In this chapter, we discuss some mathematical issues related to the theory of quantum mechanics. We will first give a quick description of the formal structure of quantum mechanics and then prove a number of mathematical (functional analytic) results that are relevant in this context. I will also make a few (slightly off-topic) more philosophical remarks on the interpretation of the formalism; I should perhaps point out that these are very close in spirit to the so-called *Copenhagen interpretation* of quantum mechanics, which, in turn, is close in spirit to Kant’s idealistic philosophy and theory of cognition. There have been other and radically different attempts, too (for example, the notorious *many worlds interpretation*), and I’m not particularly familiar with these. Quantum mechanics according to the Copenhagen interpretation, if nothing else, gives a consistent picture, and it opens up a strange new world of breathtaking elegance and awe-inspiring beauty.

We base our discussion of quantum mechanics on the following principles:

(K1) The state space of a quantum mechanical system is a Hilbert space \( H \); the quantum mechanical states correspond to the one-dimensional subspaces of \( H \). We will usually use normalized vectors (\( \|\psi\| = 1 \)) to represent states.

(K2) Physical observables correspond to self-adjoint operators on \( H \). The system interacts with the observer through measurements of observables (and in no other way). The outcome of a measurement is genuinely random. If the system is in the state \( \psi \in H, \|\psi\| = 1 \), at the time of the measurement of the observable \( A = A^* \), then the probability to observe a value in \( M \subset \mathbb{R} \) is given by

\[
P_A(M) = \|E_A(M)\psi\|^2,
\]

where \( E_A \) denotes the spectral resolution of \( A \): \( A = \int_{\mathbb{R}} t \, dE_A(t) \)

(U) When a measurement has been performed and a value in \( M \) of the observable \( A \) was observed, then the state must be updated according to

\[
\psi_{\text{new}} = \frac{E_M(A)\psi}{\|E_M(A)\psi\|}.
\]

(D) If no measurement is carried out, then the state evolves according to a group of unitary operators:

\[
\psi(t) = U(t)\psi(0), \quad U(s + t) = U(s)U(t),
\]

and here \( U(t) \) is a unitary operator on \( H \) for every \( t \in \mathbb{R} \).
(K1), (K2) show how a system is described at a fixed point in time ("kinematics"), (U) shows how to update the state when new information becomes available through a measurement, and (D) describes the dynamics of $\psi$ of an unobserved system. Note that $\psi$ is a purely formal construct; it acquires physical meaning only through (K2), when a measurement is performed. It is a list of potential knowledge about the system, perhaps similar to a weather forecast (but note the ingenious way in which probabilities for all possible measurements are encoded in $\psi$, making the crude weather forecast pale in comparison).

On a more philosophical level, we can say that $\psi$ does not describe the “system as such,” but only how it is perceived through measurements. This simple precaution immediately gets rid of a number of alleged “paradoxes” such as Schrödinger’s cat, and it also demystifies the “discontinuity” in the evolution of $\psi$ that comes with (U) and that has occasionally been perceived as troublesome: of course, probabilities have to be updated instantaneously as new information becomes available. In fact, far from introducing any mysterious discontinuity, (U) implies a very desirable continuity property: if we measure an observable, and, having completed this measurement, immediately measure the same observable again, then, with probability 1, we will observe the outcome of the first measurement again.

Notice how the emphasis shifts here, compared to classical physics. Some scientists and philosophers have found it tempting, for reasons that are unclear to me, to decree that, whether or not we choose to observe a classical system, it will evolve according to a set of physical laws, and if we do observe it, we just read off what the equations tell the system to do. Adherents of such a realistic (in the technical sense, as the opposite of idealistic) philosophy usually proudly declare that physical objects “exist,” independently of their being observed, but fail to explain what verifiable consequences this “existence” might have (or what exactly it refers to).

One can probably avoid immediate disaster by a judicious application of this unwarranted “existence” assumption in classical physics, but this does not seem to be good philosophical practice. More to the point, this philosophy becomes untenable in quantum mechanics. Matters are very different: The theory explicitly only talks about the system how it interacts with the observer, and in fact it would be cleaner to not insist on the existence of a system or anything else having a life of its own, independent of the observer.

(K1)–(D) describe the skeleton of quantum mechanics. While this style of developing the theory is sometimes referred to as axiomatic quantum mechanics, (K1)–(D) are not axioms in the sense the word is
usually used in mathematics; for a careful development of the theory, further analysis and clarifying comments are required. For example, (U) might cause difficulties if $A$ is such that $E_A(\{x\}) = 0$ for all $x \in \mathbb{R}$ (many self-adjoint operators are of this type). Indeed, one suggestion is to just admit projections as observables, corresponding to yes-no questions as the only admissible measurements. We don’t want to discuss these issues here.

Moreover, and perhaps more to the point for the working physicist, (K1)–(D) show how the theory works in principle, but not how to model concrete systems. For example, if I want to study the hydrogen atom in its ground state, and perform a measurement of the angular momentum of the electron, what are $H, \psi, A, U(t)$? Clearly, this is a totally separate issue. We will not say much about this here.

As our first mainly mathematical topic in this chapter, we discuss an important result that gives a reformulation of (D). To prepare for this, let $T$ be a self-adjoint operator and let $U(t) = e^{-itT}$. Recall that this is defined via the functional calculus as $U(t) = \int_{\mathbb{R}} e^{-its} dE(s)$, where $E$ is the spectral resolution of $T$. By the properties of the functional calculus, $U(t)$ is unitary for every $t \in \mathbb{R}$ and $U(s + t) = U(s)U(t)$.

Exercise 12.1. Prove these properties of $U(t)$.

In other words, $U(t)$ is a unitary group as in (D). Moreover, $U(t)$ is also strongly continuous: this means that for every fixed $x \in H$, the map $\mathbb{R} \to H, t \mapsto U(t)x$ is continuous. To see this, notice that

$$
\|U(t)x - U(s)x\|^2 = \int_{\mathbb{R}} |e^{-itv} - e^{-isv}|^2 d\mu_{x,x}(v)
$$

by the properties of the functional calculus. The Dominated Convergence Theorem shows that the right-hand side goes to zero as $s \to t$, as claimed.

Exercise 12.2. Let $T$ be a self-adjoint operator on a Hilbert space $H$, and let $U(t) = e^{-itT}$. Show that the map $\mathbb{R} \to B(H), t \mapsto U(t)$ is continuous if and only if $T \in B(H)$.

Exercise 12.3. Let again $T$ be self-adjoint and $U(t) = e^{-itT}$. Suppose that $x \in D(T)$. Show that then $U(t)x \in D(T)$ for all $t \in \mathbb{R}$. Moreover, show that the map $t \mapsto U(t)x$ is differentiable and $(d/dt)U(t)x = -iTU(t)x$ (this of course means that $\lim_{h \to 0} (1/h)(U(t + h) - U(t))x$ exists (in $H$) and equals $-iTU(t)x$).

Hint: Use an argument similar to the one that was used above to establish the strong continuity of $U(t) = e^{-itT}$. 

Stone’s Theorem asserts that, conversely, every strongly continuous unitary group is of this type.

**Theorem 12.1** (Stone). Let $U(t)$ be a strongly continuous unitary group. Then there exists a unique self-adjoint operator $T$ so that $U(t) = e^{-itT}$.

We call $T$ the (infinitesimal) generator of $U(t)$.

**Exercise 12.4.** Let $U(t)$ be a unitary group. Prove that $U(0) = 1$. Also, prove that $U(t)^* = U(-t)$.

**Proof.** From Exercise 12.3, we already have at least a vague idea of how to find such a $T$: we have to “differentiate” $U(t)$. It thus seems natural to define

$$D(S) = \left\{ x \in H : \lim_{h \to 0} \frac{U(h) - 1}{h} x \text{ exists} \right\},$$

$$Sx = i \lim_{h \to 0} \frac{U(h) - 1}{h} x.$$  \hspace{1cm} (12.1)

It is easy to see that $D(S)$ is a subspace and $S$ is linear. We next claim that $D(S)$ is dense. To show this, we will make use of Hilbert space valued integrals. However, we will not develop this subject carefully here; instead, we will leave some of the details to the reader.

For $x \in H$ and $f \in C_0^\infty(\mathbb{R})$, we want to define

$$x_f = \int_{-\infty}^{\infty} f(t)U(t)x \, dt.$$  \hspace{1cm} (12.2)

What exactly do we mean by this? Since the integrand takes values in $H$, this question certainly has to be asked. Fortunately, several good answers are available. Here, we only need to be able to integrate continuous functions, so (generalized) Riemann sums provide a convenient interpretation of (12.2): We take $R \in \mathbb{N}$ so large that $	ext{supp } f \subset (-R, R)$, then form $(1/N) \sum_{n=-RN}^{RN} f(n/N)U(n/N)x$ and finally take the limit $N \to \infty$ to define the right-hand side of (12.2).

Existence of this limit is an easy consequence of the continuity of the integrand, just as in the elementary theory of the Riemann integral. In the sequel, we will make use of some (very plausible) basic properties of this new integral without worrying too much about their formal verification; we leave this to the reader (see Exercise 12.7 below).
First of all, we claim that \( x_f \in D(S) \) whenever \( f \in C_0^\infty(\mathbb{R}) \). This follows from the following calculation:

\[
\frac{U(h) - 1}{h} x_f = \frac{1}{h} \int_{\mathbb{R}} f(t)(U(t + h) - U(t)) x \, dt = \int_{\mathbb{R}} \frac{f(t) - f(t)}{h} U(t) x \, dt
\]

Now as \( h \to 0 \), we have that \( (f(t) - f(t))/h \to -f'(t) \), uniformly in \( t \in \mathbb{R} \).

Exercise 12.5. Prove this. The point here of course is the uniform convergence; convergence for fixed \( t \) just follows from the definition of the derivative.

From this, it follows that \((1/h)(U(h) - 1)x_f \to x_{-f'}\); this shouldn’t be very hard to believe because Riemann integration can be interchanged with uniform limits. In particular, \( x_f \in D(S) \), as claimed.

Now if \( x \in H \) is arbitrary, fix an \( f \in C_0^\infty(\mathbb{R}) \) with \( \int f = 1 \), and let \( f_n(t) = nf(nt) \), \( x_n = x_{f_n} \). Then

\[
\|x_n - x\| = \left\| \int_{\mathbb{R}} f_n(t)(U(t)x - x) \, dt \right\| \leq \int_{\mathbb{R}} |f_n(t)| \|U(t)x - x\| \, dt.
\]

Notice that the \( f_n \) are supported by \((-R/n, R/n)\), for suitable fixed \( R > 0 \), and \( \sup_{|t|<R/n} \|U(t)x - x\| \to 0 \) as \( n \to \infty \), by the strong continuity of \( U(t) \) and Exercise 12.4. Since \( \int |f_n| \) is independent of \( n \), it follows that \( x_n \to x \), and, as observed earlier, \( x_n \in D(S) \), so \( D(S) \) is indeed dense.

Given this, it is now easy to verify that \( S \) is symmetric. Let \( x,y \in D(S) \). Then, by the continuity of the scalar product and the second part of Exercise 12.4,

\[
\langle x, Sy \rangle = i \lim_{h \to 0} \left\langle x, \frac{U(h) - 1}{h} y \right\rangle = i \lim_{h \to 0} \left\langle \frac{U(-h) - 1}{h} x, y \right\rangle = i \langle Sx, y \rangle = \langle Sx, y \rangle.
\]

Next, we claim that \( T = S \) is self-adjoint. \( T \) is symmetric by Exercise 11.6, and we will use Theorem 11.6(b) with \( z = i \) to establish self-adjointness. As a preliminary, observe that if \( x \in D(S) \), then \( U(t)x \in D(S) \) also for all \( t \in \mathbb{R} \) because

\[
\frac{U(h) - 1}{h} U(t)x = U(t) \frac{U(h) - 1}{h} x \to -iU(t)Sx.
\]
Quantum mechanics

Now suppose that \( x \in N(T^* - i) = N(S^* - i) \), fix an arbitrary vector \( y \in D(S) \), and let
\[
f(t) = \langle x, U(t)y \rangle.
\]
By the observation just made, \( f \) is differentiable, and in fact
\[
f'(t) = \langle x, -iSU(t)y \rangle = -i\langle S^*x, U(t)y \rangle = -i\langle ix, U(t)y \rangle = -f(t).
\]
This ODE has the unique solution \( f(t) = f(0)e^{-t} \). This is unbounded unless \( f(0) = 0 \), but, on the other hand, \( |f(t)| \leq \|x\|\|y\| \) is clearly bounded, so \( f(0) = \langle x, y \rangle = 0 \), and since \( y \in D(S) \) was arbitrary here, we have proved that \( x \in D(S)^\perp = \{0\} \). The proof that \( N(T^* + i) = \{0\} \) is analogous, of course.

Let \( V(t) = e^{-itT} \). We want to show that \( U(t) = V(t) \). Let \( x \in D(S) \), and put \( w(t) = U(t)x - V(t)x \). Then \( w \) is differentiable, by our discussion from the preceding paragraph (for \( U(t)x \)) and Exercise 12.3 (for \( V(t)x \)); also notice that \( D(S) \subset D(T) \). We have that
\[
w'(t) = -iSU(t)x + iTV(t)x = -iTw(t).
\]
Hence,
\[
\frac{d}{dt} \langle w(t), w(t) \rangle = \langle w', w \rangle + \langle w, w' \rangle = i\langle Tw, w \rangle - i\langle w, Tw \rangle = 0,
\]
by the symmetry of \( T \).

Exercise 12.6. Formulate and prove the product rule for derivatives of the scalar product that was used here.

Since \( w(0) = 0 \), this shows that \( U(t)x = V(t)x \) for all \( t \in \mathbb{R} \). Since this holds for all \( x \) from the dense set \( D(S) \) and since both \( U(t) \) and \( V(t) \) are bounded operators, we obtain that \( U(t) = V(t) \), as desired.

This construction of the infinitesimal generator \( T \) also yields uniqueness: If also \( U(t) = e^{-itA} \), with \( A = A^* \), then Exercise 12.3 with \( t = 0 \) (and \( A \) taking the role of \( T \)) shows that \( A \subset T \), where \( T \) is the generator constructed in this proof. As both \( A \) and \( T \) are self-adjoint, this implies that \( A = T \). \(\square\)

Exercise 12.7. Provide the omitted details in the first part of the proof. More precisely, clearly state and then prove the properties of integrals of the type \( \int \varphi(t) \ dt \) (where \( \varphi : \mathbb{R} \to H \) is a compactly supported continuous function) that were used in the proof.

In fact, this proof has shown slightly more than originally stated. By Exercise 12.3 again, the limit from the definition of \( D(S) \) exists for all \( x \in D(T) \), so \( S = T \) and the closure operation from the proof turns out to be unnecessary. In other words, (12.1) gives a description of the generator of \( U(t) \).
Once we know that every strongly continuous group has an infinitesimal generator, the statement of Exercise 12.3 becomes completely general. It is of particular interest in quantum mechanics, so we formulate it one more time.

**Corollary 12.2.** Let $U(t)$ be a strongly continuous unitary group with generator $T$. Suppose that $\psi \in D(T)$. Then $U(t)\psi \in D(T)$ for all $t \in \mathbb{R}$ and $\psi(t) \equiv U(t)\psi$ satisfies the (Hilbert space valued) differential equation

$$i\frac{d\psi(t)}{dt} = T\psi(t).$$

In the context of quantum mechanics, this is the famous (time-dependent) Schrödinger equation. The self-adjoint operator $T$ is called the Schrödinger operator (mathematicians) or the Hamilton operator (physicists) of the system. We can reformulate (D), at least, if we also assume strong continuity of $U(t)$ (an assumption that looks eminently desirable).

(D’) There exists a self-adjoint operator $T$, the Schrödinger operator of the system, so that the time-evolved state $\psi(t)$, while no measurement is performed, obeys the Schrödinger equation

$$i\frac{d\psi(t)}{dt} = T\psi(t).$$

This description of the dynamics of a quantum system is found in most physics books. Note, however, that we have sacrificed some precision here: The Schrödinger equation can be taken at face value only if $\psi(0) \in D(T)$; otherwise, one has to fall back on the unitary group $U(t) = e^{-itT}$. We can also say that with the help of the unitary group, we can “solve” the Schrödinger equation, $\psi(t) = e^{-itT}\psi(0)$, and this solution has the added benefit that it works for all $\psi(0) \in H$. We have used quotation marks here because $e^{-itT}$ is defined via the functional calculus as $e^{-itT} = \int e^{-its} dE(s)$, so it will not be easy to extract information from this unless one has very good knowledge of the spectral resolution $E$ of $T$.

Note also that only self-adjoint operators are admissible as Schrödinger operators of quantum mechanical systems because only these arise as generators of unitary groups. In particular, it is not enough to just come up with some symmetric operator and leave the matter at that. Domain issues need to be discussed very carefully.

Our second set of mathematical results in this Chapter shows how to use spectral theory to obtain valuable information on the dynamics
in cases where a complete analysis of $U(t) = e^{-itT}$ is not possible (this is all cases of interest minus perhaps one or two). We will now also, at last, make at least a few remarks about how to model concrete systems.

To keep matters digestible, we focus on one-particle systems right away. Then the Hilbert space is usually taken to be $H = L^2(\mathbb{R}^3)$, and the variable $x \in \mathbb{R}^3$ is interpreted as space. However, we will be interested almost exclusively in discrete analogs of these models, and so we will right away move on to the case $H = \ell^2(\mathbb{Z}^d)$. Again, we want to interpret $n \in \mathbb{Z}^d$ as position, so you should think of a quantum mechanical particle whose position, if it were measured, is restricted to the lattice points $n \in \mathbb{Z}^d$. We can be more explicit: the observable describing the measurement “is the particle in $M$?”, where $M \subset \mathbb{Z}^d$, is given by $(P_M \psi)(n) = \chi_M(n)\psi(n)$. This is in fact the orthogonal projection onto the subspace $\{\varphi \in \ell^2 : \varphi(n) = 0 \text{ if } n \notin M\}$; in particular, $P_M$ is self-adjoint, as it must be according to (K2).

**Exercise 12.8.** Let $P$ be an orthogonal projection. Show that the spectral resolution $E$ of $P$ is given by $E(\{0\}) = 1 - P$, $E(\{1\}) = P$, $E(\mathbb{R} \setminus \{0, 1\}) = 0$.

Suppose that the system is in the state $\psi \in \ell^2$, $\|\psi\| = 1$, and we measure $P_M$. What are the probabilities? According to (K2) and Exercise 12.8, we have that

$$\text{prob}(1) = \|P_M \psi\|^2 = \sum_{n \in M} |\psi(n)|^2,$$

$$\text{prob}(0) = \|(1 - P_M) \psi\|^2 = \sum_{n \notin M} |\psi(n)|^2.$$

We of course interpret 1 as “particle is found in $M$” and 0 as “particle is not found in $M$.” So it turns out that $|\psi(n)|^2$ has a direct physical interpretation as the probability of finding the particle at site $n$, if we choose to measure this observable (if we don’t measure it, then of course any talk about the position of the particle becomes meaningless and in fact highly destructive from a philosophical point of view, as we discussed above).

Now suppose we prepare the system in a certain state $\psi$ and we wait a long time before we carry out such a measurement. Where will we, most likely, find the particle? This clearly is a very relevant physical question, and so it seems reasonable to classify states according to their long term dynamic behavior (in this sense).

Let $T$ be a self-adjoint operator on $\ell^2(\mathbb{Z}^d)$, thought of as the Schrödinger operator of some quantum system. Evolve $\psi \in \ell^2$, according to (D), (D'), $\psi(t) = e^{-itT} \psi$, and denote the probability of finding the particle
in $M$ at time $t$ by $P_M(t) = \sum_{n \in M} |\psi(n, t)|^2$ ("if a measurement is performed", we should of course add, but, as always, very precise language easily gets tedious and convoluted, so we will sometimes omit this). We will also use the quantity $P_M$ for not necessarily normalized $\psi \in \ell^2$; then $0 \leq P_M \leq \|\psi\|^2$.

**Definition 12.3.** Let $\psi \in \ell^2(\mathbb{Z}^d)$. We call $\psi$ a **strong scattering state** if
\[
\lim_{|t| \to \infty} P_M(t) = 0
\]
for every finite set $M \subset \mathbb{Z}^d$, then we call $\psi$ a **weak scattering state**. If for every $\epsilon > 0$, we can find a finite set $M \subset \mathbb{Z}^d$ so that $P_M(t) < \epsilon$ for all $t \in \mathbb{R}$, then we call $\psi$ a **bound state**.

We write $H_{ss}, H_{ws}, H_b$ for the corresponding subsets of $\ell^2$.

So, roughly speaking, if the system is in a scattering state, then the particle will leave every bounded set if you just wait long enough (the usual qualifications apply: the system of course doesn’t do anything other than respond to questions in the form of measurements; it is really the information in the form of probabilities as encoded in $\psi(t)$ that evolves, not something with an existence of its own). In a weak scattering state, we can only make such a statement about the time averaged probabilities. If the system is in a bound state, on the other hand, it can essentially be confined to a bounded set for all times.

Obviously, $H_{ss} \subset H_{ws}$. More can be said:

**Proposition 12.4.** $H_{ss}, H_{ws}, H_b$ are closed subspaces and $\ell^2 = H_{ws} \oplus H_b$.

We postpone the proof because other results that we will develop later will come in handy here.

We want to relate the dynamically defined subspaces $H_{ss}, H_{ws}, H_b$ to spectral subspaces, so we need to discuss this topic first. We do this in an abstract setting.

So let $T$ be a self-adjoint operator on a Hilbert space $H$. Recall that a Borel measure $\rho$ on $\mathbb{R}$ is called **absolutely continuous** if $\rho(B) = 0$ for all Borel sets $B \subset \mathbb{R}$ of Lebesgue measure zero. By the Radon-Nikodym Theorem, $\rho$ is absolutely continuous if and only if $d\rho(t) = f(t) \, dt$ for some density $f \in L^1_{loc}(\mathbb{R}), f \geq 0$. If $\rho$ is supported by a Lebesgue null set (that is, there exists a Borel set $B \subset \mathbb{R}$ with $m(B) = \rho(B^c) = 0$), then we say that $\rho$ is **singular**. If $\rho$ is even supported by a countable set, then we call $\rho$ a (pure) point measure. We call $\rho$ continuous if $\rho(\{x\}) = 0$ for all $x \in \mathbb{R}$, and a **singular continuous** measure is a
measure that is both singular and continuous (the standard example being the Cantor measure).

Any Borel measure \( \rho \) on \( \mathbb{R} \) can be uniquely decomposed into absolutely continuous, singular continuous, and point parts:

\[
\rho = \rho_{ac} + \rho_{sc} + \rho_{pp}
\]

Exercise 12.9. Derive this refined decomposition from Lebesgue’s decomposition theorem (see, for example, Folland, Real Analysis, Theorem 3.8), which says that we can, in a unique fashion, write \( \rho = \rho_{ac} + \rho_s \), where \( \rho_{ac} \) is absolutely continuous and \( \rho_s \) is singular. In other words, you need to further break up \( \rho_s \).

We now apply these notions to spectral measures to define the spectral subspaces. We write \( d\mu_x(t) = \|E(t)x\|^2 \) for the spectral measure of \( T \) and \( x \) (we used to denote this by \( \mu_{x,x} \)).

**Definition 12.5.** The absolutely continuous, singular continuous, and pure point subspaces are defined as follows:

\[
H_{ac} = \{ x \in H : \mu_x \text{ absolutely continuous} \}
\]
\[
H_{sc} = \{ x \in H : \mu_x \text{ singular continuous} \}
\]
\[
H_{pp} = \{ x \in H : \mu_x \text{ pure point measure} \}
\]

**Theorem 12.6.** \( H_{ac}, H_{sc}, H_{pp} \) are closed subspaces; in fact, they are reducing subspaces for \( T \). Moreover,

\[
H = H_{ac} \oplus H_{sc} \oplus H_{pp}
\]

**Proof.** We first show that the \( H_{\cdot} \) are closed subspaces and that (12.4) holds. We will make use of the following fact, or rather the version for three subsets.

Exercise 12.10. Let \( A, B \) be subsets of a Hilbert space \( H \) and suppose that \( A \perp B \) and \( H = A + B \) (that is, every \( x \in H \) can be written in the form \( x = a + b \) with \( a \in A, b \in B \)). Show that then \( A = B^\perp, B = A^\perp \), and thus \( H = A \oplus B \).

Let \( x \in H \), and decompose \( \mu_x \) as in (12.3): \( \mu_x = \mu_{ac} + \mu_{sc} + \mu_{pp} \). By the defining properties of the individual parts, we can find disjoint Borel sets \( S_{ac}, S_{sc}, S_{pp} \) that support the corresponding \( \mu \)'s. Then their union supports \( \mu \) (or we can just assume that this union is all of \( \mathbb{R} \)), so

\[
x = E(S_{ac} \cup S_{sc} \cup S_{pp})x = E(S_{ac})x + E(S_{sc})x + E(S_{pp})x.
\]

Notice that, for example,

\[
\mu_{E(S_{ac})x}(M) = \|E(M)E(S_{ac})x\|^2 = \|E(M \cap S_{ac})x\|^2
\]
\[
= \mu_x(M \cap S_{ac}) = \mu_{ac}(M),
\]
and similarly for the other parts, so $E(S_j)x \in H_j$ for $j = ac, sc, pp$. This proves that $H = H_{ac} + H_{sc} + H_{pp}$. To prove that these sets are orthogonal to each other, let $x \in H_{ac}, y \in H_{sc},$ say. As above, the corresponding spectral measures admit disjoint supports $S_{ac}, S_{sc}$ (because one measure is absolutely continuous, the other is singular). It follows that

$$\langle x, y \rangle = \langle E(S_{ac})x, E(S_{sc})y \rangle = \langle x, E(S_{ac})E(S_{sc})y \rangle = 0.$$ 

An argument of this type works in all cases.

To prove that these subspaces are reducing, we will use the criterion from Exercise 11.21(a). So let $P$ be the projection onto $H_{ac}$, say; we want to show that $PT \subset TP$. Notice that Theorem 11.13(c) implies that

$$(12.5) \quad E(A)T \subset TE(A)$$

for all Borel sets $A \subset \mathbb{R}$.

Let $x \in D(T)$. Fix again a Borel set $S \subset \mathbb{R}$ that supports $(\mu_x)_{ac}$ and is given zero weight by the singular part of $\mu_x$. Then, as above, $Px = E(S)x$. Moreover, $E(S)x \in D(T)$, too, so $x \in D(TP)$ and $TPx = TE(S)x$. By (12.5), this equals $E(S)Tx$, so it just remains to show that $E(S)Tx = PTx$. Now (12.5) also implies that

$$(12.6) \quad d\mu_{Tx}(t) = t^2 d\mu_x(t).$$

From the first part of the proof, we know that we can obtain $P(Tx)$ as $E(M)(Tx)$, where the set $M \subset \mathbb{R}$ needs to be chosen so that it supports the ac part of $\mu_{Tx}$ and is given zero weight by the singular part of the same measure. By (12.6), a set that works for $x$ will also work for $Tx$, so we can take $M = S$.

It is useful to note that $H_{pp}$ has an alternative description. As usual, we call $x \in H, x \neq 0$ an eigenvector with eigenvalue $t \in \mathbb{R}$ if $x \in D(T)$ and $Tx = tx$.

**Proposition 12.7.** $H_{pp}$ is the closed linear span of the eigenvectors of $T$.

**Exercise 12.11.** Prove that $x \neq 0$ is an eigenvector with eigenvalue $t$ if and only if $E(\{t\})x = x$.

**Proof.** If $x$ is an eigenvector with eigenvalue $t$, then, by the Exercise, $x = E(\{t\})x$, so

$$\mu_x(M) = \|E(M)x\|^2 = \|E(M \cap \{t\})x\|^2$$

is supported by $\{t\}$ and thus definitely a pure point measure. In other words, $x \in H_{pp}$ for all eigenvectors $x$. Since $H_{pp}$ is a closed subspace,
this implies that the closed linear span of the eigenvectors is contained in $H_{pp}$.

Conversely, if $x \in H_{pp}$ and $\{t_j : j \in \mathbb{N}\}$ supports $\mu_x$, then

$$(12.7) \quad x = E(\{t_j : j \in \mathbb{N}\})x = \lim_{N \to \infty} \sum_{j=1}^{N} E(\{t_j\})x,$$

and by Exercise 12.11 again, this shows that $x$ is in the closed linear span of the eigenvectors. \hfill \square

**Exercise 12.12.** Give a careful argument for the second equality in (12.7).

We now return to the situation where $H = \ell^2(\mathbb{Z}^d)$ and consider the dynamical subspaces from Definition 12.3.

**Theorem 12.8.** $H_{ws} = H_c$, $H_b = H_{pp}$. Moreover, $H_{ac} \subset H_{ss}$.

Here, $H_c$ denotes the continuous subspace $H_c = H_{ac} \oplus H_{sc}$; in other words, $x \in H_c$ if and only if $\mu_x$ is a continuous measure.

Theorem 12.8 depends on two easy classical results on the Fourier transform of measures; in fact, it could be argued that Theorem 12.8 is essentially a rephrasing of these results. To see why Fourier transforms are relevant here, denote the standard ONB of $\ell^2(\mathbb{Z}^d)$ by $\{\delta_n : n \in \mathbb{Z}^d\}$.

So $\delta_n(m) = 1$ if $m = n$ and $\delta_n(m) = 0$ otherwise. Write

$$\hat{\mu}(t) = \int_{\mathbb{R}} e^{-its} d\mu(s)$$

for the Fourier transform of a finite (possibly complex) Borel measure $\mu$. Then

$$P_M(t) = \sum_{n \in M} |\langle \delta_n, e^{-iT\psi} \rangle|^2 = \sum_{n \in M} |\hat{\rho}_n(t)|^2,$$

where we use the notation $\rho_n(B) = \langle \delta_n, E(B)\psi \rangle$ for the (complex) spectral measure associated with the vectors $\delta_n, \psi$.

Here are the two results about Fourier transforms that will form the basis of our discussion. Our measures are still assumed to be complex Borel measures on $\mathbb{R}$; in particular, they must be finite if they are positive.

**Theorem 12.9** (Riemann-Lebesgue Lemma). Let $\mu$ be an absolutely continuous measure. Then $\lim_{|t| \to \infty} \hat{\mu}(t) = 0$.

**Theorem 12.10** (Wiener).

$$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |\hat{\mu}(t)|^2 \, dt = \sum_{x \in \mathbb{R}} |\mu(\{x\})|^2$$
Existence of the limit is part of the statement in Wiener’s Theorem. Note that \( \mu(\{x\}) \neq 0 \) for at most countably many \( x \in \mathbb{R} \), so there are no difficulties involved in defining the sum. We will not use the formula from Wiener’s Theorem, but the following immediate Corollary:

**Corollary 12.11.**

\[
\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |\hat{\mu}(t)|^2 \, dt = 0
\]

if and only if \( \mu_{pp} = 0 \).

**Exercise 12.13.** Provide proofs of the Riemann-Lebesgue Lemma and Wiener’s Theorem. You can of course look these up in the literature, but it might be more fun to try for yourself first.

**Proof of Theorem 12.8 and Proposition 12.4.** Notice that

\[
|\rho_n(B)| = |\langle \delta_n, E(B)\psi \rangle| \leq \|E(B)\psi\|
\]

so \( \rho_n \ll \mu_\psi \). Thus, if \( \psi \in H_{ac} \), then the \( \rho_n \) \((n \in \mathbb{Z}^d)\) are absolutely continuous measures, too. The Riemann-Lebesgue Lemma now shows that

\[
P_M(t) = \sum_{n \in M} |\hat{\rho}_n(t)|^2 \to 0 \quad (t \to \pm \infty)
\]

for all finite \( M \subset \mathbb{Z}^d \), so \( H_{ac} \subset H_{ss} \).

A similar argument, with Corollary 12.11 replacing the Riemann-Lebesgue Lemma, shows that \( H_c \subset H_{ws} \). Conversely, if \( \psi \in H_{ws} \), then in particular

\[
\frac{1}{2T} \int_{-T}^{T} |\hat{\rho}_n(t)|^2 \, dt \to 0
\]

for all \( n \in \mathbb{Z}^d \), so, by Corollary 12.11 again, every \( \rho_n \) is a continuous measure. But then \( \mu_\psi \) is a continuous measure, too, because if \( x \in \mathbb{R} \) is an arbitrary point, then

\[
\mu_\psi(\{x\}) = \langle \psi, E(\{x\})\psi \rangle = \sum_n \overline{\psi(n)} \langle \delta_n, E(\{x\})\psi \rangle = \sum_n \overline{\psi(n)} \rho_n(\{x\}) = 0.
\]

In other words, \( \psi \in H_c \). This identification \( H_{ws} = H_c \) also proves that \( H_{ws} \) is a closed (in fact: reducing) subspace, as claimed in Proposition 12.4.

Next, we show directly that \( H_b \) is a closed subspace also. Obviously, \( c\psi \in H_b \) if \( \psi \in H_b \), and the estimate

\[
|\langle \delta_n, e^{-it} (\psi_1 + \psi_2) \rangle|^2 \leq 2 \left| \langle \delta_n, e^{-it} \psi_1 \rangle \right|^2 + 2 \left| \langle \delta_n, e^{-it} \psi_2 \rangle \right|^2
\]
Quantum mechanics makes it clear that also \( \psi_1 + \psi_2 \in H_b \) if \( \psi_{1,2} \in H_b \). If \( \psi_j \in H_b, \psi_j \rightarrow \psi \), and \( \epsilon > 0 \) is given, pick first \( j \in \mathbb{N} \) (large enough) so that \( \| \psi_j - \psi \|^2 < \epsilon \)
and then \( M \subset \mathbb{Z}^d \), \( M \) finite, so large that
\[
\sum_{n \notin M} |\langle \delta_n, e^{-itT} \psi_j \rangle|^2 < \epsilon
\]
for all \( t \in \mathbb{R} \) (and this \( j \)). Since
\[
\sum_{n \in \mathbb{Z}^d} \left| \langle \delta_n, e^{-itT} \psi_j \rangle - \langle \delta_n, e^{-itT} \psi \rangle \right|^2 = \|e^{-itT} (\psi_j - \psi)\|^2 = \|\psi_j - \psi\|^2,
\]
by Parseval’s identity, it then follows that
\[
\sum_{n \notin M} \left| \langle \delta_n, e^{-itT} \psi \rangle \right|^2 < 4\epsilon
\]
for all \( t \in \mathbb{R} \). Thus \( \psi \in H_b \), and \( H_b \) is closed.

Exercise 12.14. Prove in a similar way that \( H_{ss} \) is a closed subspace.

Next, let \( \psi \) be an eigenvector, with eigenvalue \( s \), say. Then \( \langle \delta_n, e^{-itT} \psi \rangle = e^{-ist} \langle \delta_n, \psi \rangle \), so
\[
P_{M^c}(t) = \sum_{n \notin M} |\langle \delta_n, \psi \rangle|^2,
\]
and Parseval’s identity ensures that this can be made arbitrarily small by taking \( M \) sufficiently large. So all eigenvectors belong to \( H_b \), and \( H_b \) is a closed subspace, so Proposition 12.7 shows that \( H_{pp} \subset H_b \).

Finally, we claim that \( H_b \perp H_{ws} \). To prove this, let \( \varphi \in H_b, \psi \in H_{ws} \). Then
\[
\langle \varphi, \psi \rangle = \langle e^{-itT} \varphi, e^{-itT} \psi \rangle = \frac{1}{2T} \int_T^{-T} \langle e^{-itT} \varphi, e^{-itT} \psi \rangle \, dt
\]
\[
= \frac{1}{2T} \int_T^{-T} \sum_{n \in \mathbb{Z}^d} \langle e^{-itT} \varphi, \delta_n \rangle \langle \delta_n, e^{-itT} \psi \rangle \, dt.
\]
We split this sum into two parts, as follows. Given \( \epsilon > 0 \), we can find a finite set \( M \subset \mathbb{Z}^d \) so that \( \|\psi\|^2 \sum_{n \notin M} |\langle e^{-itT} \varphi, \delta_n \rangle|^2 < \epsilon^2 \) for all \( t \in \mathbb{R} \). With the help of the Cauchy-Schwarz inequality and Parseval’s identity, we thus see that
\[
\left| \sum_{n \notin M} \langle e^{-itT} \varphi, \delta_n \rangle \langle \delta_n, e^{-itT} \psi \rangle \right| < \epsilon,
\]
and thus the corresponding time averaged quantity is also \( \epsilon \). Since \( \psi \in H_{ws}, \) we have that \( 1/(2T) \int_T^{-T} \|\langle \delta_n, e^{-itT} \psi \rangle\|^2 \, dt \rightarrow 0 \) for all \( n \in \mathbb{Z}^d \).
Now $M$ is finite, so we can take $T > 0$ so large that also

$$\left| \sum_{n \in M} \frac{1}{2T} \int_{-T}^{T} \langle e^{-itT} \varphi, \delta_n \rangle \langle \delta_n, e^{-itT} \psi \rangle \, dt \right| < \epsilon.$$ 

We again use the Cauchy-Schwarz inequality here, and in fact we use it twice, to estimate the integral as well as the sum. Our estimates have shown that $|\langle \varphi, \psi \rangle| < 2\epsilon$, and $\epsilon > 0$ was arbitrary here, so $\langle \varphi, \psi \rangle = 0$, as claimed.

Put differently, we have that $H_b \subset H_{ws}^\perp$. We know already that $H_{ws} = H_c$ (see the first part of this proof) and $H_c^\perp = H_{pp}$ (see (12.4)), so it follows that $H_b \subset H_{pp}$. Putting things together, we obtain that $H_b = H_{pp}$, and this completes the proof. $\Box$

The decomposition from Theorem 12.6 also induces decompositions of the other quantities that are involved here. This makes use of the fact that we have reducing subspaces. More specifically, we can write

$$T = T_{ac} \oplus T_{sc} \oplus T_{pp},$$

where, for example, $T_{ac} : H_{ac} \to H_{ac}$, $D(T_{ac}) = D(T) \cap H_{ac}$, and if $x \in D(T_{ac})$, then $T_{ac}x = Tx$. Or we can say that $T_{ac} = TP_{ac}$, but restricted to the smaller Hilbert space $H_{ac}$. We also define $\sigma_j(T) := \sigma(T_j)$, where $j = ac, sc, pp$. As in Chapter 10 (see especially Proposition 10.9), we then have that

$$\sigma(T) = \sigma_{ac}(T) \cup \sigma_{sc}(T) \cup \sigma_{pp}(T).$$

The union is not necessarily disjoint, as the following Example shows. We will make use of the following fact, which is of considerable independent interest.

**Proposition 12.12.** The (pure) point spectrum, $\sigma_{pp}$, is the closure of the set of eigenvalues.

Recall that (much) earlier, we introduced the notation $\sigma_p$ for the set of eigenvalues; so we can now say that $\sigma_{pp} = \overline{\sigma_p}$. In particular, $\sigma_{pp}$ can be much larger than $\sigma_p$.

**Proof.** $H_{pp}$ contains all eigenvectors, so every eigenvalue of $T$ is an eigenvalue of $T_{pp}$ also, so $\sigma_p \subset \sigma_{pp}$. Since $\sigma_{pp}$, being a spectrum, is closed, we in fact obtain that $\overline{\sigma_p} \subset \sigma_{pp}$.

Conversely, if $t \in \sigma_{pp} = \sigma(T_{pp})$, then $E((t - r, t + r))x = x$ for some $x = x_\mu \in H_{pp}$ for all $r > 0$, but $\mu_x$ is a pure point measure if $x \in H_{pp}$ and an $s \in \mathbb{R}$ with $E(\{s\}) \neq 0$ is an eigenvalue by Exercise 12.11, so this means that there are eigenvalues arbitrarily close to $t$, so $t \in \overline{\sigma_p}$. $\Box$
Example 12.1. This is in fact very easy from an abstract point of view. We can start out with, let’s say, an absolutely continuous operator on one Hilbert space and a pure point operator on a second space and then assemble these by taking the orthogonal sum. We just need to make sure that the spectra were not disjoint to start with.

For example, let \( H_0 = L^2(0,1) \), \((T_0 f)(x) = xf(x)\). Note that \( T_0 \in B(H_0) \), \( T_0 = T_0^* \). In fact, \( T_0 \) is already given as multiplication by the variable, as in a Spectral Representation, so the spectral theory of \( T_0 \) is easy to work out. We claim that \( H_{ac} = H_0 \) for this operator. Indeed, if \( f \in H_0 \) is arbitrary, then

\[
\mu_f(M) = \|E(M)f\|^2 = \int |\chi_M f|^2 \, dx.
\]

In this context, recall the discussion following Proposition 10.10, where the spectral resolution was identified as \( E(M)f = \chi_M f \). So \( d\mu_f = |f|^2 \, dx \), and, as claimed, all spectral measures are absolutely continuous. Finally, notice that \( \sigma(T_0) = [0,1] \).

Next, let \( H_1 \) be an arbitrary separable Hilbert space. Let \( \{x_n\} \) be an ONB of \( H_1 \), and fix an enumeration \( \{q_n\} \) of \( \mathbb{Q} \cap [0,1] \). Define \( T_1 x_n = q_n x_n \) and extend linearly.

Exercise 12.15. Prove that \( T_1 \) is bounded on \( L(x_n) \) and thus extends continuously to \( L(x_n) = H_1 \). Show that (the extended) \( T_1 \in B(H_1) \), \( H_1 = H_{pp}, \sigma(T_1) = [0,1] \).

We can now let \( H = H_0 \oplus H_1 \), \( T = T_0 \oplus T_1 \). Then, basically by construction, \( T_0 \) and \( T_1 \) are the absolutely continuous and pure point parts, respectively, of \( T \). In particular, \( \sigma_{ac}(T) = \sigma_{pp}(T) = [0,1] \).

Exercise 12.16. Let \( T \) be self-adjoint. Prove that every isolated point of \( \sigma(T) \) is an eigenvalue of \( T \).

Exercise 12.17. Let \( H = \mathbb{C}^n \) be a finite-dimensional Hilbert space. Show that then \( H_{pp} = H, H_{ac} = H_{sc} = \{0\} \) for every self-adjoint \( T \) on \( H \).

Exercise 12.18. Let \( P \) be the projection onto the closed subspace \( M \subset H \), and let \( T \in B(H) \) be self-adjoint, \( T = \int t \, dE(t) \). Prove that \( M \) is a reducing subspace for \( T \) if and only if \( PE(B) = E(B)P \) for all Borel sets \( B \subset \mathbb{R} \).

Hint: By Exercise 11.21(a), \( M \) is reducing if and only if \( PT = TP \).

Remark: The statement also holds for unbounded self-adjoint \( T \), but the proof becomes more technical in this case.
Exercise 12.19. Let $\mu$ be a finite (positive) Borel measure on $\mathbb{R}$, with Lebesgue decomposition $\mu = \mu_{ac} + \mu_{sc} + \mu_{pp}$. Consider the (self-adjoint) operator of multiplication by the variable on $L^2(\mathbb{R}, \mu)$:

$$D(M_t) = \{ f \in L^2(\mu) : tf(t) \in L^2(\mu), \quad (M_t f)(t) = tf(t)\}$$

Prove that $H_j$ may be naturally identified with $L^2(\mathbb{R}, \mu_j)$, for $j = ac, sc, pp$. More precisely, proceed as follows: Pick disjoint supports $S_j$ of $\mu_j$, and show that $P_j f = \chi_{S_j} f$, where $P_j$ denotes the projection onto $H_j$ ($j = ac, sc, pp$). Given this, we then indeed have that

$$H_j = \{ f \in L^2(\mu) : f = 0 \mu\text{-a.e. on } S_j^c \},$$

which may be identified with $L^2(S_j, \mu)$ and also with $L^2(\mathbb{R}, \mu_j)$. 