ \sum_{n=0}^{\infty} [(\lambda - \lambda_0)^n R^n - (\lambda - \lambda_0)^{n+1} R^{n+1}] &= I. \end{aligned}$$

The same argument shows that

$$\left[\left(\sum_{n=0}^{\infty} (\lambda - \lambda_0)^n R^{n+1} \right) \right] (T - \lambda I) = I,$$

so we have found an explicit inverse for $(T - \lambda I)$. Thus $\lambda \in \rho(T)$. \square

Note the use of power series techniques in both the proof of this theorem and in the proof that $r(T) = \limsup_{n \rightarrow \infty} \|T^n\|^{1/n}$ in the exercise above. Such techniques from complex analysis are very useful in operator theory; $(T - \lambda I)^{-1}$ is called an analytic operator-valued function on $\rho(T)$ because it admits such local power series expansions in λ .

Spectral Theorem - continuous functional calculus. *If A is bounded self-adjoint operator on the Hilbert space \mathbf{H} there is a unique map $\pi: C(\sigma(A)) \rightarrow \mathcal{L}(\mathbf{H})$ such that:*

1. π is a ***-homomorphism**, i.e.

$$\begin{aligned} \pi(\lambda f) &= \lambda \pi(f) \quad , \quad \pi(f + g) = \pi(f) + \pi(g) \quad , \quad \pi(fg) = \pi(f)\pi(g) \quad , \\ \pi(\bar{f}) &= \pi(f)^* \quad , \quad \pi(1) = I \quad , \end{aligned}$$

2. π is continuous, i.e., $\|\pi(f)\| \leq C\|f\|_{\infty}$,

3. If h is the function $h(\lambda) = \lambda$, then $\pi(h) = A$.

We shall usually write $f(A)$ for $\pi(f)$. In this notation, π has the following properties:

4. If $f \geq 0$ then $f(A) \geq 0$.

5. $\|f(A)\| = \|f\|_{\infty}$.

6. $\sigma(f(A)) = f(\sigma(A))$. (**The spectral mapping theorem.**)

7. If $A\psi = \lambda\psi$ then $f(A)\psi = f(\lambda)\psi$.

Proof - We first define π on polynomials by $\pi(p) = p(A)$. The idea is to show that the definition of a *polynomial function* of a self-adjoint operator extends to *continuous functions*, obtaining a ‘continuous functional calculus’. The trick is to use Weierstrass’ theorem, which says that the set of polynomials is dense in $C(\sigma(A))$, since $\sigma(A)$ is compact. If we can show 2 for polynomial f , it will then follow that π extends uniquely from the polynomials to all of $C(\sigma(A))$, and satisfies 2 for all f . (Here we are using the easy but all-important **B.L.T. theorem** (for ‘bounded linear transformation’), which says that if X and Y are Banach spaces, D is a dense subspace of X , and $f: D \rightarrow Y$ satisfies $\|f(x)\| \leq C\|x\|$ for all $x \in D$, then f extends uniquely to a bounded linear function from X to Y .) Thus we first verify 2 for a polynomial p . We have

$$\begin{aligned} \|\pi(f)\| &= \|p(A)\| \\ &= (\|p(A)^*p(A)\|)^{1/2} \\ &= \|(p\bar{p})(A)\|^{1/2} \\ &= r((p\bar{p})(A))^{1/2}, \end{aligned}$$

the latter since $p(A)^*p(A) = (p\bar{p})(A)$ is self-adjoint. By the spectral mapping theorem for polynomials we have

$$\begin{aligned} r((p\bar{p})(A)) &= \sup\{p(\lambda)\bar{p}(\lambda): \lambda \in \sigma(A)\} \\ &= \|p\|_\infty^2, \end{aligned}$$

Thus

$$\|\pi(p)\| = \|p\|_\infty$$

if p is the restriction of a polynomial to $\sigma(A)$, verifying 2 for polynomials.

It follows from the B.L.T. theorem that we may uniquely extend π to continuous functions, and that 2 and 7 hold. Clearly 3 holds, and since 1 holds for polynomial f, g its easy to see that it extends by continuity to hold for all $f, g \in C(\sigma(A))$. It also follows from the B.L.T. theorem that π as we’ve defined it is the *unique* function satisfying 1 - 3.

Property 4 clearly holds for polynomial f and extends by continuity to all $f \in C(\sigma(A))$. We’ve shown that 5 holds. 7 holds for polynomials and extends by continuity to all $f \in C(\sigma(A))$.

It remains to prove 6, which we know holds for polynomials. We show that $f(\sigma(A)) \subseteq \sigma(f(A))$ as follows. Let $\lambda \in \sigma(A)$; λ is an approximate eigenvalue of A so choose $\psi_n \in \mathbf{H}$ with $\|\psi_n\| = 1$ and $\|(A - \lambda I)\psi_n\| \leq 1/n$. Choose polynomials P_m converging uniformly to

f on $\sigma(A)$, with $\|P_m - f\| \leq 1/m$. Then for any m ,

$$\begin{aligned} \|(f(A) - f(\lambda)I)\psi_n\| &\leq \|(f(A) - P_m(A))\psi_n\| + \|(P_m(A) - P_m(\lambda)I)\psi_n\| \\ &\quad + \|(P_m(\lambda) - f(\lambda))\psi_n\| \\ &\leq \|f(A) - P_m(A)\| + \|(P_m(A) - P_m(\lambda))\psi_n\| + \|P_m - f\|_{L^\infty(\sigma(A))} \\ &\leq \|(P_m(A) - P_m(\lambda))\psi_n\| + 2/m. \end{aligned}$$

Since $\|\psi_n\| = 1$ and $\|A\psi - \lambda\psi\| \leq 1/n$, we have, for any $k \geq 0$,

$$\begin{aligned} \|A^k\psi_n - \lambda^k\psi_n\| &= \|(A^{k-1} + \lambda A^{k-2} + \dots + \lambda^{k-1}I)(A - \lambda I)\psi_n\| \\ &\leq \|A^{k-1} + \lambda A^{k-2} + \dots + \lambda^{k-1}\|/n, \end{aligned}$$

which goes to zero as $n \rightarrow \infty$, so

$$\lim_{n \rightarrow \infty} \|(P_m(A) - P_m(\lambda))\psi_n\| = 0,$$

so for n large enough this implies that

$$\|(f(A) - f(\lambda)I)\psi_n\| \leq 3/m.$$

Since m is arbitrary this implies that

$$\|(f(A) - f(\lambda)I)\psi_n\| \rightarrow 0.$$

Thus $f(\lambda)$ is an approximate eigenvalue for $f(A)$, so $f(\lambda) \in \sigma(f(A))$.

To show that $\sigma(f(A)) \subseteq f(\sigma(A))$, suppose that $\lambda \notin f(\sigma(A))$. Then the function $g(x) = (f(x) - \lambda)^{-1}$ is continuous on $\sigma(A)$, so we obtain $g(A)$ by the functional calculus, and

$$g(A)(f(A) - \lambda I) = I = (f(A) - \lambda I)g(A)$$

by the *-homomorphism property. Thus $f(A) - \lambda I$ is invertible, so $\lambda \notin \sigma(f(A))$. \square

The spectral theorem has many versions, and a long and tangled history. There are three main forms of the spectral theorem: the functional calculus form, the projection-valued measure form (see the remarks on spectral projections in the next section), and the following form: if A is a self-adjoint operator on a Hilbert space \mathbf{H} , there is a measure space (X, μ) , a real-valued measurable function f on X , and a unitary transformation $U : \mathbf{H} \rightarrow L^2(X)$ such that $A = U^{-1}M_fU$, where M_f denotes multiplication by f . That is, every self-adjoint operator is unitarily equivalent to a multiplication operator.

Exercise - Prove the B.L.T. theorem.

Note that if $\psi \in \mathbf{H}$ the above theorem implies that $f \mapsto \langle \psi, f(A)\psi \rangle$ is a continuous linear functional from $C(\sigma(A))$ to \mathbf{C} , which moreover is positive by property 4. Thus by the Riesz-Markov theorem there is a unique measure $\mu_{A,\psi}$ on $\sigma(A)$, the **spectral measure**, such that

$$\langle \psi, f(A)\psi \rangle = \int_{\sigma(A)} f(\lambda) d\mu_{A,\psi}(\lambda) \quad .$$

If $\|\psi\| = 1$ we have $\int_{\sigma(A)} d\mu_{A,\psi}(\lambda) = 1$, so the spectral measure is a probability measure. This makes good on our promise that in quantum theory the spectral theorem would yield a probability measure given a state and an observable - though so far we've only shown this for bounded A .

The spectral measure also immediately allows us to strengthen the above theorem:

Spectral Theorem - measurable functional calculus. *Let A be a bounded self-adjoint operator on \mathbf{H} . Let $\mathcal{B}(\sigma(A))$ denote the bounded Borel functions on $\sigma(A)$ - not identifying those which are equal a.e.. Given A on \mathbf{H} and $f \in \mathcal{B}(\sigma(A))$, there is a unique bounded operator $\pi(f) = f(A)$ such that*

$$\langle \psi, f(A)\psi \rangle = \int_{\sigma(A)} f(\lambda) d\mu_{A,\psi}(\lambda) \quad .$$

Then $\pi : \mathcal{B}(\mathbf{R}) \rightarrow \mathcal{L}(\mathbf{H})$ is the unique map satisfying:

1. π is a $*$ -homomorphism,
2. $\|f(A)\| = \|f\|'_\infty$, where the L^∞' norm is defined with the following modified notion of 'almost everywhere': we say a property holds a.e.' if for each $\psi \in \mathbf{H}$ the property holds except for a set of measure zero relative to $\mu_{A,\psi}$.

3. If h is the function $h(\lambda) = \lambda$, then $h(A) = A$.

Moreover 4. If $f \geq 0$ then $f(A) \geq 0$.

5. $\sigma(f(A)) = f(\sigma(A))$.
6. If $A\psi = \lambda\psi$ then $f(A)\psi = f(\lambda)\psi$.
7. If $AT = TA$ for some $T \in \mathcal{L}(\mathbf{H})$, then $f(A)T = Tf(A)$.
8. Suppose that $f_n \rightarrow f$ pointwise a.e.' and $\|f_n\|'_\infty$ is bounded. Then $f_n(A) \rightarrow f(A)$ strongly.

Sketch of proof - For the uniqueness one uses the polarization identity (see below), which implies that a bounded operator is determined by the inner products $\langle \psi, T\psi \rangle$. Property 8 follows from the dominated convergence theorem and the rest of the properties follow

by limiting arguments and the previous version of the spectral theorem, using Lusin's theorem that given any Borel measure μ on \mathbb{R} , any $f \in \mathcal{B}(\mathbb{R})$ is a pointwise μ -a.e. limit of a uniformly bounded sequence of continuous functions. \square

Exercise - Prove the **polarization identity** for any $T \in \mathcal{L}(\mathbf{H})$:

$$\langle \phi, T\psi \rangle = \frac{1}{4} \left[\{ \langle \psi + \phi, T(\psi + \phi) \rangle - \langle \psi - \phi, T(\psi - \phi) \rangle \} \right. \\ \left. - i \{ \langle \psi + i\phi, T(\psi + i\phi) \rangle - \langle \psi - i\phi, T(\psi - i\phi) \rangle \} \right].$$

Exercise - Work out the details of the proof of the above theorem.

The following exercises contain important generalizations of the spectral theorem, and the results will be used in the next section:

Exercise - Prove the following:

Spectral Theorem for Collections. Let A_1, \dots, A_n be a collection of pairwise commuting bounded self-adjoint operators on \mathbf{H} . Let

$$X = \sigma(A_1) \times \dots \times \sigma(A_n) \subseteq \mathbb{R}^n.$$

Let $\mathcal{B}(X)$ denote the bounded Borel functions on X , again not identifying those which are equal a.e.. There is a map $\pi : \mathcal{B}(X) \rightarrow \mathcal{L}(\mathbf{H})$ and for each $\psi \in \mathbf{H}$ a measure μ_ψ on X such that

$$\langle \psi, \pi(f)\psi \rangle = \int_X f(x) d\mu_\psi(x)$$

and:

1. π is a $*$ -homomorphism,
2. $\|\pi(f)\| \leq \|f\|'_\infty$, where the $L^{\infty'}$ norm is defined with the following modified notion of 'almost everywhere': we say a property holds a.e.' if for each $\psi \in \mathbf{H}$ the property holds except for a set of measure zero relative to μ_ψ .
3. If $h: X \rightarrow \mathbb{R}$ is the function $h(\lambda_1, \dots, \lambda_n) = \lambda_i$, then $\pi(h) = A_i$.
4. If $f \geq 0$ then $\pi(f) \geq 0$.
5. If some $T \in \mathcal{L}(\mathbf{H})$ has $A_i T = T A_i$ for all i , then $\pi(f)T = T\pi(f)$.

6. Suppose that $f_n \rightarrow f$ pointwise a.e.' and $\|f_n\|'_\infty$ is bounded. Then $f_n(A) \rightarrow f(A)$ strongly.

One usually writes $f(A_1, \dots, A_n)$ for $\pi(f)$.

To prove this one can essentially copy the proof for a single operator. However to get the B.L.T. theorem to work don't use polynomials in the A_i , but instead linear combinations of 'rectangle functions'. Namely, define a **rectangle** in \mathbb{R}^n to be a set of the form $S_1 \times \dots \times S_n$ with each S_i a Borel subset of \mathbb{R} . Define π on characteristic functions of rectangles by:

$$\pi(\chi_R) = \chi_{S_1}(A_1) \cdots \chi_{S_n}(A_n)$$

where each $\chi_{S_i}(A_i)$ is defined by the spectral theorem. Using the spectral theorem for a single bounded operator, and in particular statement 8 of the 'measurable functional calculus' version, show that if R_i is a finite collection of disjoint rectangles then

$$\left\| \sum a_i \pi(\chi_{R_i}) \right\| \leq \left\| \sum a_i \chi_{R_i} \right\|_\infty .$$

Define π on finite linear combinations of rectangle functions by $\pi(\sum a_i \chi_{R_i}) = \sum a_i \pi(\chi_{R_i})$, and use the B.L.T. theorem to extend π to $\mathcal{B}(X)$.

Exercise - A bounded operator is **normal** if it commutes with its adjoint. E.g., unitary and bounded self-adjoint operators are normal. Show that if T is normal then $\operatorname{Re}T = \frac{1}{2}(T + T^*)$ and $\operatorname{Im}T = \frac{1}{2i}(T - T^*)$ are bounded self-adjoint and commute. Use the spectral theorem for collections to prove the following:

Spectral Theorem for Normal Operators. Let T be a normal operator on \mathbf{H} . Let $\mathcal{B}(\sigma(T))$ denote the bounded Borel functions on $\sigma(T)$. There is a map $\pi : \mathcal{B}(\sigma(T)) \rightarrow \mathcal{L}(\mathbf{H})$ and for each $\psi \in \mathbf{H}$ a measure $\mu_{T,\psi}$ on X such that

$$\langle \psi, f(T)\psi \rangle = \int_{\sigma(T)} f(\lambda) d\mu_{T,\psi}(\lambda) .$$

and:

1. π is a *-homomorphism,
2. $\|f(A)\| = \|f\|'_\infty$, where the L^∞ ' norm is defined with the following modified notion of 'almost everywhere': we say a property holds a.e.' if for each $\psi \in \mathbf{H}$ the property holds except for a set of measure zero relative to μ_ψ .

3. If h is the function $h(\lambda) = \lambda$, then $h(T) = T$.
4. If $f \geq 0$ then $f(T) \geq 0$.
5. $\sigma(f(T)) = f(\sigma(T))$.
6. If $T\psi = \lambda\psi$ then $f(T)\psi = f(\lambda)\psi$.
7. If $ST = TS$ for some $S \in \mathcal{L}(\mathbf{H})$, then $f(S)T = Tf(S)$.
8. Suppose that $f_n \rightarrow f$ pointwise a.e. and $\|f_n\|'_\infty$ is bounded. Then $f_n(T) \rightarrow f(T)$ strongly.

Exercise - Show using the theorem above that if T is normal, the spectrum of T is the closure of $\bigcup_{\psi \in \mathbf{H}} \text{supp} \mu_{T, \psi}$. (Hint: show that if λ is not in the closure of $\bigcup \text{supp} \mu_{T, \psi}$, one can construct $(T - \lambda I)^{-1}$ using the functional calculus.)

Exercise - Use the exercise above to show that the spectrum of a normal operator consists of approximate eigenvalues. Conclude that if U is unitary, $\sigma(U)$ is a subset of the unit circle in \mathbb{C} .

Exercise - Assume $A \in \mathcal{L}(\mathbf{H})$. Use the polarization identity to show that if $A \geq 0$, i.e. $\langle \psi, A\psi \rangle \geq 0$ for all $\psi \in \mathbf{H}$, then A is self-adjoint.

Exercise - Show that $A \geq 0$ if and only if $\sigma(A)$ is a subset of $[0, \infty)$ (for one direction use the previous exercise). Show that if $A \geq 0$ there exists $B \in \mathcal{L}(\mathbf{H})$ with $B \geq 0$ and $B^2 = A$.

5. The Spectral Theorem - Unbounded Case

The spectral theorem for unbounded self-adjoint operators proceeds by reduction to the bounded case. We begin by establishing a useful property of self-adjoint operators. We define an operator T on \mathbf{H} to be **symmetric** if it is densely defined and for all $\psi, \phi \in D(T)$ we have

$$\langle \phi, T\psi \rangle = \langle T\phi, \psi \rangle.$$

Note that this is equivalent to saying that $T \subseteq T^*$, since it says that for every $\phi \in D(T)$ there exists $\phi' \in \mathbf{H}$ (namely $T\phi$) such that

$$\langle \phi, T\psi \rangle = \langle \phi', \psi \rangle,$$

so that $D(T) \subseteq D(T^*)$, and $T^*\phi = \phi' = T\phi$, so that $T \subseteq T^*$. Thus every self-adjoint operator is symmetric, but as we'll see the converse does *not* hold.

Beware: physicists often say self-adjoint when they mean symmetric. This is a result of insufficient education and the following facts: 1) ‘sufficiently nice’ symmetric operators are self-adjoint, and 2) it’s much easier to check if an operator is symmetric than if its self-adjoint - in other words, it can be very difficult to show that a symmetric operator is ‘sufficiently nice’. To check that an operator T is symmetric one must show that $\langle \phi, T\psi \rangle = \langle T\phi, \psi \rangle$ for every $\phi, \psi \in D(T)$; this is often easy to do, and if T is a differential operator it usually amounts to integration by parts. To check that T is self-adjoint one must *also* show that if $\phi \in \mathbf{H}$ has a $\phi' \in \mathbf{H}$ such that $\langle \phi, T\psi \rangle = \langle \phi', \psi \rangle$, then ϕ lies in $D(T)$. This is usually done indirectly using various theorems, some of which we will discuss later, and many problems in mathematical physics consist of proving that symmetric operators are self-adjoint. This is because the spectral theorem only holds for self-adjoint operators, so self-adjoint operators are much better than merely symmetric operators.

We also note term ‘hermitian’ is also used by mathematicians to mean symmetric, and by physicists to mean self-adjoint, by which they mean symmetric. . . we will avoid this term.

We define an operator T to be **closed** if given $\psi_i \in D(T)$ such that $\psi_i \rightarrow \psi \in \mathbf{H}$ and $T\psi_i \rightarrow \psi'$ then $\psi \in D(T)$ and $T\psi = \psi'$. If T is an operator and $T \subseteq S$ for some operator S , we say S is an **extension** of T . If T is an operator with a closed extension we say T is **closable**. If T is closable the intersection of all closed extensions of T is a closed extension of T called the **closure** of T , written \bar{T} .

Exercise - Prove the claim above the intersection of a set of closed extensions is a closed extension.

Exercise - Show that if T is a densely defined operator satisfying $\|T\psi\| \leq C\|\psi\|$ then T has a unique closed extension, and that this extension is a bounded operator.

Exercise - An operator that’s not closable. Let ψ_n be an orthonormal basis for a Hilbert space \mathbf{H} and let $\phi \in \mathbf{H}$ be a vector that is not a linear combination of finitely many ψ_n . Let D be the set of finite linear combinations of ϕ and the ψ_n , and let T be the operator with domain D defined by:

$$T \left(a\phi + \sum_{n=1}^N b_n \psi_n \right) = a\phi .$$

Show that T is a densely defined linear operator. Show that any closed extension S of T would have $S\phi = \phi$ and $S\phi = 0$, obtaining a contradiction.

Exercise - A nonzero operator whose adjoint is zero (roughly speaking). Let ψ_n be an orthonormal basis for a Hilbert space \mathbf{H} and let D be the set of all finite linear combinations of the ψ_n . Let T be the operator with domain D defined by:

$$T \left(\sum_{n=1}^N a_n \psi_n \right) = \left(\sum_{n=1}^N a_n \right) \psi_1 .$$

Show that T is a densely defined linear operator. Show that the domain of T^* consists of all vectors orthogonal to ψ_1 , and that T^* is zero on this domain.

Note that in matrix language

$$T = \begin{pmatrix} 1 & 1 & 1 & \cdots \\ 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \cdot & \cdot & \cdot & \cdots \end{pmatrix}$$

So why is T^* zero?

The Basic Criterion for Self-adjointness. *Let T be a symmetric operator on a Hilbert space \mathbf{H} . Then the following three statements are equivalent:*

1. T is self-adjoint.
2. T is closed and $\text{Ker}(T^* \pm i) = \{0\}$.
3. $\text{Ran}(T \pm i) = \mathbf{H}$.

Proof. $1 \Rightarrow 2$: Assume T is self-adjoint. To show that T is closed we need to show that if $\psi_i \rightarrow \psi \in \mathbf{H}$ and $T\psi_i \rightarrow \psi'$ then $\psi \in D(T)$ and $T\psi = \psi'$. Since $T = T^*$ we only need to show that $\psi \in D(T^*)$ and $T^*\psi = \psi'$. In other words, we need to show that for all $\phi \in D(T)$,

$$\langle \psi, T\phi \rangle = \langle \psi', \phi \rangle.$$

Since T is self-adjoint and $\psi_i \in D(T)$ we have:

$$\begin{aligned} \langle \psi, T\phi \rangle &= \lim_{i \rightarrow \infty} \langle \psi_i, T\phi \rangle \\ &= \lim_{i \rightarrow \infty} \langle T\psi_i, \phi \rangle \\ &= \langle \psi', \phi \rangle \end{aligned}$$

as desired.

Next we need to show that if T is self-adjoint and $(T + i)\psi = 0$ or $(T - i)\psi = 0$ for some $\psi \in D(T)$, then $\psi = 0$. If $(T - i)\psi = 0$,

$$\begin{aligned} \langle \psi, i\psi \rangle &= \langle \psi, T\psi \rangle \\ &= \langle T\psi, \psi \rangle \\ &= \langle i\psi, \psi \rangle \\ &= -\langle \psi, i\psi \rangle, \end{aligned}$$

so $\psi = 0$, and the other case works similarly.

2 \Rightarrow 3: Assume T is symmetric and closed and $\text{Ker}(T^* \pm i) = 0$. We shall show $\text{Ran}(T \pm i)$ is closed and dense in \mathbf{H} . Since $\text{Ker}(T^* \pm i) = \text{Ran}(T \mp i)^\perp$, and the orthogonal complement of a subspace is empty if and only if the subspace is dense, clearly $\text{Ran}(T \pm i)$ is dense.

To show that $\text{Ran}(T + i)$ is closed we naturally use the hypothesis that T is closed (the proof that $\text{Ran}(T - i)$ is closed is similar). Suppose we have $(T + i)\psi_i \rightarrow \phi$ in \mathbf{H} . We need to show that there exists $\psi \in D(T)$ such that $\phi = (T + i)\psi$. We have

$$\begin{aligned} 0 &\leftarrow \|(T + i)\psi_i - (T + i)\psi_j\|^2 \\ &= \|T(\psi_i - \psi_j)\|^2 + \|\psi_i - \psi_j\|^2 \\ &\quad + \langle i(\psi_i - \psi_j), T(\psi_i - \psi_j) \rangle + \langle T(\psi_i - \psi_j), i(\psi_i - \psi_j) \rangle, \end{aligned}$$

and the last two terms cancel because T is symmetric. Thus $\psi_i \rightarrow \psi$ and $T\psi_i \rightarrow T\psi$ for some $\psi \in D(T)$, since T is closed. Thus we have $\phi = \lim(T + i)\psi_i = (T + i)\psi$.

3 \Rightarrow 1: Assume that $\text{Ran}(T \pm i) = \mathbf{H}$. Given $\phi \in D(T^*)$ we want to show that $\phi \in D(T)$, since this will imply $D(T) = D(T^*)$ which together with $T \subseteq T^*$ implies $T = T^*$.

Since $\text{Ran}(T - i) = \mathbf{H}$ there exists $\psi \in D(T)$ such that $(T^* - i)\phi = (T - i)\psi$. Since $T \subseteq T^*$ this implies $\phi - \psi \in D(T^*)$ and

$$(T^* - i)(\phi - \psi) = 0.$$

Since $\text{Ran}(T + i) = \mathbf{H}$, $\text{Ker}(T^* - i) = 0$ so $\phi = \psi \in D(T)$, as was to be shown. \square

Exercise - Let X be a measure space, μ a measure on X , and $\mathbf{H} = L^2(X, \mu)$. Let f be a real-valued measurable function on X and let M_f be the operator with domain equal to $D = \{\psi \in \mathbf{H}: f\psi \in \mathbf{H}\}$ and for $\psi \in D$ let $M_f\psi = f\psi$. M_f is called a **multiplication operator**. Show that: 1) M_f is densely defined, 2) M_f is symmetric, and 3) M_f is self-adjoint. One can prove 3) directly from the definitions or using the criterion above.

Spectral Theorem - measurable functional calculus. Suppose that A is a self-adjoint operator on \mathbf{H} . Then for each $\psi \in \mathbf{H}$ there is a **spectral measure** $\mu_{A, \psi}$, and for each $f \in \mathcal{B}(\mathbb{R})$ there is a unique bounded operator $\pi(f) = f(A)$ such that

$$\langle \psi, f(A)\psi \rangle = \int_{\mathbb{R}} f(\lambda) d\mu_{A, \psi}(\lambda)$$

and:

1. π is a $*$ -homomorphism,
2. $\|f(A)\| = \|f\|'_\infty$, where the L^∞' norm is defined with the modified notion of ‘almost everywhere’ in which we say a property holds a.e.’ if for each $\psi \in \mathbf{H}$ the property holds a.e. except for a set of measure zero relative to $\mu_{A,\psi}$.
3. If $f_n \in \mathcal{B}(\mathbb{R})$ are functions such that $f_n(\lambda) \rightarrow \lambda$ pointwise and $|f_n(\lambda)| \leq |\lambda|$, then $f_n(x)\psi \rightarrow A\psi$ for any $\psi \in D(A)$.
4. If $f \geq 0$ then $f(A) \geq 0$.
5. If $A\psi = \lambda\psi$ then $f(A)\psi = f(\lambda)\psi$.
6. Suppose that $f_n \rightarrow f$ pointwise a.e.’ and $\|f_n\|'_\infty$ is bounded. Then $f_n(A) \rightarrow f(A)$ strongly.

Sketch of proof - The idea is to reduce to the bounded case, so we need a function that ‘compresses’ A down to a normal operator without losing any information. For example, the operator consisting of multiplication by x on $L^2(\mathbb{R})$ is self-adjoint and unbounded, and our procedure will ‘compress’ it to the operator consisting of multiplication by $(x+i)(x-i)^{-1}$, which is unitary. Let $\alpha(x) = (x+i)(x-i)^{-1}$. Our idea is to construct $f(A)$ by constructing $(f \circ \alpha^{-1})(\tilde{A})$, where $\tilde{A} = (A+i)(A-i)^{-1}$ will be a well-defined unitary operator, called the *Cayley transform* of A . Composing f with α^{-1} ‘undoes’ the effect of the Cayley transform $A \mapsto \tilde{A}$. We begin with some lemmas:

Lemma 1. *If A is self-adjoint then $\tilde{A} = (A+i)(A-i)^{-1}$ is a well-defined unitary operator on \mathbf{H} .*

Proof - If A is self-adjoint the maps $(A \pm i): D(A) \rightarrow \mathbf{H}$ are 1-1 and onto. Thus the composite $(A+i)(A-i)^{-1}$ is a well-defined 1-1 and onto map from \mathbf{H} to \mathbf{H} , and to show that it is unitary it suffices to show that it is norm-preserving.

We need to show that $\|(A+i)(A-i)^{-1}\phi\| = \|\phi\|$ for all $\phi \in \mathbf{H}$. We can use the criterion for self-adjointness to write $\phi = (A-i)\psi$ for some $\psi \in D(A)$. Thus it suffices to show that $\|(A+i)\psi\| = \|(A-i)\psi\|$ for all $\psi \in D(A)$. This follows from:

$$\|(A \pm i)\psi\|^2 = \|A\psi\|^2 + \|\psi\|^2,$$

shown earlier in the proof of the criterion for self-adjointness. \square

Lemma 2. *Let S^1 denote the unit circle in the complex plane. Given $f \in \mathcal{B}(\mathbb{R})$ define $\tilde{f} \in \mathcal{B}(S^1)$ by: $\tilde{f}(x) = f(\alpha^{-1}(x))$ on the range of α , and $\tilde{f} = 0$ elsewhere. Then the map*

$f \mapsto \tilde{f}$ defines a $*$ -homomorphism from $\mathcal{B}(\mathbb{R})$ to $\mathcal{B}(S^1)$ modulo functions that are zero a.e. ι , where $f = 0$ a.e. ι means that $f = 0$ a.e. with respect to all the measures $\mu_{\tilde{A},\psi}$.

Proof - An easy calculation shows that $\alpha: \mathbb{R} \rightarrow S^1$, so that $\tilde{f} \in \mathcal{B}(S^1)$. The range of α is just $S^1 - \{1\}$. We prove one of the $*$ -homomorphism properties and leave most of the rest as easy exercises:

$$(\tilde{f} + \tilde{g})(x) = f(\alpha^{-1}(x)) + g(\alpha^{-1}(x)) = (f + g)(\alpha^{-1}(x)) = \widetilde{(f + g)}(x).$$

The only tricky point is that we need to show that $\tilde{1}$, that is the function equal to 1 on $S^1 - \{1\}$ and 0 at $\{1\}$, is equal to 1 a.e. ι . It suffices to show that $\mu_{\tilde{A},\psi}(1) = 0$ for all $\psi \in \mathbf{H}$. We use the:

Sublemma. *If T is normal, then $\lambda \in \sigma(T)$ has nonzero measure with respect to some $\mu_{T,\psi}$ if and only if $T\phi = \lambda\phi$ for some nonzero $\phi \in \mathbf{H}$.*

Proof - Suppose $\mu_{T,\psi}(\lambda) \neq 0$. Let χ_λ be the characteristic function of $\{\lambda\}$, and let $\phi = \chi_\lambda(T)\psi$. We claim that $\phi \neq 0$ and $T\phi = \lambda\phi$. First note that by the spectral theorem

$$\begin{aligned} \|\phi\|^2 &= \langle \psi, \chi_\lambda(T)^2 \psi \rangle \\ &= \langle \psi, \chi_\lambda(T) \psi \rangle \\ &= \int \chi_\lambda(\lambda') d\mu_{T,\psi}(\lambda) \\ &= u_{T,\psi}(\{\lambda\}) > 0, \end{aligned}$$

so $\phi \neq 0$. Similarly,

$$\begin{aligned} \|(T - \lambda I)\phi\|^2 &= \langle \psi, \chi_\lambda(T)^2 (T - \lambda)^* (T - \lambda) \psi \rangle \\ &= \int \chi_\lambda(\lambda')^2 |\lambda' - \lambda|^2 d\mu_{T,\psi} \\ &= 0, \end{aligned}$$

so $T\phi = \lambda\phi$.

The converse uses similar ideas and is left as an exercise. \square

By this sublemma it suffices to show that there is no nonzero $\phi \in \mathbf{H}$ with $\tilde{A}\phi = \phi$. Suppose there was, i.e. suppose that

$$\phi = (A + i)(A - i)^{-1}\phi.$$

Let $\psi = (A - i)^{-1}\phi$. Then

$$(A - i)\psi = (A + i)\psi$$

so $\psi = 0$. It follows that $\phi = 0$ because $\text{Ker}(A - i)^{-1} = 0$; this is a contradiction. \square

Note that $\sigma(\tilde{A}) \subseteq S^1$ since \tilde{A} is unitary. Thus \tilde{f} restricts to a bounded Borel function on $\sigma(\tilde{A})$, and given $f \in \mathcal{B}(\mathbb{R})$ we may define $f(A) = \tilde{f}(\tilde{A})$ using the spectral theorem for bounded operators. Most of the statements of the theorem now follow rather simply.

Statement 1 follows from Lemma 2 and the functional calculus for normal operators, e.g.: $(fg)(A) = \widetilde{(fg)}(\tilde{A}) = \tilde{f}\tilde{g}(\tilde{A}) = \tilde{f}(\tilde{A})\tilde{g}(\tilde{A}) = f(A)g(A)$.

We define the spectral measures $\mu_{A,\psi}$ using the formula given in the statement of the theorem. Note that if $\mu = \mu_{A,\psi}$ and $\tilde{\mu} = \mu_{\tilde{A},\psi}$ then unraveling the definitions implies

$$\int f d\mu = \langle \psi, f(A)\psi \rangle = \langle \psi, \tilde{f}(\tilde{A})\psi \rangle = \int \tilde{f} d\tilde{\mu}$$

Thus the μ -measure of a set S equals the $\tilde{\mu}$ -measure of the set $\alpha(S)$, and statement 2 follows from the corresponding statement for bounded operators.

Statements 4-6 also follow easily from the spectral theorem for bounded operators, and we leave them as exercises. Statement 3 is more touchy since the functions $\{f_n\}$ are not uniformly bounded. We need a lemma:

Lemma 3. *Let $g(x) = (x + i)^{-1}$. Then $g(A) = (A + i)^{-1}$.*

Proof - By definition, $g(A) = \tilde{g}(\tilde{A})$, and it is easily computed that $\tilde{g}(x) = \frac{x-1}{2ix}$. Thus by the functional calculus for normal operators:

$$\begin{aligned} g(A) &= \tilde{g}(\tilde{A}) = (\tilde{A} - 1)(2i\tilde{A})^{-1} \\ &= \frac{1}{2i}(1 - \tilde{A}^{-1}) \\ &= \frac{1}{2i}(1 - (A - i)(A + i)^{-1}) \\ &= \frac{1}{2i}((A + i) - (A - i))(A + i)^{-1} \\ &= (A + i)^{-1}; \end{aligned}$$

here as always one must take care to justify all the algebraic manipulations involving unbounded operators, but we leave this as an exercise. \square

Now suppose $f_n \in \mathcal{B}(\mathbb{R})$ are functions such that $f_n(\lambda) \rightarrow \lambda$ pointwise and $|f_n(\lambda)| \leq |\lambda|$. Let g be defined as in the lemma; then the sequence $(f_n + i)g \in \mathcal{B}(\mathbb{R})$ is uniformly bounded and converges to 1 pointwise, so by statement 6 and the lemma we have:

$$\lim_{n \rightarrow \infty} (f_n(A) + i)(A + i)^{-1}\phi = \lim_{n \rightarrow \infty} (f_n(A) + i)g(A)\phi = \phi$$

for all $\phi \in \mathbf{H}$. Given any $\psi \in D(A)$, we may write $\psi = (A + i)^{-1}\phi$, and conclude that $\lim_{n \rightarrow \infty} (f_n(A) + i)\psi = (A + i)\psi$, or

$$\lim_{n \rightarrow \infty} f_n(A)\psi = A\psi,$$

as desired. \square

The criterion for self-adjointness is due to von Neumann. The extension of the spectral theorem to unbounded operators is due to von Neumann, 'Allgemeine Eigenwerttheorie Hermitescher Functionaloperatoren,' *Math. Ann.* **102**,49-131 (1929-1930), and Stone, 'Linear Transformations in Hilbert Spaces and their Applications to Analysis,' *A. M. S. Colloq. Publ.* **15** (1932). Carleman in 1923 had pointed out that the spectral theorem would not work for arbitrary symmetric operators.

The spectral theorem has many uses; our first big application of it will be to prove Stone's theorem in the next section. Note that it fulfills our promise to associate to each state ψ and observable A of a quantum system a probability measure $\mu_{A,\psi}$. We shall now describe some basic consequences of this.

It follows from the theorem above that for all states $\psi \in D(A)$ we have

$$\langle \psi, A\psi \rangle = \int \lambda d\mu_{A,\psi}(\lambda).$$

Since $\mu_{A,\psi}$ represents the probability distribution of values that A can be measured to have in the state ψ , the quantity $\langle \psi, A\psi \rangle$ naturally corresponds to the mean value, or **expectation value**, of the observable A in the state ψ . Note that this only makes sense for $\psi \in D(A)$.

Exercise - A **projection** on a Hilbert space \mathbf{H} is a bounded self-adjoint operator P with $P^2 = P$. (Sometimes one considers projections that are not self-adjoint, but in quantum theory one typically assumes self-adjointness, so we build it into the definition.) Show that there is a 1-1 correspondence between projections and closed subspaces of \mathbf{H} given by associating to each projection its range, or conversely to each closed subspace the orthogonal projection onto that subspace. Given a self-adjoint operator A on \mathbf{H} and a

Borel set $S \subseteq \mathbb{R}$, define the **spectral projection** onto S to be the operator $\chi_S(A)$ given by the functional calculus, where χ_S is the characteristic function of S . Show that the spectral projections are indeed projections and satisfy $\chi_{S \cup T}(A) = \chi_S(A) + \chi_T(A)$ if the sets S and T are disjoint.

Note that $\langle \psi, \chi_S(A)\psi \rangle = \int_S d\mu_{A,\psi}(\lambda)$, so that $\langle \psi, \chi_S(A)\psi \rangle$ corresponds to the probability that the observable A is measured to have a value in the set S . Spectral projections are thus very important for computing probabilities.

Exercise - Show that if \mathbb{R} is the disjoint union of a countable collection of Borel sets S_i that the sum of the probabilities p_i that A will be measured to have a value lying in S_i is one.

This is only the beginning of detailed correspondence between concepts from probability theory and from quantum theory. Indeed quantum theory may be regarded as a ‘noncommutative generalization’ of probability theory - the noncommutativity of operator multiplication gives rise to numerous strange effects such as the ‘uncertainty principle’, which we will discuss later.

6. Stone’s Theorem

Stone’s theorem gives a 1-1 correspondence between one-parameter subgroups of unitary operators on \mathbf{H} and self-adjoint operators on \mathbf{H} , which in quantum mechanics connects symmetries with their associated conserved quantities. To prove it we need the notion of an ‘essentially self-adjoint’ operator.

Recall that a densely defined operator is closable if it has closed extensions, in which case the closure of the operator is its smallest closed extension. The following exercise shows that symmetric operators are always closable.

Exercise - If T is densely defined then T^* is closed. If T is symmetric it is closable, since T^* is a closed extension of T .

A symmetric operator A on a Hilbert space \mathbf{H} is **essentially self-adjoint** if its closure is self-adjoint. The importance of this is that it will be easier to show that operators are es-

essentially self-adjoint than to show that they're self-adjoint, and for some purposes essential self-adjointness is good enough. One has: self-adjoint \Rightarrow essentially self-adjoint \Rightarrow symmetric. There are a lot of useful relationships between these concepts, but for we'll just mention a couple that we'll need for Stone's theorem:

Exercise - If S and T are densely defined and $S \subseteq T$, then $T^* \subseteq S^*$. If A is self-adjoint and B is a symmetric extension of A then $B = A$. (Hint: show $A \subseteq B$ and $B \subseteq B^* \subseteq A^* = A$.)

Exercise - Copy the proof of the basic criterion for self-adjointness to show the following:

Basic Criterion for Essential Self-Adjointness. *Let T be a symmetric operator on a Hilbert space \mathbf{H} . Then the following three statements are equivalent:*

1. T is essentially self-adjoint.
2. $\text{Ker}(T^* \pm i) = \{0\}$.
3. $\text{Ran}(T \pm i)$ are dense in \mathbf{H} .

Stone's Theorem. *Let A be a self-adjoint operator and let $U(t) = e^{itA}$. Then $U(t)$ is a strongly continuous one-parameter unitary group, i. e., for all t $U(t)$ is unitary, $U(t+s) = U(t)U(s)$, and $t_\alpha \rightarrow t$ implies $U_\alpha(t)\psi \rightarrow U(t)\psi$ for all $\psi \in \mathbf{H}$. Moreover*

$$D(A) = \{\psi \in \mathbf{H}: \lim_{t \rightarrow 0} \frac{U(t)\psi - \psi}{t} \text{ exists}\}$$

and for all $\psi \in D(A)$

$$iA\psi = \lim_{t \rightarrow 0} \frac{U(t)\psi - \psi}{t}.$$

Conversely, if $U(t)$ is a strongly continuous one-parameter unitary group, there is a unique self-adjoint operator A such that $U(t) = e^{itA}$. A is called the **infinitesimal generator** of $U(t)$.

Proof - If A is a self-adjoint operator it is an easy exercise using the spectral theorem to show that e^{itA} is a strongly continuous unitary group. If $\psi \in D(A)$ then since $(e^{it\lambda} - 1)/t$

converges pointwise to λ as $t \rightarrow 0$, with $|(e^{it\lambda} - 1)/t| \leq \lambda$ for $|t| \leq 1$, statement 3 of the spectral theorem for unbounded operators implies that

$$iA\psi = \lim_{t \rightarrow 0} \frac{U(t)\psi - \psi}{t}.$$

Defining B by

$$D(B) = \{\psi \in \mathbf{H}: \lim_{t \rightarrow 0} \frac{U(t)\psi - \psi}{t} \text{ exists}\}$$

and

$$iB\psi = \lim_{t \rightarrow 0} \frac{U(t)\psi - \psi}{t},$$

it follows from the above that $A \subseteq B$, and B is symmetric since

$$\begin{aligned} \langle B\psi, \phi \rangle &= \lim_{t \rightarrow 0} \langle (it)^{-1}(U(t) - I)\psi, \phi \rangle \\ &= \lim_{t \rightarrow 0} \langle \psi, -(it)^{-1}(U(-t) - I)\phi \rangle \\ &= \lim_{t \rightarrow 0} \langle \psi, (it)^{-1}(U(t) - I)\phi \rangle \\ &= \langle \psi, B\phi \rangle. \end{aligned}$$

Since B is a symmetric extension of the self-adjoint operator A , the exercise above implies $B = A$. Thus

$$D(A) = \{\psi \in \mathbf{H}: \lim_{t \rightarrow 0} \frac{U(t)\psi - \psi}{t} \text{ exists}\}.$$

Conversely, suppose that $U(t)$ is a strongly continuous one-parameter unitary group. We will obtain an essentially self-adjoint operator by differentiating $U(t)\psi$ for ψ in a domain of ‘smooth’ vectors, and the closure of this operator will be our self-adjoint A .

We use an important trick of Gårding’s to obtain vectors for which $U(t)\psi$ is differentiable. Given $f \in C_0^\infty(\mathbb{R})$ define

$$\phi_f = \int f(t)U(t)\phi dt,$$

where the integral can be taken in the sense of Riemann (one can develop a full-blown theory of Lebesgue integration for Hilbert-space-valued functions, but for continuous functions like we have here the Riemann integral will do). Note that

$$\begin{aligned} \lim_{t \rightarrow 0} \left(\frac{U(t) - I}{t} \right) \phi_f &= \lim_{t \rightarrow 0} \int f(s) \left(\frac{U(t+s) - U(s)}{t} \right) \phi ds \\ &= \lim_{t \rightarrow 0} \int \frac{f(\tau - t) - f(\tau)}{t} U(\tau)\phi d\tau \\ &= - \int f'(\tau)U(\tau)\phi d\tau \\ &= \phi_{-f'}, \end{aligned}$$

leaving as an exercise the justification for passing the limit through the integral!

We define the **Gårding domain** by $D = \{\phi_f : f \in C_0^\infty(\mathbb{R})\}$ and define the operator B on D by

$$B\phi = \lim_{t \rightarrow 0} \left(\frac{U(t) - I}{it} \right) \phi.$$

We claim: 1) D is dense, 2) B is essentially self-adjoint, 3) if $A = \overline{B}$ then $U(t) = e^{itA}$.

For 1), given $\phi \in \mathbf{H}$ choose a sequence $f_n \in C_0^\infty(\mathbb{R})$ with $\text{supp} f_n \subset [-1/n, 1/n]$ and $\int f_n = 1$. Then each ϕ_{f_n} is in D and

$$\begin{aligned} \|\phi_{f_n} - \phi\| &= \left\| \int f_n(t)(U(t)\phi - \phi) dt \right\| \\ &\leq \sup_{t \in [-1/n, 1/n]} \|U(t)\phi - \phi\|; \end{aligned}$$

and since $U(t)$ is strongly continuous it follows that $\phi_{f_n} \rightarrow \phi$, so D is dense.

For 2), first note that B can be proved symmetric in the same way that the operator B appearing earlier in the proof was shown to be symmetric. Thus to show that B is essentially self-adjoint we can use the basic criterion and show that $\text{Ker}(B^* \pm i) = \{0\}$. We treat $(B^* - i)$, the case of $(B^* + i)$ being analogous. Suppose that $\psi \in D(B^*)$ has $B^*\psi = i\psi$. Then for each $\phi \in D(B) = D$ we have

$$\begin{aligned} \frac{d}{dt} \langle U(t)\phi, \psi \rangle &= \langle iBU(t)\phi, \psi \rangle \\ &= -i \langle U(t)\phi, B^*\psi \rangle \\ &= -i \langle U(t)\phi, i\psi \rangle \\ &= \langle U(t)\phi, \psi \rangle, \end{aligned}$$

where in the first line we use the fact that $U(t)\psi \in D$, so that $\frac{d}{dt}U(t)\psi = iBU(t)\psi$. Thus the function $f(t) = \langle U(t)\phi, \psi \rangle$ satisfies $f' = f$, so $f(t) = ce^t$. But $f(t)$ is bounded, since $U(t)$ is unitary, so we must have $c = 0$, and in particular $\langle \phi, \psi \rangle = 0$ for all $\phi \in D$. Since D is dense this means that $\psi = 0$ as desired.

For 3) we compare the two unitary groups $U(t)$ and e^{itA} . Given $\phi \in D$ let $w(t) = U(t)\phi - e^{itA}\phi$. Then $w(t)$ is differentiable and

$$w'(t) = iBU(t)\phi - iAe^{itA}\phi = iAU(t)\phi - iAe^{itA}\phi = iAw(t)$$

since $A|_D = B$. Thus

$$\begin{aligned} \frac{d}{dt} \|w(t)\|^2 &= \langle w'(t), w(t) \rangle + \langle w(t), w'(t) \rangle \\ &= \langle iAw(t), w(t) \rangle + \langle w(t), iAw(t) \rangle \\ &= 0 \end{aligned}$$

so $w(t) = 0$ for all t , since $w(0) = 0$. Thus for all $\phi \in D$ we have $U(t)\phi = e^{itA}\phi$. Since D is dense we must have $U(t) = e^{itA}$ (two bounded operators which agree on a dense set are equal). \square

Note that if $U(t)$ is a strongly continuous one-parameter unitary group with infinitesimal generator A , for all t and all $f \in \mathcal{B}(\mathbb{R})$ we have

$$\begin{aligned} \int f(\lambda) d\mu_{A,\psi}(\lambda) &= \langle \psi, f(A)\psi \rangle \\ &= \langle e^{itA}\psi, f(A)e^{itA}\psi \rangle \\ &= \int f(\lambda) d\mu_{A,U(t)\psi}(\lambda) \end{aligned}$$

so

$$\mu_{A,U(t)\psi} = \mu_{A,\psi}.$$

This makes good on our promise that Stone's theorem would associate to each one-parameter subgroup of symmetries an 'associated conserved quantity' - the infinitesimal generator. In the most basic case $U(t)\psi$ represents the state obtained from ψ by time translation - that is, waiting t units of time. Then the infinitesimal generator of $U(t)$ is called the Hamiltonian, and is often denoted H . The fact that the probability distribution of measured values of H doesn't change with time, i.e. $\mu_{H,U(t)\psi} = \mu_{H,\psi}$, is called **conservation of energy**.

Also note that if $\psi \in D(H)$ then

$$\frac{d}{dt}U(t)\psi = HU(t)\psi.$$

This is called (the abstract) **Schrödinger's equation** and is the basic equation of quantum mechanics, just as Hamilton's equations are fundamental in classical mechanics. Note also how much Schrödinger's equation resembles the modern version of Hamilton's equations:

$$\frac{d}{dt}U(t)\psi = v_H U(t)\psi.$$

This is naturally no coincidence; there is a subject of mathematics called 'geometric quantization' which arose from investigations into the relationship between these two equations.

Note that one can't *do* anything with Schrödinger's equation until one knows what the Hamiltonian H is! In what follows we will describe the Hamiltonians for various important physical systems, for example atoms and molecules. We will begin, however, with simpler

examples designed to get a firmer grip on the meaning of the spectral theorem and Stone's theorem.

7. Self-adjoint Matrices and the Spin $\frac{1}{2}$ -Particle

So far our only examples of self-adjoint operators are multiplication operators. If X is a measure space, f is a real-valued measurable function on X , and M_f is the operator on $L^2(X)$ with domain $D = \{\psi \in L^2(X) : f\psi \in L^2(X)\}$ defined by:

$$M_f\psi = f\psi, \quad \psi \in D$$

then M_f is self-adjoint.

Another sort of example comes from finite-dimensional matrices. If \mathbf{H} is finite-dimensional, a dense subspace must be the whole space so a densely defined operator T on \mathbf{H} must have $D(T) = \mathbf{H}$, and any such operator must be bounded. Thus all self-adjoint operators on finite-dimensional \mathbf{H} are bounded. For any self-adjoint operator A on \mathbf{H} there is a unitary U such that UAU^{-1} is a diagonal matrix with real coefficients, the eigenvalues of A , and $\sigma(A)$ is the set of eigenvalues of A , or the set of numbers along the diagonal of UAU^{-1} . The following exercise is crucial for getting a feeling for spectral measures:

Exercise - Let A be a self-adjoint operator on \mathbf{C}^n and let $\{\psi_i\}_{i=1}^n$ be an orthonormal basis of eigenvectors with $A\psi_i = \lambda_i\psi_i$. Let ψ be the state (= unit vector) with $\psi = \sum_{i=1}^n a_i\psi_i$. Show that the spectral measure $\mu_{A,\psi}$ is given by:

$$\mu_{A,\psi} = \sum |a_i|^2 \delta_{\lambda_i}.$$

The simplest and also most important use of finite-dimensional matrices in quantum theory is in the study of 'spin- $\frac{1}{2}$ ' particles. During 1921-1924, Stern and Gerlach did experiments in which they sent a beam of silver atoms through an inhomogeneous magnetic field. To their surprise it split into two components. The same effect occurs in a simpler context if one uses electrons rather than silver atoms. Goudsmit and Uhlenbeck hypothesized in 1925 that this strange effect was due to an intrinsic angular momentum of the electron, called 'spin'. The electrons are either 'spin up' or 'spin down' (relative to the gradient of the magnetic field). Those with 'spin up' are bent one way by the magnetic

field, and those with ‘spin down’ are bent the other. But this happens *no matter from which angle the magnetic field is applied!* It’s as if one had a spinning tennis ball whose axis of rotation was vertical *in any rotated coordinate system*, which is clearly impossible.

Wolfgang Pauli, Enrico Fermi, and Paul Dirac developed a mathematical theory of spin- $\frac{1}{2}$ particles (which include protons and neutrons as well as electrons), and we present part of it here in a summarized form. In what follows we ignore the position and momentum aspects of the electron’s state and concentrate on the description of its ‘spin’.

The spin of the electron about *any* axis is always measured to be $\pm\frac{1}{2}\hbar$, where $\hbar = h/2\pi$ and h is **Planck’s constant**, a physical constant equal to 6.626196×10^{-27} erg sec in the metric system. This constant is so important that mathematical physicists typically work in a modified system of units in which $\hbar = 1$, and we shall do so from now on.

How can the spin of something *always* be measured to be $\pm\frac{1}{2}$ about *any* axis? This does not occur in classical mechanics, where the angular momenta (of, say, a spinning tennis ball) about the x, y , and z axes form a vector $\vec{J} = (J_1, J_2, J_3)$ whose components can be *arbitrary* real numbers, and transform in the usual way under rotations. We need to specify mathematical descriptions of the states, observables, and symmetry group for the spinning electron. We begin with the states and observables, and discuss the symmetry group later.

In quantum mechanics the states of an electron (ignoring its position and momentum) are described by the unit vectors in a 2-dimensional Hilbert space $\mathbf{H} = \mathbb{C}^2$, and the spin angular momenta about the x, y , and z axes are observables described by 2×2 self-adjoint matrices,

$$J_i = \frac{1}{2}\sigma_i,$$

where the σ_i are the **Pauli matrices**:

$$\begin{aligned}\sigma_1 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \sigma_2 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ \sigma_3 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.\end{aligned}$$

We develop the properties of these matrices in the following exercises:

Exercise - Show that the each of the matrices σ_i has eigenvalues ± 1 (thus the matrices J_i have eigenvalues $\pm\frac{1}{2}$). Determine the eigenvectors of the σ_i . Note that no two σ_i have a (nonzero) eigenvector in common.

Exercise - Show that the operator J defined by

$$J = (J_1^2 + J_2^2 + J_3^2)^{1/2}$$

has $J = \sqrt{3/4}I$. (In general if \vec{J} are the angular momentum operators for a ‘spin- s particle’ one has $J = \sqrt{s(s+1)}I$.)

Exercise - The *commutator* $[S, T]$ of $S, T \in \mathcal{L}(\mathbf{H})$ is defined to be $ST - TS$. Show that if $A, B \in \mathcal{L}(\mathbf{H})$ are self-adjoint then $i[A, B]$ is self-adjoint. Show that

$$\sigma_i^2 = I$$

and

$$\sigma_i \sigma_j = -\sigma_j \sigma_i = i\sigma_k$$

if (i, j, k) is a cyclic permutation of $(1, 2, 3)$, hence $[J_i, J_j] = iJ_k$.

In particular, this implies that real linear combinations of the matrices $i\sigma_1, i\sigma_2, i\sigma_3$ and the identity I form an algebra over \mathbb{R} that is isomorphic to Hamilton’s **quaternions**, which recall is an algebra generated by elements i, j, k satisfying

$$i^2 = j^2 = k^2 = -1,$$

$$ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j.$$

In the late 1800’s Hamilton’s quaternions (invented 1843) were in strong competition with Gibbs’ vector notation as a tool for the mathematics of 3-dimensional space. Vector notation, being more elegant for the problems at hand, eventually won out. But it is now clear that quaternions (now often represented as matrices) are necessary for describing certain features of the geometry of \mathbb{R}^3 . The appropriate generalization to higher dimensions are the ‘Clifford algebras’, known in the four-dimensional case by physicists as ‘Dirac matrices’.

Exercise - Let $\psi \in \mathbf{H}$ be the unit vector (a, b) , $|a|^2 + |b|^2 = 1$. Determine the spectral measures $\mu_{\sigma_i, \psi}$, using the exercise above on spectral measures.

In particular, the ‘spin-up’ and ‘spin-down’ states

$$\uparrow = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \downarrow = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

are the eigenvectors of the operator J_3 :

$$J_3 \uparrow = \frac{1}{2} \uparrow, \quad J_3 \downarrow = -\frac{1}{2} \downarrow,$$

and they correspond to an electron with spin pointing up or down, i.e., rotating counter-clockwise or clockwise along the z axis. An electron in any state can be written as a linear combination or, in physics language, **superposition** of these states:

$$\begin{pmatrix} a \\ b \end{pmatrix} = a\uparrow + b\downarrow,$$

and has the spectral measure

$$\mu_{J_3, \psi} = |a|^2 \delta_{-1/2} + |b|^2 \delta_{1/2}.$$

Thus it has probability $|a|^2$ of being measured to have spin $\frac{1}{2}$ and probability $|b|^2$ of being measured to have spin $-\frac{1}{2}$, *if one measures its angular momentum along the z axis*. If one measures its spin along the x axis, one will measure it to be $\pm\frac{1}{2}$, distributed according to the probability distribution $\mu_{J_1, \psi}$.

Of course there's nothing special about the x, y , and z axes. Let $\vec{v} \in \mathbb{R}^3$ be the unit vector (v_1, v_2, v_3) . Then the self-adjoint operator corresponding to spin in the \vec{v} direction is

$$\vec{v} \cdot \vec{J} = \sum_{i=1}^3 v_i J_i.$$

Now we introduce the symmetry group. Recall that the **special orthogonal group** $SO(n)$ is the group of $n \times n$ real orthogonal matrices with determinant one, and the **special unitary group** $SU(n)$ is the group of $n \times n$ complex unitary matrices with determinant one. $SO(3)$ is the group of rotations of 3-dimensional space, and is the obvious candidate for being the symmetry group of the spinning electron, but in fact the group of symmetries is $SU(2)$, which we shall show is a **double cover** of $SO(3)$. That is, there is an onto, two-to-one homomorphism $\alpha: SU(2) \rightarrow SO(3)$.

Note first that the group $SU(2)$ has a natural strongly continuous unitary representation U on \mathbf{H} given by

$$U(g)\psi = g\psi.$$

(Here as always we give groups of finite-dimensional matrices the topology such that $T_\alpha \rightarrow T$ iff $(T_\alpha)_{ij} \rightarrow T_{ij}$ for all i, j .) Next note that if A is a self-adjoint operator on \mathbf{H} which is **traceless**, i.e. $\text{tr}A = 0$ (where $\text{tr}A = \sum A_{ii}$), then $e^{itA} \in SU(2)$, by the following:

Exercise - If \mathbf{H} is an arbitrary Hilbert space, show that for self-adjoint $A \in \mathcal{L}(\mathbf{H})$,

$$e^{itA} = \sum_{n=0}^{\infty} \frac{(itA)^n}{n!}.$$

Show that for any self-adjoint $A \in \mathcal{L}(\mathbb{C}^n)$,

$$\det e^{itA} = e^{it \operatorname{tr} A}.$$

For any unit vector $\vec{v} \in \mathbb{R}^3$ the one-parameter subgroup $e^{i\theta\vec{v}\cdot\vec{J}}$ ($\theta \in \mathbb{R}$) corresponds to rotations about the \vec{v} axis. In other words, given an electron with state described by ψ , if one rotates it by an angle θ about the \vec{v} axis one obtains an electron with state $e^{i\theta\vec{v}\cdot\vec{J}}\psi$. Note however that

$$e^{i\theta J_3} = \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix},$$

so rotating the state ψ by 2π gives, *not* ψ , but $-\psi$! Rotating by 4π , we get back to the state ψ ! This is one reason why it makes sense to call such particles ‘spin- $\frac{1}{2}$ ’. And it’s a physical fact that if you leave an electron lying on the table and someone sneaks in when you’re not looking and rotates it by 360 degrees, when you come back you can tell - if you have the right apparatus! But if he it turns around twice you can’t.

Here we are glossing over an important subtlety. For every observable A on the Hilbert space \mathbf{H} we have $\mu_{A,\psi} = \mu_{A,-\psi}$, so there is no observable property of an electron in *isolation* that changes upon rotating it 360 degrees. But its interactions with other parts of a larger system may differ.

The spin of a particle can be $0, 1/2, 1, 3/2$, etc. Mathematically these various spins correspond to the different irreducible strongly continuous unitary representations of $SU(2)$. Various mesons such as pions, kaons, etc. are spin-zero; electrons, protons, neutrons, neutrinos and various baryons are spin- $1/2$; and photons are spin-one. Assemblages of particles can act like particles with various spins, thus atomic nuclei may be said to have spin $5/2$ and so on.

We have:

Theorem. *There is an onto, two-to-one homomorphism $\alpha: SU(2) \rightarrow SO(3)$ such that the element $e^{i\theta\vec{v}\cdot\vec{J}}$, where \vec{v} is a unit vector, is mapped onto the matrix representing a rotation of angle θ about the \vec{v} axis.*

Sketch of Proof - We construct α as follows. There is an isomorphism between \mathbb{R}^3 and the 2×2 self-adjoint traceless matrices given by:

$$\vec{v} \mapsto \vec{v} \cdot \vec{\sigma} = \begin{pmatrix} v_3 & v_1 - iv_2 \\ v_1 + iv_2 & v_2 \end{pmatrix}.$$

Note that

$$\|\vec{v}\|^2 = -\det(\vec{v} \cdot \vec{J}).$$

The group $SU(2)$ has a representation on the 2×2 traceless self-adjoint matrices given by

$$(g, A) \mapsto gAg^{-1};$$

we leave it as an exercise to verify this. Via the above 1-1 correspondence this gives rise to a representation of $SU(2)$ on \mathbb{R}^3 , which we call α . Note that

$$\begin{aligned}\|V(g)\vec{v}\|^2 &= -\det g(\vec{v} \cdot \vec{J})g^{-1} \\ &= -\det(\vec{v} \cdot \vec{J}) \\ &= \|\vec{v}\|^2,\end{aligned}$$

so that $\alpha(g)$ is an orthogonal 3×3 matrix.

It's clear that $\alpha(I) = I$, and $\alpha(gh) = \alpha(g)\alpha(h)$ because

$$(gh)A(gh)^{-1} = g(hAh^{-1})g^{-1}.$$

Thus α is a homomorphism from $SU(2)$ to $O(3)$, the group of 3×3 orthogonal matrices. It remains to show that: 1) $Ran(\alpha) = SO(3)$, 2) α is two-to-one, 3) α maps $e^{i\theta\vec{v}\cdot\vec{J}}$ to the rotation of angle θ about \vec{v} .

1) can be shown by a calculation, but it's easier to use a few facts about these groups. First, since $SU(2)$ is connected and α is continuous, α must map $SU(2)$ into the connected component of $O(3)$ that contains the identity. This is just $SO(3)$; the other connected component consists of matrices with determinant -1. Since every element of $SO(3)$ is a rotation of some angle about some axis, if we show 3) it will follow that α is onto $SO(3)$.

2) is equivalent to saying that the kernel of α consists of two elements. Unraveling the definitions, $g \in SU(2)$ is in $Ker \alpha$ iff for all traceless self-adjoint matrices A we have $gAg^{-1} = A$. It is easily seen that the only matrices which commute with all traceless self-adjoint matrices are those of the form cI . The only ones of these which lie in $SU(2)$ are $\pm I$. Thus $Ker \alpha = \pm I$.

To show 3) it suffices to show - differentiating with respect to θ - that:

$$\frac{d}{d\theta}\alpha(e^{i\theta\vec{v}\cdot\vec{J}})\vec{w}|_{\theta=0} = -\vec{v} \times \vec{w}.$$

(Here the minus sign is an unimportant consequence of our sign conventions; the main point is that when one rotates the vector \vec{w} about the \vec{v} axis the derivative is $\pm\vec{v} \times \vec{w}$, with the sign depending on whether one rotates clockwise or counterclockwise.) Unraveling the definitions, this is equivalent to:

$$\frac{d}{d\theta}e^{i\theta\vec{v}\cdot\vec{J}}(\vec{w} \cdot \vec{J})e^{-i\theta\vec{v}\cdot\vec{J}}|_{\theta=0} = -(\vec{v} \times \vec{w}) \cdot \vec{J}.$$

Doing the differentiation on the left side, it suffices to show:

$$[i\vec{v} \cdot \vec{J}, \vec{w} \cdot \vec{J}] = -(\vec{v} \times \vec{w}) \cdot \vec{J}.$$

This follows from the result of an exercise above, that

$$[\sigma_i, \sigma_j] = 2i\sigma_k .$$

□

Note also that by general principles the operator $\vec{v} \cdot \vec{J}$ is the conserved quantity corresponding to the one-parameter group $e^{i\theta\vec{v} \cdot \vec{J}}$. This means that the probability distribution of measured spins in the \vec{v} direction doesn't change if you rotate the electron by a fixed angle about the \vec{v} axis - which makes sense.

Rather than meditating further on the oddity of the quantum theory of spin- $\frac{1}{2}$ particles, let us turn to infinite-dimensional diagonal matrices. Recall that ℓ^2 is the usual Hilbert space of complex sequences $\psi = (\psi_i)_{i=1}^{\infty}$ such that $\|\psi\|^2 = \sum |\psi_i|^2$ is finite, with the inner product $\langle \psi, \phi \rangle = \sum \bar{\psi}_i \phi_i$.

Theorem. Let $(a_i)_{i=1}^{\infty}$ be a sequence of real numbers. Let A be the operator on ℓ^2 with domain

$$D(A) = \{\psi \in \ell^2 : \sum |a_i \psi_i|^2 < \infty\}$$

given on this domain by

$$(A\psi)_i = a_i \psi_i .$$

Then A is self-adjoint. The restriction of A to the domain $D_0 = \{\psi \in \ell^2 : \text{only finitely many } \psi_i \text{ are nonzero}\}$ is essentially self-adjoint.

Proof - We can think of ℓ^2 as $L^2(N)$, where $N = 1, 2, 3, \dots$ with the 'counting measure' that assigns to each point the measure 1. Then A is a multiplication operator and is self-adjoint by the general theory of such. To show that $A_0 = A|_{D_0}$ is essentially self-adjoint, we use the basic criterion: it is easily seen that A_0 is symmetric, and to show that $\text{Ran}(A \pm i)$ are dense in \mathbf{H} it suffices to show that they contain D_0 , which is dense. Suppose $\psi \in D_0$, and define ϕ by $\phi_i = (a_i \pm i)^{-1} \psi_i$. Then $\phi \in D_0$, and $(A_0 \pm i)\phi = \psi$. Thus $\text{Ran}(A \pm i) \supseteq D_0$.
□

Exercise - Show that if A is the self-adjoint operator on ℓ^2 given above and $\psi \in \ell^2$ then $\mu_{A,\psi} = \sum |\psi_i|^2 \delta_{a_i}$.

The spectrum of an unbounded self-adjoint operator may be defined in more or less the same way as for bounded operators. We give two definitions and prove that they are equivalent, after giving an exercise used in the proof:

Exercise - Show that if T is a densely defined operator on \mathbf{H} then $(T - \lambda I)^* = T^* - \bar{\lambda}I$, where $T - \lambda I$ is the operator with domain $D(T)$, defined in the obvious way.

Theorem. *Let A be a self-adjoint operator on a Hilbert space \mathbf{H} . Then the following are equivalent:*

1) λ is in the closure of $\bigcup \text{supp} \mu_{A,\psi}$.

2) There does not exist $R \in \mathcal{L}(\mathbf{H})$ such that $R(A - \lambda I)\psi = \psi$ for all $\psi \in D(A)$ and $(A - \lambda I)R\psi = \psi$ for all $\psi \in \mathbf{H}$. (Note that for the latter equation to make sense we need to have $\text{Ran}(R) \subseteq D(A)$.)

If either of these holds we say λ is in the **spectrum** of A , denoted $\sigma(A)$.

Proof - 1) \Rightarrow 2): Suppose 1) holds. Then given any $\epsilon > 0$ there exists $\psi \in \mathbf{H}$ such that λ is distance $\leq \epsilon$ from $\text{supp} \mu_{A,\psi}$. This implies that the intersection of the interval $I = [\lambda - \epsilon, \lambda + \epsilon]$ with $\text{supp} \mu_{A,\psi}$ is nonzero, so by the spectral theorem

$$\begin{aligned} \|\chi_I(A)\psi\|^2 &= \langle \psi, \chi_I(A)\psi \rangle \\ &= \int_I \mu_{a,\psi} > 0. \end{aligned}$$

Let $\phi = \chi_I(A)\psi$; by the above $\phi \neq 0$, and clearly $\phi = \chi_I(A)\phi$. It is easily seen using the spectral theorem that $\phi \in D(A)$. Moreover

$$\begin{aligned} \|(A - \lambda)\phi\|^2 &= \langle \phi, (A - \lambda)^2 \chi_I(A)\phi \rangle \\ &= \int_I (x - \lambda)^2 d\mu_{A,\phi}(\lambda) \\ &\leq \epsilon^2 \int_I d\mu_{A,\phi} = \epsilon^2 \|\phi\|^2. \end{aligned}$$

Thus for any $\epsilon > 0$ there exists nonzero $\phi \in D(A)$ such that $\|(A - \lambda I)\phi\| \leq \epsilon \|\phi\|$. It is an easy exercise to show that no bounded R can have $R(A - \lambda I)\phi = \phi$ for all such ϕ .

2) \Rightarrow 1): If 1) does not hold, then the function $(x - \lambda)^{-1}$ has finite L^∞ norm, and we can define R_λ by the functional calculus to be $(A - \lambda I)^{-1}$. Note that by the functional calculus we have $R_{\bar{\lambda}} = R_\lambda^*$. If we let f_n be a sequence of functions with $f_n(x) \rightarrow x$ for

all $x \in \mathbb{R}$ and $|f_n(x)| \leq x$, then by the functional calculus we have $f_n(A)\psi \rightarrow A\psi$ for all $\psi \in D(A)$, and $R_\lambda(f_n(A) - \lambda I) \rightarrow I$ strongly, hence for all $\psi \in D(A)$

$$\begin{aligned} R_\lambda(A - \lambda I)\psi &= R_\lambda \lim_{n \rightarrow \infty} (f_n(A) - \lambda I)\psi \\ &= \lim_{n \rightarrow \infty} R_\lambda(f_n(A) - \lambda I)\psi \\ &= \psi. \end{aligned}$$

Furthermore, if $\psi \in \mathbf{H}$, for all $\phi \in D(A - \lambda I)$ we have

$$\begin{aligned} \langle \psi, \phi \rangle &= \langle \psi, R_{\bar{\lambda}}(A - \bar{\lambda}I)\phi \rangle \\ &= \langle R_\lambda \psi, (A - \bar{\lambda}I)\phi \rangle. \end{aligned}$$

hence by the definition of adjoint, $R_\lambda \psi \in D((A - \bar{\lambda}I)^*)$, and

$$(A - \lambda I)R_\lambda \psi = (A - \bar{\lambda}I)^* R_\lambda \psi = \psi.$$

Thus 2) does not hold. \square

We apply this to infinite-dimensional diagonal matrices in the following exercises:

Exercise - Show that a self-adjoint operator A on \mathbf{H} is bounded if and only if $\sigma(A)$ is a bounded set, in which case $\|A\| = r(A)$.

Exercise - If A is the self-adjoint operator on ℓ^2 given by

$$(A\psi)_i = a_i \psi_i,$$

show that $\sigma(A) = \{a_i\}$. Show that A is bounded if and only if a_i is a bounded sequence, in which case $\|A\| = \sup |a_i|$.

The theorems above have many applications to differential equations. For example:

Exercise - Show that there is an orthonormal basis of $L^2[0, 1]$ consisting of $\psi_n \in C^\infty[0, 1]$ such that

$$-\frac{d^2}{dx^2}\psi_n = \lambda_n \psi_n$$

satisfying the **Dirichlet boundary conditions**:

$$\psi(0) = \psi(1) = 0.$$

Let D be the space of finite linear combinations of the ψ_n . Show that $-\frac{d^2}{dx^2}$ is essentially self-adjoint on D , and let Δ_D be the closure of this operator. Determine the spectrum of Δ_D .

Exercise - Show that there is an orthonormal basis of $L^2[0, 1]$ consisting of $\psi_n \in C^\infty[0, 1]$ such that

$$-\frac{d^2}{dx^2}\psi_n = \lambda_n\psi_n$$

satisfying the **Neumann boundary conditions**:

$$\psi'(0) = \psi'(1) = 0.$$

Let N be the space of finite linear combinations of the ψ_n . Show that $-\frac{d^2}{dx^2}$ is essentially self-adjoint on N , and let Δ_N be the closure of this operator. Determine the spectrum of Δ_N .

One can then use the functional calculus to define operators such as $e^{-t\Delta}$ (the ‘heat kernel’), which are very useful in mathematics and physics.

8. Translation and Differentiation

Another simple and important self-adjoint operator arises from translation:

Theorem. Let $U(t): L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ defined for $t \in \mathbb{R}$ by

$$U(t)\psi(x) = \psi(x + t).$$

Then $U(t)$ is a strongly continuous one-parameter unitary group, and the infinitesimal generator p of $U(t)$ is a self-adjoint operator such that for all $\psi \in D(p)$,

$$p\psi = i^{-1} \frac{d\psi}{dx}$$

where the derivative is taken in the distributional sense.

Proof - Clearly $U(t)$ is a 1-parameter unitary group. Note that to show that a unitary group is strongly continuous it suffices to show that for all $\psi \in H$, $U(t)\psi \rightarrow \psi$ as $t \rightarrow 0$.

Given $\psi \in L^2(\mathbb{R})$, for all $\epsilon > 0$ there exists $\phi \in C_0^\infty(\mathbb{R})$ such that $\|\psi - \phi\| \leq \epsilon$. Let $M^2 = \int_{\mathbb{R}} |d\phi/dx|^2$. Then

$$\begin{aligned} \|U(t)\psi - \psi\| &\leq \|U(t)\psi - U(t)\phi\| + \|U(t)\phi - \phi\| + \|\phi - \psi\| \\ &\leq 2\epsilon + \left(\int |\phi(x+t) - \phi(x)|^2 dx \right)^{\frac{1}{2}} \\ &\leq 2\epsilon + t|M| \end{aligned}$$

so if $|t|M \leq \epsilon$ then

$$\|U(t)\psi - \psi\| \leq 3\epsilon.$$

Thus $U(t)$ is strongly continuous.

By Stone's theorem $U(t) = e^{itp}$ for some self-adjoint p and

$$D(p) = \{ \psi \in L^2(\mathbb{R}) : \lim_{t \rightarrow 0} \frac{U(t)\psi - \psi}{t} \text{ exists in } L^2(\mathbb{R}) \}$$

and

$$ip\psi = \lim_{t \rightarrow 0} \frac{U(t)\psi - \psi}{t}.$$

Thus we only need to show that if the limit

$$\lim_{t \rightarrow 0} t^{-1}(U(t)\psi - \psi)$$

exists in $L^2(\mathbb{R})$ then it equals the distributional derivative $d\psi/dx$.

Suppose the limit exists in $L^2(\mathbb{R})$. Then for all test functions $\phi \in C_0^\infty(\mathbb{R})$

$$\begin{aligned} \int \frac{d\bar{\psi}}{dx} \phi dx &= - \int \bar{\psi} \frac{d\phi}{dx} dx \\ &= \lim_{t \rightarrow 0} \int t^{-1} \bar{\psi}(x) (\phi(x-t) - \phi(x)) dx \\ &= \lim_{t \rightarrow 0} \int t^{-1} (\bar{\psi}(x+t) - \bar{\psi}(x)) \phi(x) dx \\ &= \lim_{t \rightarrow 0} \langle t^{-1} (U(t)\psi - \psi), \phi \rangle \\ &= \langle ip\psi, \phi \rangle, \end{aligned}$$

so the distributional derivative $\frac{d\psi}{dx}$ equals $ip\psi$. \square

Translation is of course an important physical symmetry, and in the context of physics the infinitesimal generator p is called 'momentum'. Of course real physics happens in \mathbb{R}^3 ,

not \mathbb{R} . Ignoring spin now, the state of a point particle in \mathbb{R}^3 is described by a unit vector in the Hilbert space $\mathbf{H} = L^2(\mathbb{R}^3)$, and the group \mathbb{R}^3 of translations has a strongly continuous unitary representation on \mathbf{H} given by

$$(U(g)\psi)(x) = \psi(x + g), \quad g \in \mathbb{R}^3, \quad \psi \in \mathbf{H}.$$

(The proof of the latter is exactly analogous to that of the theorem above.) The infinitesimal generators of translation in the x , y , and z directions respectively are self-adjoint operators, the x , y , and z **momenta**, denoted p_1, p_2, p_3 .

We can also obtain the momentum operators, as well as other self-adjoint differential operators, in terms of the Fourier transform. Recall that the Fourier transform is a unitary operator on $L^2(\mathbb{R}^n)$ given by

$$\widehat{f}(\vec{k}) = (2\pi)^{-n/2} \int f(\vec{x}) e^{-i\vec{k}\cdot\vec{x}} d^n x,$$

where the integral is absolutely convergent for $f \in L^2(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$ but is defined by a limit for arbitrary $f \in L^2$.

Theorem. *Let p be the self-adjoint generator of translations on $L^2(\mathbb{R})$. Then the domain of p is equal to $\{f \in L^2(\mathbb{R}) : k\widehat{f} \in L^2(\mathbb{R})\}$.*

Proof - Let $U(t)$ be the one-parameter group of translations on $\mathbf{H} = L^2(\mathbb{R})$, and let $F: \mathbf{H} \rightarrow \mathbf{H}$ be the Fourier transform. We have

$$\begin{aligned} (U(t)\widehat{f})(k) &= (2\pi)^{-1/2} \int f(x+t) e^{-ikx} dx \\ &= (2\pi)^{-1/2} \int f(x) e^{-ik(x-t)} dx \\ &= e^{ikt} \widehat{f}(k), \end{aligned}$$

or, in short,

$$FU(t) = M_{e^{ikt}} F,$$

or

$$U(t) = F^{-1} M_{e^{ikt}} F,$$

where $M_{e^{ikt}}$ denotes the operator of multiplication by e^{ikt} (which is a unitary operator on \mathbf{H}). Thus $\lim_{t \rightarrow 0} t^{-1}(U(t) - I)\psi$ exists in \mathbf{H} if and only if $\lim_{t \rightarrow 0} t^{-1}(e^{ikt} - I)\widehat{\psi}$ exists in \mathbf{H} , which is easily seen to hold if and only if $k\widehat{\psi} \in \mathbf{H}$. \square

Exercise - As a spot-check on your real analysis skills, prove the ‘easily seen’ claim above.

Exercise - Determine the spectrum of the operator p . Find approximate eigenvectors.

We can use the following theorem to show that $i^{-1}\frac{d}{dx}$ is essentially self-adjoint on $C_0^\infty(\mathbb{R})$. More precisely, if A is a self-adjoint operator on \mathbf{H} we say $D \subseteq D(A)$ is a **domain of essential self-adjointness** for A or a **core** for A if the closure of $A|_D$ is A . (In particular this implies that $A|_D$ is essentially self-adjoint.)

Theorem. Let $e^{itA} = U(t)$ be a strongly continuous 1-parameter unitary group on \mathbf{H} and suppose $D \subset \mathbf{H}$ be a dense space of **differentiable vectors** for $U(t)$, i.e., if $\psi \in D$ then the limit

$$\lim_{t \rightarrow 0} \frac{U(t)\psi - \psi}{t}$$

exists in \mathbf{H} . Suppose also that D is **invariant under** $U(t)$, i.e., $U(t)D \subseteq D$ for all t . Then D is a core for A .

Proof - The proof is much like part of that of Stone’s theorem. Note that by Stone’s theorem $D \subseteq D(A)$. Next note that $B = A|_D$ is symmetric since $\langle A\psi, \phi \rangle = \langle \psi, A\phi \rangle$ for all $\psi, \phi \in D(A)$, hence for all $\psi, \phi \in D$.

Next we show that B is essentially self-adjoint by the basic criterion, by showing that $\text{Ker}(B^* \pm i) = \{0\}$. We treat $(B^* - i)$, the case of $(B^* + i)$ being analogous. Suppose that $\psi \in D(B^*)$ has $B^*\psi = i\psi$. Then for each $\phi \in D(B) = D$ we have

$$\begin{aligned} \frac{d}{dt} \langle U(t)\phi, \psi \rangle &= \langle iBU(t)\phi, \psi \rangle \\ &= -i \langle U(t)\phi, B^*\psi \rangle \\ &= -i \langle U(t)\phi, i\psi \rangle \\ &= \langle U(t)\phi, \psi \rangle, \end{aligned}$$

where in the first line we use the fact that $U(t)\psi \in D$, so that $\frac{d}{dt}U(t)\psi = iBU(t)\psi$. Thus the function $f(t) = \langle U(t)\phi, \psi \rangle$ satisfies $f' = f$, so $f(t) = ce^t$. But $f(t)$ is bounded, since $U(t)$ is unitary, so we must have $c = 0$, and in particular $\langle \phi, \psi \rangle = 0$ for all $\phi \in D$. Since D is dense this means that $\psi = 0$ as desired.

Thus we have the following situation: A is self-adjoint, $B \subseteq A$ and B is essentially self-adjoint. We’ll be done if we show the closure of B is A . This follows from:

Lemma. Suppose A is a self-adjoint operator, B is an essentially self-adjoint operator and $B \subseteq A$. Then $\overline{B} = A$.

Proof - Since A is self-adjoint it is closed, and it is easily seen that in general

$$S \subseteq T \implies \overline{S} \subseteq \overline{T},$$

so $\overline{B} \subseteq \overline{A} = A$. To show the opposite inclusion, note that $\overline{B} \subseteq A$ implies $A^* \subseteq \overline{B}^*$, but A and \overline{B} are self-adjoint, so $A \subseteq \overline{B}$. \square

Corollary. *The space $C_0^\infty(\mathbb{R})$ is a core for the operator p .*

Proof - This follows immediately from the theorem. \square

Differentiation on the unit interval $[0, 1]$ is a more subtle matter. Let $\mathbf{H} = L^2[0, 1]$. By Stone's theorem we know any self-adjoint operator corresponding to $i^{-1} \frac{d}{dx}$ has to generate a one-parameter group of unitaries, and by analogy with the \mathbb{R} case we expect this group to consist of translations, but if one translates a function on $[0, 1]$ it will 'fall off the edge'!

Suppose one considers, by analogy with the case of \mathbb{R} , an operator B by:

$$D(B) = \{\psi \in \mathbf{H}: \psi \in C^\infty[0, 1], \phi(0) = \phi(1) = 0\}$$

$$B\psi = i^{-1} \frac{d\psi}{dx}.$$

Is B essentially self-adjoint? An integration by parts shows that it is symmetric:

$$\begin{aligned} \langle \psi, B\phi \rangle &= \int_0^1 \overline{\psi(x)} i^{-1} \phi'(x) dx \\ &= \int_0^1 \overline{i^{-1} \psi'(x)} \phi(x) dx \\ &= \langle B\psi, \phi \rangle; \end{aligned}$$

note that no boundary terms appear in the integration by parts since $\phi(0) = \phi(1) = 0$. Unfortunately B is not essentially self-adjoint because the kernels $\text{Ker}(B^* \pm i)$ are nonzero.

Consider the case of $B^* + i$ (the other is similar). We claim that the function $e^x \in \mathbf{H}$ is in $D(B^*)$ and $B^* e^x = -ie^x$, which implies $e^x \in \text{Ker}(B^* + i)$. To prove our claim it suffices to show that

$$\langle e^x, B\phi \rangle = \langle -ie^x, \phi \rangle$$

for all $\phi \in D(B)$. This follows from integration by parts:

$$\begin{aligned}\langle e^x, B\phi \rangle &= \int_0^1 e^x i^{-1} \phi'(x) dx \\ &= \int_0^1 \overline{i^{-1} e^x} \phi(x) dx \\ &= \langle -ie^x, \phi \rangle.\end{aligned}$$

The moral here is that the domain $D(B)$ is too small, so the function e^x is orthogonal to $(B - i)\phi$ for all $\phi \in \mathbf{H}$, giving $B^* + i$ a nonzero kernel.

So why not try a bigger domain? Define B instead by:

$$D(B) = \{\psi \in \mathbf{H}: \psi \in C^\infty[0, 1]\}$$

$$B\psi = i^{-1} \frac{d\psi}{dx}.$$

This domain is unfortunately too big: now B isn't symmetric! For if $\psi, \phi \in D(B)$, integration by parts gives:

$$\begin{aligned}\langle \psi, B\phi \rangle &= \int_0^1 \overline{\psi(x)} i^{-1} \phi'(x) dx \\ &= \int_0^1 \overline{i^{-1} \psi(x)} \phi(x) dx + \overline{\psi} \phi \Big|_0^1 \\ &\neq \langle B\psi, \phi \rangle.\end{aligned}$$

Don't despair, however! Suppose we define B by

$$D(B) = \{\psi \in \mathbf{H}: \psi \in C^\infty[0, 1], \psi(1) = \alpha\psi(0)\},$$

$$B\psi = i^{-1} \frac{d\psi}{dx},$$

where α is a complex number with $|\alpha| = 1$. Now B is symmetric because, integrating by parts:

$$\begin{aligned}\langle \psi, B\phi \rangle &= \int_0^1 \overline{\psi(x)} i^{-1} \phi'(x) dx \\ &= \int_0^1 \overline{i^{-1} \psi(x)} \phi(x) dx + \overline{\psi} \phi \Big|_0^1 \\ &= \langle B\psi, \phi \rangle,\end{aligned}$$

since

$$\overline{\psi(1)} \phi(1) = |\alpha|^2 \overline{\psi(0)} \phi(0) = \overline{\psi(0)} \phi(0).$$

Let θ be any real number with $e^{i\theta} = \alpha$. Note that the functions ψ_n defined by

$$\psi_n(x) = e^{i(2\pi n + \theta)x}$$

are in $D(B)$, and

$$B\psi_n = (2\pi n + \theta)\psi_n.$$

By Fourier theory it's easy to see that the functions ψ_n form an orthonormal basis of \mathbf{H} , so finite linear combinations of the ψ_n , which lie in $D(B)$, are dense in \mathbf{H} . In particular, given such a combination, say $\sum a_n \psi_n$, we can write it as

$$\sum a_n \psi_n = (B \pm i) \left(\sum a_n (2\pi n + \theta \pm i)^{-1} \psi_n \right),$$

where the sum on the right side is in $D(B)$, so we see that $Ran(B \pm i)$ are dense. Thus B is essentially self-adjoint. As one would expect from the above calculations, we have:

Exercise - \overline{B} is a self-adjoint operator with spectrum equal to $\{2\pi n + \theta\}_{-\infty < n < \infty}$.

Thus we see that in this example the expression $i^{-1} \frac{d}{dx}$ may be given many different interpretations as a self-adjoint operator, depending on what domain we choose, and that the spectrum depends on the domain we choose! Note also that the choice of domain is closely related to boundary conditions. This is typical in elliptic operator theory; an elliptic operator on a bounded open set is self-adjoint if we choose the right domain, but its spectrum depends on the domain.

Exercise - Describe the unitary group $e^{it\overline{B}}$ generated by the operator \overline{B} above, and explain how it depends on α . Hint: evaluate $e^{it\overline{B}}\psi_n$.

9. The Canonical Commutation Relations

Ignoring spin, the state of a particle in \mathbb{R}^3 is described by a unit vector in $L^2(\mathbb{R}^3)$. The function ψ on \mathbb{R}^3 is often called the particle's **wave function**. As mentioned above, the momenta are the observables p_1, p_2, p_3 generating translations in the x, y , and z directions. Other fundamental observables are the **position** operators, that is, multiplication operators by the coordinate functions x, y , and z . These self-adjoint operators are denoted q_1, q_2 , and q_3 . (The generalization to \mathbb{R}^n should be obvious.)

Recall that $\psi \in L^2(\mathbb{R}^n)$ is said to be in the **Schwartz space** $\mathcal{S}(\mathbb{R}^n)$ if for all $n, m \geq 0$ we have

$$(1 + r^2)^n (1 + \Delta)^m \psi \in L^2(\mathbb{R}^n).$$

where

$$\Delta = - \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}$$

is taken in the distributional sense. Another equivalent definition is commonly used:

Exercise - Show that $\psi \in \mathcal{S}(\mathbb{R}^n)$ if and only if for all $n, m \geq 0$ we have

$$(1 + r^2)^n (1 + \Delta)^m \psi \in L^\infty(\mathbb{R}^n).$$

Much of the importance of the Schwartz space in operator theory comes from the following facts:

Exercise - Show that if $\psi \in \mathcal{S}(\mathbb{R}^n)$ then $p_i \psi \in \mathcal{S}(\mathbb{R}^n)$ and $q_i \psi \in \mathcal{S}(\mathbb{R}^n)$ for all i , $1 \leq i \leq n$.

Exercise - Show that $\mathcal{S}(\mathbb{R}^n)$ is a core for the operators p_i, q_i . (Hint: use the theorem on cores in the previous section.)

Exercise - Let F denote the Fourier transform as a unitary operator on $L^2(\mathbb{R}^n)$. Show that F maps $\mathcal{S}(\mathbb{R}^n)$ to itself in a 1-1 and onto manner.

Suppose that $\psi \in \mathcal{S}(\mathbb{R}^n)$. Then note that for all i, j ,

$$\begin{aligned} (p_i q_j - q_j p_i) \psi &= i^{-1} \left(\frac{\partial}{\partial x_i} x_j \psi - x_j \frac{\partial}{\partial x_i} \psi \right) \\ &= i^{-1} \frac{\partial x_j}{\partial x_i} \psi \\ &= i^{-1} \delta_{ij} \psi. \end{aligned}$$

We also have

$$(p_i p_j - p_j p_i) \psi = 0, \quad (q_i q_j - q_j q_i) \psi = 0.$$

These relations, often summarized as

$$[p_i, q_j] = i^{-1} \delta_{ij}, \quad [p_i, p_j] = [q_i, q_j] = 0,$$

are called the **canonical commutation relations**. They were discovered by Werner Heisenberg's supervisor, Max Born, while studying a paper of Heisenberg's. At the time (1925) Born was one of the few physicists who had a good knowledge of matrices. He interpreted some quantum-mechanical calculations of Heisenberg's in terms of infinite-dimensional matrices, explaining in this way the odd non-commutativity of multiplication

that Heisenberg had gotten himself into. This made physicists learn about matrices and develop the new ‘matrix mechanics’, as an alternative to Schrödinger’s ‘wave mechanics’. Only later were Heisenberg’s position and momentum operators correlated by Pauli to the operators of multiplication and differentiation by coordinates x_i , in which form they had appeared in Schrödinger’s equation for the hydrogen atom (more on which later). Heisenberg had, in effect, been describing differential operators in terms of an orthonormal basis for $L^2(\mathbb{R}^n)$.

10. The Uncertainty Principle

We begin by pointing out how to multiply and add densely defined operators. Given two densely defined operators A and B , we define the product AB by:

$$D(AB) = \{\psi \in D(B) : B\psi \in D(A)\},$$

$$AB\psi = A(B\psi).$$

It is not necessarily the case that $D(AB)$ is dense! This needs to be proved in each case; for example, properties of the Schwartz space make it clear that products of position and momentum operators such as pq , p^2 , qp^2q and so on are all densely defined operators.

Similarly we define the sum $A + B$ by:

$$D(A + B) = D(A) \cap D(B),$$

$$(A + B)\psi = A\psi + B\psi.$$

Again it needs to be proved that $D(A+B)$ is dense - two dense subspaces of a Hilbert space can have 0 as their intersection! The properties of the Schwartz space imply that all polynomials in p and q are densely defined operators.

Exercise - Show that if A is a self-adjoint operator then any polynomial in A is densely defined, and any real polynomial in A is essentially self-adjoint.

Exercise - Define \tilde{q} by

$$D(\tilde{q}) = C_0^\infty(\mathbb{R}^n),$$

$$(\tilde{q}\psi)(x) = x\psi(x).$$

Define \tilde{p} by

$$D(\tilde{p}) = \{\psi \in L^2(\mathbb{R}): \widehat{\psi} \in C_0^\infty(\mathbb{R})\}$$

$$(\tilde{p}\psi)(x) = i^{-1} \frac{d}{dx} \psi(x).$$

Show that \tilde{p} and \tilde{q} are essentially self-adjoint and have as closures the usual p and q . Show however that

$$D(\tilde{p}) \cap D(\tilde{q}) = \{0\}$$

so $\tilde{p} + \tilde{q}$ has as domain only 0! (Hint: for the last part use the fact that if $\widehat{\psi} \in L^2(\mathbb{R})$ is compactly supported then ψ is analytic, which follows from the Paley-Wiener theorem.)

Exercise - Show that the operator $\Delta = p_1^2 + p_2^2 + p_3^2$ is essentially self-adjoint. (Hint: show using the Fourier transform that it's even essentially self-adjoint when restricted to the smaller domain $\mathcal{S}(\mathbb{R}^3)$.)

Given an observable A on \mathbf{H} and a state $\psi \in D(A^2)$, we can define the **standard deviation** of A in the state ψ , usually written ΔA even though it depends on ψ , by

$$(\Delta A)^2 = \langle \psi, A^2 \psi \rangle - \langle \psi, A \psi \rangle^2.$$

Note that this corresponds to the usual definition of standard deviation because $\langle \psi, A \psi \rangle$ is the mean value of the observable A in the state ψ , as described in section 5.

The **uncertainty principle** gives a lower bound on the product $\Delta A \Delta B$ for non-commuting observables A and B :

Uncertainty Principle. *Let A and B be self-adjoint operators on the Hilbert space \mathbf{H} , and suppose $\psi \in D(A^2) \cap D(B^2) \cap D(AB) \cap D(BA)$ is a unit vector. Then*

$$\Delta A \Delta B \geq \frac{1}{2} |\langle \psi, [A, B] \psi \rangle|.$$

Proof - First we note that it suffices to show this for

$$A' = A - \langle \psi, A \psi \rangle$$

and

$$B' = B - \langle \psi, B \psi \rangle$$

instead. This is because $[A', B'] = [A, B]$ and

$$\langle \psi, A' \psi \rangle = \langle \psi, A \psi \rangle - \langle \psi, \langle \psi, A \psi \rangle \psi \rangle = 0,$$

hence

$$\begin{aligned}
(\Delta A')^2 &= \langle \psi, A'^2 \psi \rangle \\
&= \langle \psi, (A^2 - 2\langle \psi, A \psi \rangle A + \langle \psi, A \psi \rangle^2) \psi \rangle \\
&= \langle \psi, A^2 \psi \rangle - \langle \psi, A \psi \rangle^2 \\
&= (\Delta A)^2,
\end{aligned}$$

and similarly for B .

We have

$$\begin{aligned}
|\langle \psi, [A', B'] \psi \rangle| &= |\langle A' \psi, B' \psi \rangle - \langle B' \psi, A' \psi \rangle| \\
&= 2 |\operatorname{Im} \langle A' \psi, B' \psi \rangle| \\
&\leq 2 |\langle A' \psi, B' \psi \rangle| \\
&\leq 2 \|A' \psi\| \|B' \psi\|,
\end{aligned}$$

the last step using Cauchy-Schwarz; but

$$\|A' \psi\| = \langle \psi, A'^2 \psi \rangle^{1/2} = \Delta A'$$

and similarly for B' , so

$$|\langle \psi, [A', B'] \psi \rangle| \leq 2 \Delta A \Delta B,$$

as was to be shown. \square

Recall that the momentum and position operators have $[p, q]\psi = i^{-1}\psi$ for all $\psi \in \mathcal{S}(\mathbb{R})$. Starting from this one can show:

Exercise - For all $\psi \in D(p^2) \cap D(q^2) \cap D(pq) \cap D(qp)$, $[p, q]\psi = i^{-1}\psi$.

This implies the famous ‘Heisenberg uncertainty principle’:

$$\Delta p \Delta q \geq \frac{1}{2},$$

for all states ψ for which both sides make sense. Putting the units back in for once:

$$\Delta p \Delta q \geq \frac{\hbar}{2}.$$

The consequence of this is that one can never measure *both* the position and momentum of a particle to arbitrary accuracy; accuracy in one must be obtained at the expense of accuracy in the other past a certain point, which is measured by Planck’s constant.

Exercise - Show that the Heisenberg uncertainty principle is sharp: if

$$\psi(x) = \frac{1}{\sqrt{\pi c}} e^{-x^2/c}$$

show that $\|\psi\| = 1$ and $\Delta p \Delta q = \frac{1}{2}$.

Exercise - Given one of the functions ψ as in the previous exercise, show that for any $t \in \mathbb{R}$ the states $e^{itp}\psi$ and $e^{itq}\psi$ also have $\Delta p \Delta q = \frac{1}{2}$. These are translated and ‘boosted’ versions of the state ψ , respectively. Such states are often called ‘wave packets’ since they are states for which the wave function ψ has minimal uncertainty in position and momentum, hence resembles as closely as possible a classical particle.

11. Quantization, Schrödinger’s Equation, and the Hydrogen Atom

To describe the dynamics of a quantum system, i.e. to specify how states evolve in time, it suffices to specify the Hamiltonian, a self-adjoint operator H on the Hilbert space \mathbf{H} whose unit vectors correspond to states: given the state $\psi \in \mathbf{H}$, $e^{itH}\psi$ describes the state t units of time later.

How does one guess the Hamiltonian of a quantum system? I say ‘guess’ because only the real world *knows* the Hamiltonian. Many quantum systems correspond to classical systems for which one knows a Hamiltonian (see section 2). For example, if one has a *classical* particle in \mathbb{R}^3 in a **potential** $V: \mathbb{R}^3 \rightarrow \mathbb{R}$, the Hamiltonian or energy from classical mechanics is a function $H: \mathbb{R}^6 \rightarrow \mathbb{R}$ given by

$$H_{class} = \frac{1}{2m}(p_1^2 + p_2^2 + p_3^2) + V(q_1, q_2, q_3)$$

where the coordinate functions on the state space \mathbb{R}^6 are $(q_1, q_2, q_3, p_1, p_2, p_3)$. The first term in H_{class} is called the **kinetic energy**; m is a positive number called the **mass** of the particle. The second term is called the **potential energy**. An important example is an electron in the electric field of a stationary proton; here the energy is

$$H_{class} = \frac{1}{2m_e}(p_1^2 + p_2^2 + p_3^2) - \frac{e^2}{r}$$

where m_e is the mass of the electron, e is the electric charge of the electron (which is minus that of the proton), and $r = (q_1^2 + q_2^2 + q_3^2)^{1/2}$. A hydrogen atom consists of an electron and

a proton, so classically it would be described by this Hamiltonian. One can show that the solutions of Hamilton's equations with this Hamiltonian correspond to circular, elliptical, parabolic and hyperbolic motions of the electron about the proton: this is essentially what Newton did in the study of a mathematically isomorphic problem, that of the planets (and comets) moving in the gravitational potential of the sun. (Newton did not, of course, use Hamilton's formalism!)

As always when speaking of real physical systems, we are glossing over some subtleties. In classical mechanics the earth does not revolve about a *stationary* sun; rather, the earth and sun revolve about their common center of mass. Since the sun is much more massive than the earth the stationary sun is a good first approximation; similarly, the proton is about 1836 times as massive as the electron. One can solve the problem exactly using a mathematical trick called separation of variables, and it turns out the only effect is to replace the mass of the electron by the **reduced mass** $m = m_e m_p / (m_e + m_p)$, where m_p is the mass of the proton.

Also, we are temporarily ignoring effects due to the electron's spin: the proton produces not only an electric field but also a magnetic field, and the latter interacts with the spin of the electron. In general, physical systems are described by a hierarchy of increasingly accurate models which take into account more and more effects, so when someone says 'the Hamiltonian is...' this must always be taken with a grain of salt.

One guesses the Hamiltonian for the corresponding *quantum* system in the most simple-minded manner possible: one formally substitutes the position and momentum *operators* for the q 's and p 's in the expression H_{class} . Thus for the hydrogen atom the states of the electron are unit vectors in $\mathbf{H} = L^2(\mathbb{R}^3)$, and the Hamiltonian is

$$H = \frac{1}{2m}(p_1^2 + p_2^2 + p_3^2) - \frac{e^2}{r},$$

where in this equation p_i stands for the momentum *operator* ($i^{-1}\partial/\partial x_i$) and r^{-1} stands for the *operator* of multiplication by r^{-1} . Thus H is a differential operator on $L^2(\mathbb{R}^3)$. As Schrödinger wrote it, keeping in Planck's constant (which we normally set equal to 1):

$$H = -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{r}.$$

The original **Schrödinger equation** for the hydrogen atom is:

$$\frac{1}{i} \frac{d\psi}{dt} = -\frac{\hbar^2}{2m}\nabla^2\psi - \frac{e^2}{r}\psi,$$

where $\psi \in D(H)$ and $\psi(t, \vec{x}) = (e^{itH}\psi)(\vec{x})$.

This process of substituting operator p 's and q 's for functions p and q is called **quantization**. Historically it started as a trick with little justification other than that it seemed to work, but now there is a well-developed mathematical theory of it, one aspect of which goes under the name of **geometric quantization**. (For a mathematically sophisticated

treatment see, for example, Jędrzej Śniatycki, *Geometric Quantization and Quantum Mechanics*.) Note that this process is ambiguous in general, because while $pq = qp$ classically, in quantum mechanics $pq \neq qp$, leading to **operator ordering problems** in the quantization of expressions involving products of p 's and q 's. These ambiguities can sometimes be resolved through other criteria; note however that in expressions of the form

$$H = \frac{1}{2m}(p_1^2 + p_2^2 + p_3^2) + V(q_1, q_2, q_3)$$

they do not occur! Indeed, one wonders what the history of quantum mechanics would have been like if there has been operator ordering problems in the Hamiltonian of the hydrogen atom (the first system to which quantum theory was successfully applied).

After deciding upon this strategy for quantization, the first problem we must address is: is the quantized Hamiltonian a self-adjoint operator? The next problem is: what is its spectrum? These questions can be answered by computations in the case of the hydrogen atom Hamiltonian, explicitly determining eigenvectors and approximate eigenvectors. For more complicated systems this 'explicit' approach is hardly ever possible. So rather than wasting time solving the hydrogen atom, which can be found in any quantum mechanics text, we simply describe the answer.

The answer is: the operator H_0 given by

$$D(H_0) = C_0^\infty(\mathbb{R}^3);$$

$$H_0\psi = \left(-\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{r}\right)\psi,$$

is essentially self-adjoint. Note that because of the singularity of the potential e^2/r at $r = 0$, some work is needed to show that H_0 is really well-defined:

Exercise - Show that if $\psi \in D(H_0)$ then $H_0\psi \in L^2(\mathbb{R}^3)$. Show that H_0 is symmetric.

Let $\overline{H_0} = H$. Then the spectrum of the self-adjoint operator H is

$$\left\{-\frac{me^4}{2\hbar^2 n^2}\right\}_{n=1}^\infty \cup [0, \infty).$$

For $n = 1, 2, 3, \dots$ there are n^2 linearly independent vectors ψ such that

$$H\psi = -\frac{me^4}{2\hbar^2 n^2}\psi;$$

the integer n is called the **principal quantum number**. Eigenvectors of the Hamiltonian are also called **bound states**. The $n = 1$ and $n = 2$ states are given as follows:

Exercise - Show that the function

$$\psi(\vec{x}) = e^{-r/a_0},$$

where $r = \|\vec{x}\|$ and $a_0 = \hbar^2/me^2$, satisfies

$$H\psi = -\frac{me^4}{2\hbar^2}\psi.$$

Thus ψ is an eigenvector of H with $n = 1$. Show that the functions

$$\psi(\vec{x}) = xe^{-r/2a_0}, ye^{-r/2a_0}, ze^{-r/2a_0}, \text{ and } \left(2 - \frac{r}{a_0}\right)e^{-r/2a_0}$$

are eigenvectors of H with $n = 2$. The constant a is called the **Bohr radius** of the hydrogen atom and is about .529 angstroms, that is, $.529 \times 10^{-8}$ centimeters.

In general the states with principal quantum number n are of the form: e^{-r/na_0} multiplied by a polynomial of degree n in x, y, z and r . Thus the probability of finding the electron in one of these states at a distance greater than r from the proton decreases rapidly with r . These states correspond physically to states in which the electron is ‘in orbit’ around the proton, so they are also called, especially in chemistry, **orbitals**. The ‘orbit’ idea shouldn’t be taken too seriously, however, because *none* of the bound states is an eigenvector of the position operators q_i (which have no eigenvectors!), so the electron does not follow a curve as time goes on: rather, at any time its position is described by the probability density $|\psi(\vec{x})|^2$ (this follows from the probability interpretation of the spectral measure). If at time t the state is an eigenfunction ψ with $H\psi = \lambda\psi$, then at time t the state is given by

$$\psi(t) = e^{itH}\psi = e^{it\lambda}\psi,$$

so the probability distribution $|\psi(t, \vec{x})|^2$ is independent of time! Thus thinking of the electron as a little ball moving around the proton is misleading for such states.

For $\lambda \geq 0$ there is *no* eigenvector $\psi \in D(H)$ such that $H\psi = \lambda\psi$, rather, these λ are only *approximate* eigenvalues of H . The states which are orthogonal to the bound states are called **scattering states** and correspond physically to states in which the electron is **ionized**, that is, has enough energy to move arbitrarily far from the proton, rather than orbit it. Approximate eigenvectors for H corresponding to the approximate eigenvalues $\lambda \geq 0$ are given as follows:

Exercise - Let $f \in C_0^\infty(\mathbb{R})$ be a decreasing function with $f(r) = 1$ for $0 \leq r \leq 1/2$ and $f(r) = 0$ for $r \geq 1$. Given $\vec{k} \in \mathbb{R}^3$ let

$$\psi_n(\vec{x}) = cf(\|\vec{x} - \vec{x}_0\|/n)e^{i\vec{k} \cdot \vec{x}/\hbar}$$

where \vec{x}_0 is any point with $\|\vec{x}_0\| = 2n$ and c is chosen such that the L^2 norm of ψ_n equals 1. Show that $\psi_n \in D(H)$ and that

$$\lim_{n \rightarrow \infty} \|H\psi_n - \lambda\psi_n\| = 0,$$

where $\lambda = \|\vec{k}\|^2/2m$. Conclude that $[0, \infty)$ is in the spectrum of H .

12. The Kato-Rellich Theorem

In this section we describe a general method for showing that ‘sufficiently small’ perturbations of self-adjoint operators are self-adjoint. Using it we will show that the hydrogen atom Hamiltonian of the previous section is self-adjoint.

Suppose that A and B are densely defined operators on a Hilbert space \mathbf{H} and that:

- 1) $D(B) \supseteq D(A)$; 2) for some $a, b \in \mathbb{R}$,

$$\|B\phi\| \leq a\|A\phi\| + b\|\phi\|$$

for all $\phi \in D(A)$. Then B is said to be **A -bounded**. The infimum of such a is called the **relative bound** of B with respect to A . Typically b will have to be chosen larger as a is chosen smaller.

We begin with two exercises that we’ll need:

Exercise - Show that if A is a self-adjoint operator then so is cA for any real $c \neq 0$ (where $D(cA) = D(A)$). Conclude that if A is densely defined and symmetric then A is self-adjoint if and only if $\text{Ran}(A \pm is) = \mathbf{H}$ for some $s > 0$, in which case it holds for all $s > 0$.

Exercise - Show that if A is densely defined and symmetric then A is essentially self-adjoint if and only if $\text{Ran}(A \pm is)$ is dense in \mathbf{H} for some $s > 0$, in which case it holds for all $s > 0$.

Kato-Rellich Theorem. Suppose that A is self-adjoint and B is A -bounded with relative bound < 1 . Then $A + B$ is self-adjoint on $D(A)$ and essentially self-adjoint on any core for A .

Proof - To show that $A + B$ is self-adjoint we will show that $\text{Ran}(A + B + is) = \mathbf{H}$ for all $s \in \mathbb{R}$ with $|s|$ large enough. Since A is self-adjoint, $(A + is): D(A) \rightarrow \mathbf{H}$ is onto and

$$(A + B + is)\phi = (I + B(A + is)^{-1})(A + is)\phi$$

for all $\phi \in D(A)$. Thus to show that $\text{Ran}(A + B + is) = \mathbf{H}$ we need to show that $\text{Ran}(I + B(A + is)^{-1}) = \mathbf{H}$. For this it suffices to show that $-1 \notin \sigma(B(A + is)^{-1})$, and since the spectral radius is less than or equal to the norm it suffices to show that $\|B(A + is)^{-1}\| < 1$.

For $\phi \in D(A)$ we have

$$\|(A + is)\phi\|^2 = \|A\phi\|^2 + s^2\|\phi\|^2.$$

Letting $\phi = (A + is)^{-1}\psi$, it follows that $\|A(A + is)^{-1}\| \leq 1$ and $\|(A + is)^{-1}\| \leq |s|^{-1}$. Since B is A -bounded, we have

$$\begin{aligned} \|B(A + is)^{-1}\psi\| &\leq a\|A(A + is)^{-1}\psi\| + b\|(A + is)^{-1}\psi\| \\ &\leq (a + b|s|^{-1})\|\psi\|. \end{aligned}$$

Thus for $|s|$ sufficiently large, $B(A + is)^{-1}$ has norm less than one, since we can choose $a < 1$.

Suppose D_0 is a core for A . Then $(A + is)D_0$ is dense in \mathbf{H} if $s \neq 0$, $s \in \mathbb{R}$. Choosing $|s|$ large enough, $\text{Ran}(I + B(A + is)^{-1}) = \mathbf{H}$ by the above. It follows from the exercise below that $\text{Ran}(A + B + is) = (I + B(A + is)^{-1})(A + is)D_0$ is dense in \mathbf{H} . Thus D_0 is a core for $A + B$. \square

Exercise - Show that if $T \in \mathcal{L}(\mathbf{H})$, $\text{Ran}T = \mathbf{H}$, and $D \subseteq \mathbf{H}$ is dense in \mathbf{H} , then TD is dense in \mathbf{H} .

To apply the Kato-Rellich theorem to the hydrogen atom we take

$$A = -\frac{1}{2m}\nabla^2, \quad B = -\frac{e^2}{r}.$$

The work will consist of showing that B is A -bounded. More precisely, recall that $-\nabla^2 = p_1^2 + p_2^2 + p_3^2$ is essentially self-adjoint (an exercise in section 10). We let Δ denote the closure, a self-adjoint operator on $L^2(\mathbb{R}^3)$. By the following exercise C_0^∞ is a core for Δ :

Exercise - Use the fact that $C_0^\infty(\mathbb{R}^3)$ is dense in $\mathbf{S}(\mathbb{R}^3)$ to show that Δ is essentially self-adjoint when restricted to $C_0^\infty(\mathbb{R}^3)$. (Hint: express Δ in terms of the Fourier transform.)

Theorem. Suppose $V \in L^\infty(\mathbb{R}^3) + L^2(\mathbb{R}^3)$, that is, $V = V_1 + V_2$ with $V_1 \in L^\infty(\mathbb{R}^3)$ and $V_2 \in L^2(\mathbb{R}^3)$. Then the operator $\Delta + V$, where here V denotes the operator of multiplication by V , is essentially self-adjoint on $C_0^\infty(\mathbb{R}^3) \subset L^2(\mathbb{R}^3)$ and is self-adjoint on $D(\Delta)$.

Lemma. For all $a > 0$ there exists b such that

$$\|\psi\|_\infty \leq a\|\Delta\psi\|_2 + b\|\psi\|_2$$

for all $\psi \in D(\Delta) \subseteq L^2(\mathbb{R}^3)$.

Proof - If f is the Fourier transform of ψ , we have

$$\|\psi\|_\infty \leq \|f\|_1$$

by the Riemann-Lebesgue lemma and

$$\|\psi\|_2 = \|f\|_2, \quad \|\Delta\psi\|_2 = \|k^2 f\|_2$$

by the Plancherel theorem, so it suffices to show that for all $a > 0$ there exists b such that

$$\|f\|_1 \leq a\|k^2 f\|_2 + b\|f\|_2.$$

Since $\psi \in D(\Delta)$, $k^2 f \in L^2(\mathbb{R}^3)$, so for any $c > 0$, $((ck)^2 + 1)f \in L^2(\mathbb{R}^3)$. It is easily seen that $((ck)^2 + 1)^{-1} \in L^2(\mathbb{R}^3)$, so by the Cauchy-Schwartz inequality

$$\|f\|_1 \leq \|((ck)^2 + 1)^{-1}\|_2 \|((ck)^2 + 1)f\|_2$$

Note that by a change of variables

$$\begin{aligned} \|((ck)^2 + 1)^{-1}\|_2 &= \left(\int \frac{d^3 k}{((ck)^2 + 1)^2} \right)^{1/2} \\ &= \left(\int \frac{c^{-3} d^3 k}{(k^2 + 1)^2} \right)^{1/2} \\ &= c^{-3/2} \|(k^2 + 1)^{-1}\|_2, \end{aligned}$$

so letting $d = \|(k^2 + 1)^{-1}\|_2$, we have

$$\begin{aligned} \|f\|_1 &\leq c^{-3/2} d \|((ck)^2 + 1)f\|_2 \\ &\leq dc^{1/2} \|k^2 f\|_2 + dc^{-3/2} \|f\|_2. \end{aligned}$$

Choosing d such that $dc^{1/2} = a$,

$$\|f\|_1 \leq a \|k^2 f\|_2 + b \|f\|_2. \quad \square$$

Proof of Theorem - It suffices to show that V is Δ -bounded with relative bound less than one. We have

$$\|V\psi\|_2 = \|(V_1 + V_2)\psi\|_2 \leq \|V_1\|_2 \|\psi\|_\infty + \|V_2\|_\infty \|\psi\|_2,$$

so by the lemma, for any $a > 0$ there exists b such that

$$\|V\psi\|_2 \leq \|V_1\|_2 (a \|\Delta\psi\|_2 + b \|\psi\|_2) + \|V_2\|_\infty \|\psi\|_2,$$

so V is Δ -bounded with relative bound zero. \square

Corollary. *The hydrogen atom Hamiltonian,*

$$\frac{1}{2m} \Delta - \frac{e^2}{r},$$

is essentially self-adjoint on $C_0^\infty(\mathbb{R}^3)$ and self-adjoint on $D(\Delta)$.

Proof - It suffices to show that $\Delta - 2me^2 r^{-1}$ has these properties, so by the Kato-Rellich theorem it suffices to show that $r^{-1} \in L^2(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$. We may write

$$r^{-1} = \chi_{\{r \leq 1\}} r^{-1} + \chi_{\{r > 1\}} r^{-1},$$

and the first term is in L^2 while the second is in L^∞ . \square

Kato's theorem is an extension of the corollary above to Hamiltonians describing arbitrary atoms and molecules, or more generally, collections of charged particles. To describe a system of n particles in \mathbb{R}^3 one uses the Hilbert space $L^2(\mathbb{R}^{3n})$, where we give

\mathbb{R}^{3n} the coordinates $(x_{1i}, x_{2i}, x_{3i})_{i=1}^n$. The position and momentum operators for the i th particle are

$$q_{1i} = M_{x_{1i}}, \quad q_{2i} = M_{x_{2i}}, \quad q_{3i} = M_{x_{3i}}$$

and

$$p_{1i} = i^{-1} \frac{\partial}{\partial x_{1i}}, \quad p_{2i} = i^{-1} \frac{\partial}{\partial x_{2i}}, \quad p_{3i} = i^{-1} \frac{\partial}{\partial x_{3i}}.$$

If the i th particle has mass m_i and charge e_i , the Hamiltonian is

$$H = \sum_{i=1}^n \frac{1}{2m_i} (p_{1i}^2 + p_{2i}^2 + p_{3i}^2) - \sum_{1 \leq i < j \leq n} \frac{e_i e_j}{\|\vec{q}_i - \vec{q}_j\|}.$$

Exercise - Prove **Kato's Theorem**: the Hamiltonian H above is essentially self-adjoint on $C_0^\infty(\mathbb{R}^{3n})$. (Hint: copy the proof for the hydrogen atom).

The kinetic energy term, $H_0 = \sum_{i=1}^n \frac{1}{2m_i} (p_{1i}^2 + p_{2i}^2 + p_{3i}^2)$, is often called the **free Hamiltonian**, since it is the Hamiltonian for n 'free' or non-interacting particles:

Exercise - Let the *classical* free Hamiltonian be given by

$$H_{class} = \sum_{i=1}^n \frac{1}{2m_i} (p_{1i}^2 + p_{2i}^2 + p_{3i}^2),$$

where here the p 's are functions on \mathbb{R}^{6n} as in section 2. Show that the solutions of Hamilton's equations correspond to particles moving with constant velocity along straight lines.

Showing that the potential energy term or **interaction Hamiltonian**

$$H_{int} = - \sum_{1 \leq i < j \leq n} \frac{e_i e_j}{\|\vec{q}_i - \vec{q}_j\|}$$

is H_0 -bounded is a way of making precise the idea that it is in some sense a 'small perturbation' of the free Hamiltonian. This situation is typical of mathematical physics: one tries to model a complicated 'interacting' system as a perturbation of a simpler 'free' system.

This approach has failed so far to prove self-adjointness of the quantum-field-theoretic Hamiltonians describing the basic laws of physics, i.e., quantum electrodynamics, or more comprehensively, the 'standard model' of the known particles, interacting via the electromagnetic, weak, and strong forces. These Hamiltonians are not even known to be well-defined operators on a dense domain in some Hilbert space! This is probably *the* problem of mathematical physics today, and the subject of 'constructive field theory' is

the hardest testing-ground for new techniques in analysis. There is no reason (any more) to think that in quantum field theory the interaction Hamiltonian is a ‘small’ perturbation of the free Hamiltonian, but there is no other clear way to begin an attack on the problem. Recently, however, I. Segal has proposed a method which amounts to considering the free Hamiltonian as a small perturbation of the interacting Hamiltonian! While this idea is unlikely to crack the problem without a lot of work being put into it, it is an amusing twist.

The Kato-Rellich theorem is originally due to F. Rellich, ‘Störungstheorie der Spektralzerlegung, II’, *Math. Ann.* **116** (1939) 555-570. Kato’s application of this result to atomic Hamiltonians appeared in ‘Fundamental properties of Hamiltonian operators of Schrödinger type’, *Trans. Amer. Math. Soc.* **70** (1951), 195-211. For introductions to mathematically rigorous quantum field theory, try *Introduction to Axiomatic Quantum Field Theory* by Bogolubov, Logonov and Todorov, or *Quantum Physics: A Functional Integral Point of View* by Glimm and Jaffe, or - please forgive the advertising - *Introduction to Algebraic and Constructive Quantum Field Theory* by Baez, Segal and Zhou.