An introduction to Quantum Mechanics for mathematicians

Daniel Ueltschi

Version of 14 May 2020

I am grateful to **Peter Mühlbacher** for useful remarks and for correcting many typos.

Contents

CHAPTER 1

Introduction

1. Origins of Quantum Mechanics

This account is based on the beautiful introduction in the famous textbook of Albert Messiah [**6**]; the interested reader can find a longer description there.

Until around 1900, Classical Theory reigned supreme. The physical world consists of *matter* (corpuscular) and of *radiation* (wave-like behaviour). Let M denote the "phase space", that is, the set of all possible states of the physical system. The evolution between times 0 and T is described by a trajectory $[0,T] \to M$ that satisfies a first-order differential equation, such as Hamilton equation for point particles or Maxwell equations for electromagnetic fields. It was then believed that electromagnetic waves take place in some medium called "ether", which might have been a form of matter. It looked like classical theory could describe the whole physical world.

1887: Mickelson and Morley design a precise experiment in order to measure the velocity of the Earth through ether: the result is surprising, there is no velocity!

This result was upsetting, as the prospect of the Earth carrying ether with it did not see appealing. The concept of ether was totally abandoned in 1905 with Einstein's article on special relativity.

At about the same time, the existence of atoms was being confirmed. Thomson had detected the electron in 1897. Einstein and Smochulowski had explained the Brownian motion using the atomic hypothesis, and had found estimates of Avogadro's number (1905).

1896: Discovery of radioactivity. This is an important step, conceptually, and it is even more important experimentally.

1911: Rutherford studies the scattering of α-particles (i.e. He^{++}) by atoms, thus gaining much more information on the properties of the atoms. These consist of a nucleus and of electrons that "gravitate" around it.

It should be remarked that the Coulomb force $\vec{F} = q\vec{E}$) and the Lorentz force ($\vec{F} = q\vec{v} \times \vec{B}$) for charged particles in electric and magnetic fields were well understood.

1900: Max Planck's theory of blackbody radiation. In order to explain the observed distribution of frequencies, he postulates that the exchange of

energy ε_ν between matter and radiation is **quantised**:

$$\varepsilon_\nu = h\nu.$$

Here, ν is the frequency and the new constant h is equal to $6.626 \cdot 10^{-34}$ J·s.

For most physicists, Planck's theory was a "lucky mathematical artifice". But it got further confirmations, notably by Einstein's explanation of the photoelectric effect in 1905 (Einstein got his Nobel prize for this).

1924: The Compton effect, about the scattering of photons by electrons. It is observed that the change in wavelength $\Delta\lambda$ only depends on physical constants and on the angle θ of observation:

$$\Delta\lambda = \tfrac{2h}{mc} \sin^2 \tfrac{\theta}{2}.$$

Planck's constant appears again! Compton and Debye showed that this formula can be explained if photons make a single elastic collision with an electron. It confirms Planck's law and the corpuscular nature of light. On the other hand, experiments about diffraction and interference of light shows that its evolution is wave-like. A purely corpuscular theory of light does not work.

There are further intriguing facts, such as the spectra of light emission and absorption of atoms. Hydrogen is characterised by Balmer's formula

$$\nu = \tfrac{me^4}{2\hbar^2}\left(\tfrac{1}{n^2} - \tfrac{1}{m^2}\right),$$

with $m, n \in \mathbb{N}, m > n$. Here, $\hbar = h/2\pi$. Bohr suggested that atoms have "energy levels", $E_n = -\tfrac{me^4}{2\hbar^2}\,\tfrac{1}{n^2}$, $n \in \mathbb{N}$, in the case of Hydrogen.

1922: The experiment of Stern and Gerlach about atoms carrying magnetic moments ("spins") and moving through a magnetic field. The deviation of the trajectory appears to be *quantised*, which cannot be explained by a classical theory.

1923: De Broglie suggests that matter has also wave-like behaviour (this should be contrasted to Einstein's suggestion that light has also corpuscular behaviour). This can be summarised by the existence of of a "wave-corpuscle duality.

1927: Experiments on the diffraction of electrons by Davisson and Gerner. Even when electrons are sent one by one, their evolution is wave-like.

The years 1910–1925 correspond to the "Old Quantum Theory" whose central tenets are the "Bohr-Sommerfeld quantisation rules". These are ad hoc hypotheses made in order to fit experimental data. In order to apply them correctly, one needs to make some arbitrary decisions. They are useful nonetheless as they allow to describe many different experiments in a unified way.

From 1923 to 1927, Heisenberg and Schrödinger formulate the basic setting of Quantum Mechanics. Dirac also made major contributions. Later, the mathematical framework was much enhanced by von Neumann. The influence of

Quantum Mechanics on the development of mathematics is huge, it was the main motivation for functional analysis.

2. A short survey of Classical Mechanics

In this survey we choose a mathematical structure so that Quantum Mechanics can be seen as a rather natural generalisation. This material is very well presented in Vassili Gelfreich's lecture notes [**2**].

The *state* of a classical particle at time $t \in \mathbb{R}$ is described by its position $q(t)$ and momentum $p(t)$. Let $M \subseteq \mathbb{R}^d \times \mathbb{R}^d$ be the set of all states $\big(q(t), p(t)\big)$. The *energy* of the particle is given by the Hamilton function (or "hamiltonian") $H : M \to \mathbb{R}$. This function, which we assume to be differentiable, determines the evolution through *Hamilton equations*:

$$\dot{q}_i = \frac{\partial}{\partial p_i} H(p,q), \qquad\qquad \dot{p}_i = -\frac{\partial}{\partial q_i} H(p,q). \tag{1.1}$$

(As usual, $\dot{q}_i$ denotes the time-derivative of the ith component of $q(t) \in \mathbb{R}^d$.) An important special case is a particle moving in an external potential. With the potential function given by $V : \mathbb{R}^d \to \mathbb{R}$, the hamiltonian is $H(p,q) = \frac{p^2}{2m} + V(q)$. Then $\dot{q}_i = \frac{1}{m} p_i$ and $\dot{p}_i = -\frac{\partial V}{\partial q_i}$, which are Newton equations.

The evolution is described by the flow

$$\begin{aligned} G_t :& M \to M \\ & (q_0, p_0) \mapsto \big(q(t), p(t)\big). \end{aligned} \tag{1.2}$$

We now expand the formalism so as to link it with differential geometry, and so that it can be suitably extended to a quantum setting. Let us consider the algebra of "observables" $\mathcal{A} = C^\infty(M)$, the set of smooth functions $M \to \mathbb{R}$. It is equipped with pointwise addition and multiplication, so it is indeed an algebra. Rather than looking at the trajectory $\big(q(t), p(t)\big)$, we can consider the evolution of observables. Thus we introduce the evolution operator

$$\begin{aligned} U_t :& \mathcal{A} \to \mathcal{A} \\ & U_t f = f \circ G_t. \end{aligned} \tag{1.3}$$

The evolution operator satisfies the group structure

(i) $U_0 = \mathbb{1}$;
(ii) $U_s U_t = U_{s+t}$, $s, t, \in \mathbb{R}$;
(iii) $U_t^{-1} = U_{-t}$, $t \in \mathbb{R}$.

In addition, we have $U_t fg = U_t f \, U_t g$ for all $f, g \in \mathcal{A}$, so U_t is an automorphism of the algebra $\mathcal{A}$.

We define a **state** to be a normalised, positive, linear functional on M. It can be represented by a probability measure by Riez-Markov theorem. With

μ denoting the probability measure, and $\big(q(t), p(t)\big)$ the trajectory with initial conditions (q_0, p_0), the expectation of the observable f at time $t \in \mathbb{R}$ is equal to

$$\int_M f\big(q(t), p(t)\big) \mathrm{d}\mu(q_0, p_0) = \int_M U_t f \mathrm{d}\mu. \tag{1.4}$$

Among the relevant states are the Dirac measures δ_{q_0, p_0} that describe deterministic states.

We see that the evolution is characterised by $U_t f$. It is then natural to consider the trajectories in $\mathcal{A}$ given by $f_t = U_t f$ where f is a fixed initial observable of $\mathcal{A}$. The evolution equation makes use of the *Poisson bracket*, which is the bilinear map $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$ defined by

$$\{f, g\} = \sum_{i=1}^{d} \Big(\frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} \Big). \tag{1.5}$$

The Poisson bracket is antisymmetric and it satisfies

- the product rule: $\big\{f, \{g, h\}\big\} = g\{f, h\} + \{f, g\}h$;
- Jacobi's identity: $\big\{f, \{g, h\}\big\} + \big\{g, \{h, f\}\big\} + \big\{h, \{f, g\}\big\} = 0$.

(Then $\mathcal{A}$ is a Lie algebra.) It follows from Hamilton's equations that the trajectory f_t in $\mathcal{A}$ satisfies the differential equation

$$\tfrac{\mathrm{d}}{\mathrm{d}t} f_t = \{H, f_t\}. \tag{1.6}$$

As we see in the next chapter, the structure of Quantum Mechanics looks similar. The algebra of smooth functions is replaced by an algebra of operators, and the Poisson bracket by commutators. Then the evolution equation for quantum systems takes a form reminiscent of (1.6); see Eq. (2.33).

Exercise 1.1. *Show that the energy is conserved, that is, $H(q(t), p(t)) = H(q_0, p_0)$ where $(q(t), p(t))$ is the trajectory defined by Hamilton equations and starting at (q_0, p_0).*

Exercise 1.2. *Show that the evolution equation for observables, Eq. (1.6), indeed follows from the Hamilton equations (1.1).*

CHAPTER 2

Finite-dimensional quantum systems

The description of a quantum system involves a (complex) Hilbert space and operators on this Hilbert space. The Hilbert space for a quantum particle in physical space is $L^2(\mathbb{R}^3)$; it is infinite-dimensional and many relevant operators are unbounded. Before considering the theory in this case, we look at a simpler setting where the Hilbert space is finite-dimensional; that is, we consider $\mathbb{C}^n$ for some $n \in \mathbb{N}$.

1. Operators on $\mathbb{C}^n$

Let $\varphi, \psi \in \mathbb{C}^n$. The **inner product** is

$$\langle \varphi, \psi \rangle = \sum_{i=1}^{n} \overline{\varphi}_i \psi_i. \tag{2.1}$$

It is linear in the second variable and skew-linear in the first variable.[1] We say that ψ is perpendicular or orthogonal to φ, $\psi \perp \varphi$, if $\langle \psi, \varphi \rangle = 0$. The **norm** of a vector φ is $\|\varphi\| = \sqrt{\langle \varphi, \varphi \rangle}$. A collection of vectors $\{e_i\}_{i=1}^n$ in $\mathbb{C}^n$ is an **orthonormal basis** (or just a **basis**) if $\langle e_i, e_j \rangle = \delta_{i,j}$. Any vector $\varphi \in \mathbb{C}^n$ can be written as

$$\varphi = \sum_{i=1}^{n} \langle e_i, \varphi \rangle e_i. \tag{2.2}$$

An **operator** A is a linear map $\mathbb{C}^n \to \mathbb{C}^n$. The corresponding matrix with respect to the basis $\{e_i\}_{i=1}^n$ is $(A_{i,j})_{1 \le i,j \le n}$ where

$$A_{i,j} = \langle e_i, A e_j \rangle. \tag{2.3}$$

The **operator norm** of A is

$$\|A\| = \sup_{\varphi \neq 0} \frac{\|A\varphi\|}{\|\varphi\|} = \sup_{\|\varphi\|=1} \|A\varphi\|. \tag{2.4}$$

The **adjoint** of A is the operator A^* such that

$$\langle A^* \varphi, \psi \rangle = \langle \varphi, A\psi \rangle \tag{2.5}$$

[1] This is the usual convention in physics; one advantage is that linear functionals look nice, $f(x) = \langle f, x \rangle$.

for all $\varphi, \psi \in \mathbb{C}^n$. The adjoint always exists and is unique. Its matrix representation is

$$(A^*)_{i,j} = \overline{A_{j,i}}. \tag{2.6}$$

An operator A on a finite-dimensional Hilbert space is **hermitian** or **symmetric** or **self-adjoint** if $A = A^*$. The **trace** and the **determinant** of an operator A can be defined with the help of its matrix representation $(A_{i,j})_{1 \le i,j \le n}$. Namely

$$\operatorname{Tr} A = \sum_{i=1}^{n} A_{i,i} \tag{2.7}$$

and

$$\det A = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} A_{i,\sigma(i)}; \tag{2.8}$$

the sum is over permutations of n elements. Trace and determinant do not depend on the choice of the basis. One can check that

$$\operatorname{Tr} AB = \operatorname{Tr} BA, \qquad \det AB = \det A \det B. \tag{2.9}$$

An operator A is **invertible** if there exists an operator A^{-1} such that $AA^{-1} = A^{-1}A = \mathbb{1}$. It is known that A is invertible if and only if $\det A \neq 0$ and that its inverse is unique. An operator U is **unitary** if it is invertible and $U^{-1} = U^*$. Equivalently, U is unitary if and only if $\|U\varphi\| = \|\varphi\|$ for all vectors φ. Any unitary operator can be written as $U = \mathrm{e}^{\mathrm{i}A}$ with A hermitian.

If there exists $\varphi \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$ such that

$$A\varphi = \lambda\varphi, \tag{2.10}$$

then φ is called an **eigenvector** and λ an **eigenvalue** of A. Some important properties of eigenvalues and eigenvectors:

- An operator A on $\mathbb{C}^n$ has at most n distinct eigenvalues.
- The eigenvalues of a hermitian operator are all real; the eigenvalues of a unitary operator have all modulus 1.
- If $AA^* = A^*A$ (i.e. the operator A is **normal**), there exist eigenvectors that form an orthonormal basis of $\mathbb{C}^n$.

An operator A is **positive-definite**, noted $A \ge 0$, if $\langle \varphi, A\varphi \rangle \ge 0$ for all $\varphi \in \mathbb{C}^n$. One can show that positive-definite operators are necessarily hermitian with nonnegative eigenvalues.

An operator P is a **projector** if $P^2 = P$; it is an **orthogonal projector** if in addition $P = P^*$; its eigenvalues are 0 or 1. Given $\varphi \in \mathbb{C}^n$ with $\|\varphi\| = 1$, the orthogonal projector P_φ onto the one-dimensional subspace spanned by φ is given by

$$P_\varphi \psi = \langle \varphi, \psi \rangle \varphi. \tag{2.11}$$

If A is a normal operator and $(\varphi_i)_{i=1}^n$ is an orthonormal basis of eigenvectors with eigenvalues $(\lambda_i)_{i=1}^n$, then we have

$$A = \sum_{i=1}^{n} \lambda_i P_{\varphi_i}. \tag{2.12}$$

Finally, let us point out that the map

$$\begin{aligned} \mathcal{B} \times \mathcal{B} &\to \mathbb{C} \\ (A, B) &\mapsto \langle A, B\rangle = \operatorname{Tr} A^* B \end{aligned} \tag{2.13}$$

is an **inner product** on $\mathcal{B}$. Its Cauchy-Schwarz inequality is useful.

2. Quantum states

In the review of classical systems, we considered the algebra of observables given by smooth functions on the phase space. In Quantum Mechanics, observables are given by self-adjoint operators on a Hilbert space. Then we let $\mathcal{A}$ denote the vector space of hermitian operators on $\mathbb{C}^n$. The usual product of operators is not a multiplication in $\mathcal{A}$ because $(AB)^* = B^*A^*$ which is not always equal to AB, even when A and B are hermitian. A possibility is to introduce the following multiplication:

$$A \bullet B = \tfrac{1}{2}(AB + BA). \tag{2.14}$$

We are not going to use this product much. Instead, we mainly work with the algebra $\mathcal{B}$ of all operators on $\mathbb{C}^n$, with the usual product. The **commutator** of two operators $A, B \in \mathcal{B}$ is

$$[A, B] = AB - BA. \tag{2.15}$$

For $\alpha, \beta \in \mathbb{C}$ and $A, B, C \in \mathcal{B}$, the commutator satisfies the following properties.

(i) Antisymmetric: $[B, A] = -[A, B]$.
(ii) Bilinear: $[\alpha A + \beta B, C] = \alpha[A, C] + \beta[B, C]$, same with the second variable.
(iii) Product rule: $[A, B \bullet C] = [A, B] \bullet C + B \bullet [A, C]$.
(iv) Jacobi's identity: $\big[A, [B, C]\big] + \big[B, [C, A]\big] + \big[C, [A, B]\big] = 0$

The commutator of hermitian operators is not always hermitian. But we can consider the "quantum Poisson bracket"

$$\{A, B\} = \mathrm{i}[A, B]. \tag{2.16}$$

The quantum Poisson bracket of hermitian operators is hermitian; then $\mathcal{A}$ and $\mathcal{B}$ are Lie algebras.

A state is a real linear functional on the algebra of observables $\mathcal{A}$. But it is more convenient, and is actually equivalent, to consider complex linear functionals on $\mathcal{B}$ (see Exercise 2.2). We then define a **state** to be a map $\omega : \mathcal{B} \to \mathbb{C}$ that is

(i) linear: $\omega(\alpha A + \beta B) = \alpha\omega(A) + \beta\omega(B)$ for all $\alpha, \beta \in \mathbb{C}$ and $A, B \in \mathcal{B}$;
(ii) positive: $\omega(A^*A) \geq 0$ for all $A \in \mathcal{B}$;

(iii) normalised: $\omega(\mathbb{1}) = 1$.

It turns out that states can be represented by density operators, a very useful property. A **density operator** ρ is a positive-definite hermitian operator such that $\operatorname{Tr}\rho = 1$. Given a density operator, there corresponds the state

$$\omega(A) = \operatorname{Tr}\rho A. \tag{2.17}$$

The converse is also true, each state is represented by a density operator.

PROPOSITION 2.1 (Riesz representation of states). *Let ω be a state. Then there exists a unique density operator ρ such that $\omega(A) = \operatorname{Tr}\rho A$ for all $A \in \mathcal{B}$.*

PROOF. Recall the operator inner product defined in (2.13). This inner product turns $\mathcal{B}$ into a Hilbert space. The standard Riesz representation theorem implies that $\mathcal{B}$ is self-dual, that is, every linear functional $\omega : \mathcal{B} \to \mathbb{C}$ is represented by a unique operator ρ such that

$$\omega(A) = \langle \rho, A\rangle = \operatorname{Tr}\rho^* A. \tag{2.18}$$

There remains to check that ρ is a density operator. We have for all $\varphi \in \mathbb{C}^n$ with $\|\varphi\| = 1$ that

$$\langle \varphi, \rho^*\varphi\rangle = \operatorname{Tr} P_\varphi \rho^* = \omega(P_\varphi) = \omega(P_\varphi^2) \geq 0. \tag{2.19}$$

Then ρ^* is positive-definite; it is therefore hermitian so $\rho \geq 0$ as well. Finally, $1 = \omega(\mathbb{1}) = \operatorname{Tr}\rho\mathbb{1}$, so ρ is indeed a density operator. □

The set of states is *convex*: If ω_1, ω_2 are two states, the convex combination $\alpha\omega_1 + (1-\alpha)\omega_2$ is also a state for all $\alpha \in [0, 1]$. A state is **mixed** if it can be written as a convex combination of distinct states. A state is **pure** if it is not mixed; in other words, pure states are the extremal points of the convex set of states.

Given $\varphi \in \mathbb{C}^n$, the corresponding projector P_φ is a special case of a density operator, hence it gives a state. It is perhaps expected that this state is pure, and that all pure states are represented by projectors.

PROPOSITION 2.2. *A state is pure if and only if its density operator is equal to P_φ for some $\varphi \in \mathbb{C}^n$.*

PROOF. The density operator ρ of the state is hermitian and it can be written as

$$\rho = \sum_{i=1}^{n} \lambda_i P_{\varphi_i}, \tag{2.20}$$

where the φ_is form a basis of eigenvectors. This can be viewed as a convex combination of density operators. This shows that if ρ is not equal to a projector, then the corresponding state is mixed.

There remains to show that if $\rho = P_\varphi$ is a projector, then the state is pure.[2]

Assume that $\omega(\cdot) = \alpha\omega_1(\cdot) + (1-\alpha)\omega_2(\cdot)$ with $\alpha \in (0,1)$. By Proposition 2.1, there exist density matrices ρ_1, ρ_2 such that $\omega_i(\cdot) = \operatorname{Tr} \rho_i \cdot$ for $i = 1, 2$. Then

$$P_\varphi = \alpha\rho_1 + (1-\alpha)\rho_2. \tag{2.21}$$

We have

$$\begin{aligned} 1 = \operatorname{Tr}\rho &= \operatorname{Tr} P_\varphi = \operatorname{Tr} P_\varphi^2 \\ &= \operatorname{Tr} P_\varphi\big(\alpha\rho_1 + (1-\alpha)\rho_2\big) \\ &= \alpha\langle\varphi, \rho_1\varphi\rangle + (1-\alpha)\langle\varphi, \rho_2\varphi\rangle \\ &\leq \alpha\|\rho_1\| + (1-\alpha)\|\rho_2\|. \end{aligned} \tag{2.22}$$

The last inequality follows from Cauchy-Schwarz. Since $\|\rho_i\| \leq 1$, we find that $\|\rho_i\| = 1$ (so ρ_i is a projector) and that the inequality is actually an identity. Both ρ_1 and ρ_2 must then project onto φ, so we necessarily have $\rho_1 = \rho_2 = P_\varphi$. □

Often in Quantum Mechanics, the state of the system is given by a normalised vector φ in $\mathbb{C}^n$ (or in another Hilbert space) and the corresponding expectation of an observable A is $\langle\varphi, A\varphi\rangle$. This is a special case of the above setting; the state corresponding to φ is the pure state with density operator P_φ.

We now discuss the Heisenberg uncertainty principle. It is often formulated for pure states, which is interesting and relevant, but it can be extended to arbitrary states.

Consider an observable A and assume that the system is in the state ω. An experiment can yield many possible values. If the experiment is repeated (with the same initial state ω), the average of the outcomes converges to $\omega(A)$. It is natural to quantify the fluctuations with respect to the average by introducing the **standard deviation** σ_ω, defined by

$$\sigma_\omega(A)^2 = \omega(A^2) - \omega(A)^2 = \omega\big([A - \omega(A)\mathbb{1}]^2\big). \tag{2.23}$$

Classically, there exist states such that the standard deviations are zero for all observables (this corresponds to Dirac measures on the phase space). But this is no longer possible in Quantum Mechanics! This is a consequence of the Heisenberg principle, which gives a lower bound for the standard deviations of two observables, in terms of expectation of their commutator.

THEOREM 2.3 (Heisenberg Uncertainty Principle). *For any state ω on $\mathcal{B}$, and any operators $A, B \in \mathcal{B}$, we have*

$$\sigma_\omega(A)\sigma_\omega(B) \geq \tfrac{1}{2}\big|\omega([A,B])\big|.$$

[2] An older proof was more complicated but Peter Mühlbacher has suggested this simpler one.

PROOF. With ρ the density operator of the state ω, we have

$$\big|\omega([A,B])\big| \leq |\operatorname{Tr}\rho AB| + |\operatorname{Tr}\rho BA|. \tag{2.24}$$

Since ρ is positive-definite it has a square-root; using the Cauchy-Schwarz inequality for the operator inner product defined in (2.13), we have

$$|\operatorname{Tr}\rho AB|^2 = |\operatorname{Tr} B\rho^{1/2}\rho^{1/2}A|^2 \leq \operatorname{Tr} B\rho^{1/2}\rho^{1/2}B \operatorname{Tr} A\rho^{1/2}\rho^{1/2}A = \omega(A^2)\omega(B^2). \tag{2.25}$$

The same holds for $\operatorname{Tr}\rho BA$ and we obtain

$$\omega(A^2)\omega(B^2) \geq \tfrac{1}{4}\big|\omega([A,B])\big|^2. \tag{2.26}$$

We can substitute A with $A - \omega(A)\mathbb{1}$ and B with $B - \omega(B)\mathbb{1}$. The left side then involves standard deviations, and the commutator in the right side does not change. □

3. Quantum evolution

We now discuss the equations that govern the dynamics of quantum systems. We first consider the traditional framework of Quantum Mechanics where the possible states are described by unit vectors in a Hilbert space; we take it to be $\mathbb{C}^n$ in this chapter. The evolution to time t is given by a map $U_t : \mathbb{C}^n \to \mathbb{C}^n$ that needs to satisfy the following properties. For all $s, t \in \mathbb{R}$ and $\varphi \in \mathbb{C}^n$,

- U_t is linear;
- $\|U_t\varphi\| = \|\varphi\|$;
- $U_{s+t} = U_sU_t$ and $U_0 = \mathbb{1}$.

The last property means that the laws of motion do not change in time. Then U_t is unitary and it can be written as

$$U_t = \mathrm{e}^{-\mathrm{i}tH} \tag{2.27}$$

for some hermitian operator H. (This can be established using Stone's theorem, that can be found below in a more general context.) Denoting $\varphi_t = U_t\varphi_0$ the vector at time t, we easily check that it satisfies **Schrödinger equation**:

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t = H\varphi_t. \tag{2.28}$$

The evolution given by U_t carries over to quantum states. One can check that

$$P_{U_t\varphi} = U_tP_\varphi U_t^*. \tag{2.29}$$

(Indeed, $U_tP_\varphi U_t^*$ is clearly a projector onto some vector of $\mathbb{C}^n$. Then $U_tP_\varphi U_t^*U_t\varphi = U_t\varphi$, so this vector must be to $U_t\varphi$.) The evolution of an arbitrary quantum state follows by linearity. From (2.29), its density operator at time t is given by

$$\rho_t = U_t\rho_0U_t^*. \tag{2.30}$$

The expectation in the state ω_t at time t is then

$$\omega_t(A) = \operatorname{Tr} U_t \rho_0 U_t^* A = \operatorname{Tr} \rho_0 U_t^* A U_t = \omega_0(U_t^* A U_t). \tag{2.31}$$

If we decide to let the observables evolve rather than the state, we get that $\omega_t(A) = \omega_0(A_t)$ with

$$A_t = U_t^* A U_t. \tag{2.32}$$

The time-dependent operator A_t satisfies the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} A_t = \mathrm{i}[H, A_t]. \tag{2.33}$$

(Check it!) This equation can also be written using the quantum Poisson bracket, namely $\frac{\mathrm{d}}{\mathrm{d}t} A_t = \{H, A_t\}$. This is the same equation as in the classical case! The objects here are operators instead of smooth functions.

It is clear from the equations above that the hermitian operator H plays an important rôle. What is its meaning? Recall Noether's theorem in classical mechanics, that states that to each symmetry there corresponds a conserved observable. This also applies here. Now H is the conserved observable that appears because of time-invariance. By definition, this observable is *the energy*. From now on, we will refer to H as the **hamiltonian**.

A **stationary state** is a state that is invariant over time. Its density operator is constant; since

$$\frac{\mathrm{d}}{\mathrm{d}t} \rho_t = -\mathrm{i}[H, \rho_t] \tag{2.34}$$

(this follows from (2.30)), we see that a state is stationary if and only if its density operator commutes with the hamiltonian.

If the stationary state is pure, i.e. if its density operator is equal to P_φ for some $\varphi \in \mathbb{C}^n$, we have

$$H\varphi = H P_\varphi \varphi = P_\varphi H \varphi. \tag{2.35}$$

Then $H\varphi$ is an eigenvector of P_φ with eigenvalue 1, so $H\varphi = \lambda\varphi$ for some $\lambda \in \mathbb{C}$. This shows that φ is an eigenvector of H (and the eigenvalue λ is real since H is hermitian). Conversely, if the state is given by P_φ with φ an eigenvector of H, then $[H, P_\varphi] = 0$ and the state is stationary. To summarise, the stationary pure states are given by eigenvectors of the hamiltonian.

If the initial state φ_0 is eigenvector of the hamiltonian with eigenvalue λ, $H\varphi_0 = \lambda\varphi_0$, its evolution is given by $\varphi_t = \mathrm{e}^{-\mathrm{i}t\lambda}\varphi_0$. It is not time-invariant, but the corresponding projector P_{φ_t} is. This shows that quantum states are more convenient.

We conclude this chapter with a word on **measurement** in Quantum Mechanics. According to the Copenhagen interpretation, the physical system, initially described by the vector φ_0, evolves deterministically according to the Schrödinger equation until an experiment is performed. To the experiment is associated the observable A (a self-adjoint operator). Assume for simplicity that the spectrum

of A only consists of eigenvalues of multiplicity 1. Then the only possible outcomes of the experiment are these eigenvalues; the outcome is random and the probability of measuring the eigenvalue a is $|\langle\varphi_t, \psi_a\rangle|^2$, where φ_t is the vector at time t and ψ_a is the eigenvector of A with eigenvalue a. Further, the state of the system after the experiment if given by ψ_a.

This description is compatible with all experimental observations but it has rattled many physicists ever since it was formulated. Einstein did not like it, who claimed that the Moon is physically present even when nobody is staring at it. There is also an obvious contradiction: Evolution under the Schrödinger equation is smooth, that is, φ_t is continuous in t, but the experiment brings a discontinuity from φ_t to ψ_a. The experiment and its observer can be viewed as a single physical systems whose evolution is given by a Schrödinger equation, and is therefore continuous; there is a contradiction. In recent years physicists have attempted to solve this conundrum with the notion of "decoherence", which purports that large systems have less randomness and more continuity. The debate continues.

Exercise 2.1. *Exponential of an operator on $\mathbb{C}^n$.*

(i) *Show that the following Taylor series converges, so it can serve as the definition of e^A:*

$$\sum_{k\geq 0} \tfrac{1}{k!} A^k.$$

(ii) *Show that $(\mathrm{e}^A)^{-1} = \mathrm{e}^{-A}$, and that $\mathrm{e}^A\,\mathrm{e}^B = \mathrm{e}^{A+B}$ if $[A, B] = 0$.*

(iii) *Show that*

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{e}^{tA} = A\,\mathrm{e}^{tA}.$$

(iv) *Show that if A is a normal operator with eigenvectors $(\varphi_i)_{i=1}^n$ and corresponding eigenvalues $(\lambda_i)_{i=1}^n$, its exponential can be written as*

$$\mathrm{e}^A = \sum_{i=1}^n \mathrm{e}^{\lambda_i} P_{\varphi_i}.$$

Exercise 2.2. *Let $\omega : \mathcal{A} \to \mathbb{R}$ be a (real) linear functional on $\mathcal{A}$. Show that there exists a unique extension to a (complex) linear functional $\hat{\omega} : \mathcal{B} \to \mathbb{C}$. Hint: An operator can be uniquely decomposed as $A = B + \mathrm{i}C$ with B, C hermitian; indeed, take $B = \frac{1}{2}(A^* + A)$ and $C = \frac{\mathrm{i}}{2}(A^* - A)$.*

Exercise 2.3. *Give a simpler proof of the Heisenberg Uncertainty Principle (Theorem 2.3) in the case of pure states. The state is given by $\omega(A) = \langle\varphi, A\varphi\rangle$ for a $\varphi \in \mathbb{C}^n$ and it is enough to invoke the Cauchy-Schwarz inequality in $\mathbb{C}^n$.*

CHAPTER 3

Quantum particle in the continuum

1. Square-integrable functions, position and momentum operators

The description of a particle in the continuum involve an infinite-dimensional Hilbert space and is therefore mathematically — and physically — more intricate and interesting.

The state of a quantum particle in $\mathbb{R}^d$ is given by a complex-valued square-integrable function. Let $\mathcal{H} = L^2(\mathbb{R}^d)$ be the Hilbert space of (Lebesgue-) integrable functions with finite L^2 norm:

$$\|f\|^2 = \int_{\mathbb{R}^d} |f(x)|^2 \mathrm{d}x. \tag{3.1}$$

The inner product in $L^2(\mathbb{R}^d)$ is

$$\langle f, g\rangle = \int_{\mathbb{R}^d} \overline{f(x)} g(x) \mathrm{d}x. \tag{3.2}$$

As is well-known, $L^2(\mathbb{R}^d)$ is an infinite-dimensional Hilbert space; there exist countable bases (in other words, this Hilbert space is separable). The space C^∞_{c} of infinitely-many differentiable functions with compact support is dense in $L^2(\mathbb{R}^d)$.

The **Fourier transform** plays a special rôle in quantum mechanics. Let $\mathcal{F}$ be the operator that maps an *integrable* function f to

$$\mathcal{F}f(k) = \widehat{f}(k) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}kx} f(x) \mathrm{d}x. \tag{3.3}$$

A remarkable property is that it conserves the L^2 norm, that is, $\|\mathcal{F}f\| = \|f\|$ for all $f \in L^1 \cap L^2$. It is therefore a unitary map, so it is bounded and continuous, and it can be extended by continuity to the whole of $L^2(\mathbb{R}^d)$. If $\widehat{f}$ is an integrable function, the inverse Fourier transform is

$$(\mathcal{F}^{-1}\widehat{f})(x) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}kx} \widehat{f}(k) \mathrm{d}k. \tag{3.4}$$

Since $\mathcal{F}$ is unitary, we have **Plancherel formula**

$$\langle f, g\rangle = \langle \widehat{f}, \widehat{g}\rangle. \tag{3.5}$$

The state of a quantum particle is described by a function $\psi \in L^2(\mathbb{R}^d)$ with norm $\|\psi\| = 1$. The *probability density* of finding the particle at $x \in \mathbb{R}^d$ is equal to $|\psi(x)|^2$ (i.e. the probability of finding the particle in a set $\Lambda \subset \mathbb{R}^d$ is given by $\int_\Lambda |\psi(x)|^2 \mathrm{d}x$). Further, the probability density for the particle to have momentum $k \in \mathbb{R}^d$ is $|\widehat{\psi}(k)|^2$. We now identify the operators associated with position and momentum. We discuss below a heuristic justification towards the momentum operator.

We simplify the setting that taking $d = 1$; the extension to general d is straightforward. If X denotes the position operator, the average position is equal to $\langle \psi, X\psi\rangle$. It is also given by the expectation with respect to the probability density, so that

$$\langle \psi, X\psi\rangle = \int_{\mathbb{R}} x|\psi(x)|^2 \mathrm{d}x. \tag{3.6}$$

It follows that the position operator is given by the multiplication operator X such that

$$(X\psi)(x) = x\psi(x). \tag{3.7}$$

This presents a mathematical difficulty. While $\psi \in L^2(\mathbb{R})$, the function $x\psi(x)$ is not necessarily square-integrable. This difficulty occurs will all unbounded operators in infinite-dimensional spaces; it is resolved by specifying a **domain** $D(X) \subset L^2(\mathbb{R})$, which is a dense linear subspace of the Hilbert space. Here we can take

$$D(X) = \Big\{ f \in L^2(\mathbb{R}) : \int_{\mathbb{R}} x^2 |f(x)|^2 \mathrm{d}x < \infty \Big\}. \tag{3.8}$$

Then the operator X is a linear map $D(X) \to L^2(\mathbb{R})$. Regarding the momentum, and recalling that $\widehat{\frac{\mathrm{d}}{\mathrm{d}x} f}(k) = \mathrm{i}k\widehat{f}(k)$, the expectation in the state ψ is equal to

$$\int_{\mathbb{R}} k|\widehat{\psi}(k)|^2 \mathrm{d}k = \langle \widehat{\psi}, \widehat{-\mathrm{i}\tfrac{\mathrm{d}}{\mathrm{d}x}\psi}\rangle = \langle \psi, -\mathrm{i}\tfrac{\mathrm{d}}{\mathrm{d}x}\psi\rangle. \tag{3.9}$$

It follows that the operator associated to the momentum is

$$P = -\mathrm{i}\nabla \tag{3.10}$$

where $\nabla = \frac{\mathrm{d}}{\mathrm{d}x}$. This operator is also unbounded and it should be defined on a domain of functions. For now we can take the space of differentiable functions whose derivative is square-integrable.

We discuss a heuristic for the momentum operator. Let us introduce the **translation operator** T_a, $a \in \mathbb{R}$, by

$$(T_a f)(x) = f(x - a). \tag{3.11}$$

Then $\{T_a\}_{a\in\mathbb{R}}$ is a one-parameter group of unitary operators. By Stone's theorem, there exists a self-adjoint operator P such that

$$T_a = \mathrm{e}^{-\mathrm{i}aP}. \tag{3.12}$$

Conservation of momentum is associated with translation-invariance, so P is the momentum. If f is an analytic function, we have the Taylor series

$$f(x-a) = \sum_{n\geq 0} \frac{(-a)^n}{n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n} f(x) \doteq \mathrm{e}^{-a\frac{\mathrm{d}}{\mathrm{d}x}} f(x). \tag{3.13}$$

Then we indeed have $P = -\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}x}$. The identity $\doteq$ is mathematically dubious, but it is possible to construct the above objects rigorously (see Section 4.3), and to establish the identity.

If the particle moves in an external potential, the latter is represented by a **multiplication operator** V. Given a function $V : \mathbb{R} \to \mathbb{R}$, the corresponding multiplication operator is the map

$$\begin{aligned} V :& D(V) \subset L^2(\mathbb{R}) \to L^2(\mathbb{R}) \\ & (Vf)(x) = V(x) f(x). \end{aligned} \tag{3.14}$$

The domain of the operator V must be chosen so that it is dense in $L^2(\mathbb{R})$ and $Vf \in L^2(\mathbb{R})$ for all $f \in D(V)$.

Classically, the energy of a particle in an external potential V is the function $H(q,p) = \frac{1}{2m}p^2 + V(q)$. In Quantum Mechanics, the energy is given by the following **Hamilton operator**, or **hamiltonian**:

$$H = -\Delta + V, \tag{3.15}$$

where $\Delta = \nabla^2 = -P^2$ is the **laplacian**. Notice that we drop all physical constants such as Planck's constant and the mass of the particle; this amounts to a rescaling of space and time.

2. Schrödinger equation

The Schrödinger equation for a quantum particle in an external potential V is

$$\mathrm{i}\frac{\partial}{\partial t}\psi(x,t) = H\psi(x,t) = -\Delta\psi(x,t) + V(x)\psi(x,t), \tag{3.16}$$

with initial condition given by a function $\psi_0(x)$. A solution $\psi(x,t)$ is a **classical solution** if it satisfies the following regularity conditions:

- For each fixed x, $\psi(x,\cdot) \in C^1([0,\infty))$. Further, for each $s > 0$ there exist $\tau > 0$ and $g \in L^1(\mathbb{R})$ (they may depend on s) such that $|\frac{\partial}{\partial t}\psi(x,t)|^2 \leq g(x)$ for all $t \in [s-\tau, s+\tau]$ and all $x \in \mathbb{R}$.
- For each fixed $t \geq 0$, $\psi(\cdot,t) \in C^2(\mathbb{R})$ and $\Delta\psi(\cdot,t) \in L^2(\mathbb{R})$.
- For each fixed $t \geq 0$, $V(\cdot)\psi(\cdot,t) \in L^2(\mathbb{R})$.

We shall see later that classical solutions exist for a large class of potentials V and initial conditions. We now check that the L^2 norm of classical solutions is conserved (so that $|\psi(\cdot,t)|^2$ is a probability density for all times) and that the evolution is deterministic (for given initial conditions the solution is unique).

LEMMA 3.1. *If ψ is a classical solution of the Schrödinger equation, then $\int_{\mathbb{R}} |\psi(x,t)|^2 \mathrm{d}x$ is constant in time.*

PROOF. We have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}} |\psi(x,t)|^2 \mathrm{d}x &\overset{(a)}{=} \int_{\mathbb{R}} \frac{\partial}{\partial t}|\psi(x,t)|^2 \mathrm{d}x \\ &= \int_{\mathbb{R}} \Big[\frac{\partial}{\partial t}\overline{\psi(x,t)}\,\psi(x,t) + \overline{\psi(x,t)}\frac{\partial}{\partial t}\psi(x,t)\Big]\mathrm{d}x \\ &\overset{(b)}{=} \int_{\mathbb{R}} \Big[-\mathrm{i}\Delta\overline{\psi(x,t)}\,\psi(x,t) + \mathrm{i}V(x)|\psi(x,t)|^2 + \mathrm{i}\overline{\psi(x,t)}\Delta\psi(x,t) - \mathrm{i}V(x)|\psi(x,t)|^2\Big]\mathrm{d}x \\ &\overset{(c)}{=} \mathrm{i}\int_{\mathbb{R}} \Big[\nabla\overline{\psi(x,t)}\,\nabla\psi(x,t) - \nabla\overline{\psi(x,t)}\,\nabla\psi(x,t)\Big]\mathrm{d}x \\ &= 0. \end{aligned} \tag{3.17}$$

We used (a) e.g. [**1**, Theorem 2.27]; (b) the Schrödinger equation; (c) integration by parts. □

COROLLARY 3.2. *If φ, ψ are two classical solutions of the Schrödinger equation with the same initial condition, i.e. $\varphi(x,0) = \psi(x,0)$ for all $x \in \mathbb{R}$, then $\varphi(x,t) = \psi(x,t)$ for all x, t.*

PROOF. $\varphi - \psi$ is a classical solution and $\int |\varphi(x,0) - \psi(x,0)|^2 \mathrm{d}x = 0$. Then $\int |\varphi(x,t) - \psi(x,t)|^2 \mathrm{d}x = 0$ for all t by Lemma 3.1, so $\varphi = \psi$. □

The simplest case is the evolution of a free particle, i.e., without external potential ($V = 0$). Let $\psi_0(x)$ denote the initial state. We first solve the Schrödinger equation with formal calculations; once we write down the solution, we can verify that it satisfies all properties for being a classical solution.

The Schrödinger equation is $\mathrm{i}\frac{\partial}{\partial t}\psi(x,t) = -\Delta\psi(x,t)$. Writing ψ with the help of the inverse Fourier transform, $\psi(x,t) = \frac{1}{\sqrt{2\pi}}\int \mathrm{e}^{\mathrm{i}kx}\,\widehat{\psi}(k,t)\mathrm{d}k$, we get

$$\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}kx}\,\mathrm{i}\frac{\partial\widehat{\psi}(k,t)}{\partial t}\mathrm{d}k = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}kx}\,k^2\widehat{\psi}(k,t)\mathrm{d}k. \tag{3.18}$$

This must hold for all x, so that $\mathrm{i}\frac{\partial\widehat{\psi}(k,t)}{\partial t} = k^2\widehat{\psi}(k,t)$. The solution is $\widehat{\psi}(k,t) = \mathrm{e}^{-\mathrm{i}k^2t}\,\widehat{\psi}_0(k)$. This allows to write the candidate solution as

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}kx - \mathrm{i}k^2t}\,\widehat{\psi}_0(k)\mathrm{d}k. \tag{3.19}$$

Under suitable conditions on the initial state ψ_0, e.g. ψ_0 is a Schwartz function, the function above can be seen to be a classical solution. This establishes existence of a solution; uniqueness follows from Corollary 3.2.

The trajectory of a free classical particle is ballistic, that is, it has constant velocity. The quantum counterpart is more complicated. It is not possible to specify an initial condition with fixed position and velocity because of the Heisenberg uncertainty principle (see Theorem 3.4 below). The most classical states are the "wave packets" and they can be shown to have ballistic motion; see Exercise 3.3.

We now consider the *average* motion of a quantum particle in an external potential and shows that it follows the classical equations of motion. This is known as Ehrenfest equations. Given a classical solution $\psi(x,t)$ of the Schrödinger equation (3.16), let

$$\langle A\rangle(t) = \langle \psi(\cdot,t), A\psi(\cdot,t)\rangle \tag{3.20}$$

denote the average of the observable A at time t.

THEOREM 3.3 (Ehrenfest equations). *Let $\psi(x,t)$ be a classical solution of the Schrödinger equation (3.16) such that $\int |\psi(x,t)|^2 \mathrm{d}x = 1$, $\int |x|\,|\psi(x,t)|^2 \mathrm{d}x < \infty$, and $\int |\nabla V(x)|\,|\psi(x,t)|^2 \mathrm{d}x < \infty$. Then*

(i) $\frac{\mathrm{d}}{\mathrm{d}t}\langle X\rangle(t) = 2\langle P\rangle(t)$.
(ii) $\frac{\mathrm{d}}{\mathrm{d}t}\langle P\rangle(t) = -\langle \nabla V\rangle(t)$.
(iii) $\frac{\mathrm{d}}{\mathrm{d}t}\langle H\rangle(t) = 0$.

The first two equations are the counterparts of Hamilton's equations — the factor 2 in (i) is present because we have set $2m = 1$. In (ii) the average is taken with respect to the multiplication operator with function ∇V. Notice that the claim (ii) is not straightforward; a natural guess could have been $\frac{\mathrm{d}}{\mathrm{d}t}\langle P\rangle(t) = -\nabla V(\langle X\rangle(t))$. This theorem is easily extended to higher spatial dimensions.

PROOF. Calculations are essentially straightforward. For (i), we have

$$\begin{aligned}
\tfrac{\mathrm{d}}{\mathrm{d}t}\langle X\rangle(t) &\overset{(a)}{=} \int_{\mathbb{R}} x\big[\tfrac{\partial}{\partial t}\overline{\psi(x,t)}\,\psi(x,t) + \overline{\psi(x,t)}\tfrac{\partial}{\partial t}\psi(x,t)\big]\mathrm{d}x \\
&\overset{(b)}{=} \mathrm{i}\int_{\mathbb{R}} \Big[\nabla\big(x\,\psi(x,t)\big)\nabla\overline{\psi(x,t)} - \nabla\big(x\,\overline{\psi(x,t)}\big)\nabla\psi(x,t)\Big]\mathrm{d}x \\
&= \mathrm{i}\int_{\mathbb{R}} \nabla x\,\big[\nabla\overline{\psi(x,t)}\,\psi(x,t) - \overline{\psi(x,t)}\nabla\psi(x,t)\big]\mathrm{d}x \\
&\overset{(c)}{=} -2\mathrm{i}\int_{\mathbb{R}} \overline{\psi(x,t)}\,\nabla\psi(x,t)\mathrm{d}x \\
&= 2\langle P\rangle(t).
\end{aligned} \tag{3.21}$$

Identity (a) follows e.g. from [**1**, Theorem 2.27]; we get (b) exactly as in the proof of Lemma 3.1, using Schrödinger's equation and integrating by parts; for (c) we integrated by parts and used the fact that $P = -\mathrm{i}\nabla$ is a symmetric operator.

For (ii) the calculations are

$$\begin{aligned}
\tfrac{\mathrm{d}}{\mathrm{d}t}\langle P\rangle(t) &= -\mathrm{i}\tfrac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}} \overline{\psi(x,t)}\,\nabla\psi(x,t)\mathrm{d}x \\
&= -\mathrm{i}\int \big[\tfrac{\partial}{\partial t}\overline{\psi(x,t)}\,\nabla\psi(x,t) + \overline{\psi(x,t)}\nabla\tfrac{\partial}{\partial t}\psi(x,t)\big]\mathrm{d}x \\
&= \ldots \\
&= -\int_{\mathbb{R}} \nabla V(x)\,|\psi(x,t)|^2\mathrm{d}x \\
&= -\langle \nabla V\rangle(t).
\end{aligned} \tag{3.22}$$

The dots hide the substitution of $\mathrm{i}\frac{\partial}{\partial t}\psi(x,t)$ with $-\Delta\psi(x,t) + V(x)\psi(x,t)$ and integration by parts.

For (iii) we have

$$\begin{aligned}
\tfrac{\mathrm{d}}{\mathrm{d}t}\langle \psi(\cdot,t), H\psi(\cdot,t)\rangle &= \langle \tfrac{\partial}{\partial t}\psi(\cdot,t), H\psi(\cdot,t)\rangle + \langle \psi(\cdot,t), H\tfrac{\partial}{\partial t}\psi(\cdot,t)\rangle \\
&= 2\mathrm{Re}\,\langle H\psi(\cdot,t), \tfrac{\partial}{\partial t}\psi(\cdot,t)\rangle \\
&= -2\mathrm{Re}\,\mathrm{i}\langle H\psi(\cdot,t), H\psi(\cdot,t)\rangle \\
&= 0.
\end{aligned} \tag{3.23}$$

We used the assumptions on regularity in order to exchange the order of integration and derivatives; we also used the fact that H is a symmetric operator, which is checked in Exercise 3.1. □

3. Heisenberg Uncertainty Principle

The usual Heisenberg uncertainty principle states that a particle cannot be localised in space and simultaneously have a definite momentum. We present two results. Theorem 3.4 is similar to the one in the finite-dimensional case, Theorem 2.3, while Theorem 3.5 is a bit different. Both results can be viewed as interesting inequalities in Fourier analysis.

THEOREM 3.4 (Heisenberg Uncertainty Principle). *For any $f \in L^2(\mathbb{R})$ with $\|f\| = 1$, and any $x_0, k_0 \in \mathbb{R}$, we have*

$$\int_{\mathbb{R}} (x-x_0)^2 |f(x)|^2 \mathrm{d}x \int_{\mathbb{R}} (k-k_0)^2 |\widehat{f}(k)|^2 \mathrm{d}k \geq \frac{1}{4}.$$

Equivalently, we have

$$\|(X-x_0)f\| \cdot \|(P-k_0)f\| \geq \tfrac{1}{2}.$$

This theorem holds also in higher spatial dimensions. It quantifies the impossibility to localise the wave function simultaneously in space and momentum. Taking x_0, k_0 to be the average position and momentum, the left side is equal to

the product of the squares of the standard deviations. The theorem is optimal in the sense that the inequality is saturated by gaussian functions, i.e. there is equality if and only if

$$f(x) = a\,\mathrm{e}^{\mathrm{i}(x-x_0)k_0}\ \mathrm{e}^{-b(x-x_0)^2} \tag{3.24}$$

with $b>0$ and $a=\sqrt{b/\pi}$.

Proof. As in the finite-dimensional version, the proof is based on the Cauchy-Schwarz inequality. We can shift f so that $x_0=0$ and add a phase so that $k_0=0$. It is then enough to prove that $\|Xf\|\cdot\|Pf\|\geq\frac{1}{2}$ whenever $\|f\|=1$. Since $[X,P]=\mathrm{i}$, we have

$$\begin{aligned} 1 &= \langle f,f\rangle = -\mathrm{i}\langle f,XPf\rangle + \mathrm{i}\langle f,PXf\rangle \\ &= -\mathrm{i}\langle Xf,Pf\rangle + \mathrm{i}\langle Pf,Xf\rangle = 2\mathrm{Re}\,\mathrm{i}\langle Pf,Xf\rangle \\ &\leq 2\|Pf\|\,\|Xf\|. \end{aligned} \tag{3.25}$$

□

In a way, Theorem 3.4 is a special case of Theorem 2.3 with observables A,B such that $[A,B]=c\mathbb{1}$. But such an identity is impossible in finite-dimensional Hilbert space, as can be seen by taking the trace in both sides.

Our second uncertainty principle has some shortcomings — it applies in dimension $d=1$ only and it does not quite scale correctly. But it is easy to prove and has some nice applications, see Exercise 3.6.

Theorem 3.5 (Alternate Uncertainty Principle). *For any continuous $f\in L^2(\mathbb{R})$ with $\|f\|=1$, we have*

$$\sup_{x\in\mathbb{R}}|f(x)|^2 \leq \Big(\int_{\mathbb{R}} k^2|\widehat{f}(k)|^2\mathrm{d}k\Big)^{1/2}.$$

Proof. By the inverse Fourier transform, we have for any a,

$$f(x) = \frac{1}{\sqrt{2\pi}}\int \sqrt{k^2+a^2}\,\mathrm{e}^{\mathrm{i}kx}\,\widehat{f}(k)\frac{\mathrm{d}k}{\sqrt{k^2+a^2}}. \tag{3.26}$$

Then, using the Cauchy-Schwarz inequality,

$$\begin{aligned} |f(x)|^2 &\leq \frac{1}{2\pi}\Big(\int_{\mathbb{R}}(k^2+a^2)|\widehat{f}(k)|^2\mathrm{d}k\Big)\Big(\int_{\mathbb{R}}\frac{\mathrm{d}k}{k^2+a^2}\Big) \\ &= \frac{1}{2|a|}\int_{\mathbb{R}} k^2|\widehat{f}(k)|^2\mathrm{d}k + \frac{|a|}{2}. \end{aligned} \tag{3.27}$$

The optimal choice is $a^2=\int_{\mathbb{R}}k^2|\widehat{f}(k)|^2\mathrm{d}k$ and we get the upper bound. □

Exercise 3.1. *Check these properties of the position, momentum, and Hamilton operators:*

(i) *they are unbounded, i.e., their operator norms are infinite;*
(ii) *they are symmetric, $\langle f,Xg\rangle=\langle Xf,g\rangle$, same for P and H;*
(iii) *the commutator between X and P is*

$$[X,P]=\mathrm{i}\mathbb{1}. \tag{3.28}$$

Check that $(XP-PX)\psi=\mathrm{i}\psi$ for a set of functions that is dense in $L^2(\mathbb{R})$.

Exercise 3.2. *Show that the classical solution of the free particle in Eq. (3.19) can be written as*

$$\psi(x,t) = \frac{1}{\sqrt{4\pi\mathrm{i}t}}\int_{\mathbb{R}}\mathrm{e}^{\mathrm{i}(x-y)^2/4t}\,\psi_0(y)\mathrm{d}y. \tag{3.29}$$

(Hint: Fubini theorem and gaussian integration.)

Exercise 3.3. *Let ψ_0 be a "wave-packet" centred at x_0 with velocity v_0 and standard deviation σ; that is,*

$$\psi_0(x) = \Big(\frac{2}{\pi\sigma^2}\Big)^{1/4}\mathrm{e}^{\mathrm{i}v_0(x-x_0)}\ \mathrm{e}^{-(x-x_0)^2/\sigma^2}. \tag{3.30}$$

Show that the classical solution of the free Schrödinger equation with initial condition ψ_0 is

$$\psi(x,t) = \Big(\frac{2}{\pi\sigma^2}\Big)^{1/4}\Big(1+\frac{4\mathrm{i}t}{\sigma^2}\Big)^{-1/2}\mathrm{e}^{-\frac{1}{4}\sigma^2v_0^2}\ \exp\Big\{-\frac{1}{\sigma^2+4\mathrm{i}t}\Big[x-x_0-\tfrac{\mathrm{i}}{2}\sigma^2v_0\Big]^2\Big\}. \tag{3.31}$$

Check that the average position is

$$\int_{\mathbb{R}} x|\psi(x,t)|^2\mathrm{d}x = x_0+v_0t; \tag{3.32}$$

this is ballistic motion, as in classical motion! Finally, check that the standard deviation is equal to the square root of

$$\int_{\mathbb{R}}(x-x_0-v_0t)^2|\psi(x,t)|^2\mathrm{d}x = \sigma^2+t^2/\sigma^2. \tag{3.33}$$

We observe that the spread of the wave packet increases over time.

Exercise 3.4. *Show that both inequalities in Theorem 3.4 are indeed equivalent.*

Exercise 3.5. *(Step in the proof of Theorem 3.4.) Given $f\in L^2(\mathbb{R})$ and $x_0,k_0\in\mathbb{R}$, find a function g such that*

(i) $\|f\|=\|g\|$;
(ii) $\|(X-x_0)f\|=\|Xg\|$;
(iii) $\|(P-k_0)f\|=\|Pg\|$.

Then it is enough to show that $\|Xg\|\,\|Pg\| \geq \frac{1}{2}$ *in the proof of Theorem 3.4.*

EXERCISE 3.6. *Give an example of a potential* $V : \mathbb{R} \to \mathbb{R}$ *that is continuous on* $\mathbb{R} \setminus \{0\}$ *with* $\lim_{x\to 0} V(x) = -\infty$*, but such that the hamiltonian* $-\Delta + V$ *is bounded below.*

This exercise is related to the question of stability of matter: Why is the energy of electrons orbiting a nucleus bounded below, even though the Coulomb potential is not bounded? This prevents us from pumping arbitrarily much energy out of a single atom!

Hint: Use Theorem 3.5 and get a lower bound that involves $\int |V|$*, which could be finite.*

CHAPTER 4

The evolution operator

The Schrödinger equation is $\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\psi(t) = H\psi(t)$ where the hamiltonian is the (unbounded) Schrödinger operator $H = -\Delta + V$. Formally, a solution to the Schrödinger equation is $\psi(t) = \mathrm{e}^{-\mathrm{i}tH}\,\psi(0)$. We would like to make sense of the operator $\mathrm{e}^{-\mathrm{i}tH}$. It should be unitary, so it is bounded and can be applied to all initial conditions $\psi(0) \in L^2(\mathbb{R}^d)$, even quite irregular ones.

We need some notions of analysis and functional analysis.

1. Lebesgue space $L^2(\mathbb{R}^d)$, weak derivatives, Sobolev spaces

The Lebesgue space $L^2(\mathbb{R}^d)$ is the linear space of all measurable functions $f : \mathbb{R}^d \to \mathbb{C}$ such that

$$\int_{\mathbb{R}^d} |f(x)|^2 \mathrm{d}x < \infty. \tag{4.1}$$

This is a Hilbert space with the inner product

$$\langle f, g\rangle = \int_{\mathbb{R}^d} \overline{f(x)}\, g(x)\, \mathrm{d}x. \tag{4.2}$$

The L^2 norm is then given by $\|f\|^2 = \langle f, f\rangle = \int_{\mathbb{R}^d} |f(x)|^2 \mathrm{d}x$. (More precisely, functions are defined up to sets of zero measure, so that $\|f\| = 0$ implies that $f = 0$.) The space $L^2(\mathbb{R}^d)$ is *separable*, meaning that there exists a dense countable subset; this also means that there exist countable orthonormal bases.

The set C_{c}^∞ of infinitely-differentiable functions with compact support is dense. Another useful dense subspace is the Schwartz space of infinitely-differentiable functions that decay fast; more precisely, a function $f \in C^\infty(\mathbb{R}^d)$ is a **Schwartz function** if for all $\alpha \in \mathbb{N}^d$ and all $k \in \mathbb{N}$, we have

$$\sup_{x\in\mathbb{R}^d} \|x\|^k \Big| \frac{\partial^{\alpha_1+\alpha_d}}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} f(x) \Big| < \infty. \tag{4.3}$$

As is well-known, the Fourier transform $f \mapsto \widehat{f}$ is a bijection from the Schwartz space into itself.

The **weak derivative** of f in the direction $j = 1, \dots, d$ is the function $\frac{\partial f}{\partial x_j}$ that satisfies

$$\int_{\mathbb{R}^d} \phi(x)\, \frac{\partial f}{\partial x_j}(x)\, \mathrm{d}x = -\int_{\mathbb{R}^d} \frac{\partial \phi}{\partial x_j}(x)\, f(x)\, \mathrm{d}x \tag{4.4}$$

for all smooth function of compact support $\phi \in C_{\mathrm{c}}^\infty(\mathbb{R}^d)$. If the function $\frac{\partial f}{\partial x_j}$ exists, it is unique. The **Sobolev embedding theorem** implies that

$$f \in C^k \quad \Longrightarrow \quad f \text{ has } k \text{ weak derivatives} \quad \Longrightarrow \quad f \in C^{k-1}. \tag{4.5}$$

In particular, a weakly differentiable function is continuous.

The **Sobolev space** $H^1(\mathbb{R}^d)$ is the space of all $f \in L^2(\mathbb{R}^d)$ such that $\frac{\partial f}{\partial x_j}$ exists and belongs to $L^2(\mathbb{R}^d)$ for all $j = 1, \dots, d$. It is a Hilbert space with the Sobolev inner product

$$\langle f, g\rangle_{H^1} = \langle f, g\rangle_{L^2} + \sum_{j=1}^d \langle \tfrac{\partial f}{\partial x_j}, \tfrac{\partial g}{\partial x_j}\rangle_{L^2} = \int_{\mathbb{R}^d} (1 + \|k\|^2)\, \overline{\widehat{f}(k)}\, \widehat{g}(k)\, \mathrm{d}k. \tag{4.6}$$

We will also consider the Sobolev space H^2, the space of all functions whose weak first and second derivatives are L^2 functions. Equivalently, we can define the Sobolev spaces using the Fourier transform:

$$\begin{aligned} H^1(\mathbb{R}^d) &= \big\{ f \in L^2(\mathbb{R}^d) : (1 + \|k\|^2)^{1/2} \widehat{f} \in L^2(\mathbb{R}^d) \big\}, \\ H^2(\mathbb{R}^d) &= \big\{ f \in L^2(\mathbb{R}^d) : (1 + \|k\|^2) \widehat{f} \in L^2(\mathbb{R}^d) \big\}. \end{aligned} \tag{4.7}$$

2. Unbounded operators and their spectrum

While the concept of derivative can be generalised so it applies to functions that are not continuously differentiable, it cannot be generalised to arbitrarily irregular L^2 functions. The corresponding operators are only defined on a subspace of functions. This is a general feature of unbounded operators, which can only be defined on a dense domain.

Definition 4.1. *A* **densely-defined operator** *on $\mathcal{H}$ is a pair $(A, \mathcal{D}(A))$ where the* **domain** *$\mathcal{D}(A)$ is a dense linear subspace of $\mathcal{H}$, and A is a linear map $\mathcal{D}(A) \to \mathcal{H}$.*

An operator A has an adjoint A^*. In words, A^* is an operator $\mathcal{D}(A^*) \to \mathcal{H}$ such that $\langle f, Ag\rangle = \langle A^* f, g\rangle$ for all $f \in \mathcal{D}(A^*), g \in \mathcal{D}(A)$, and the domain of A^* is the largest possible. Notice that $\mathcal{D}(A^*)$ is not necessarily dense.

Definition 4.2. *Let $A : \mathcal{D}(A) \to \mathcal{H}$ where $\mathcal{D}(A)$ is dense in $\mathcal{H}$. The* **domain of the adjoint** *is*

$$\mathcal{D}(A^*) = \big\{ f \in \mathcal{H} : \exists h \in \mathcal{H} \text{ such that } \langle f, Ag\rangle = \langle h, g\rangle \ \forall g \in \mathcal{D}(A) \big\}. \tag{4.8}$$

If such an h exists (for given f), it is unique because $\mathcal{D}(A)$ is dense. Then the **adjoint** *is the operator A^* that assigns $A^* f = h$ to each $f \in \mathcal{D}(A^*)$.*

When A is bounded, Riesz representation theorem implies that $\mathcal{D}(A^*) = \mathcal{H}$. And if $\dim \mathcal{H} < \infty$, the adjoint of a matrix is the hermitian conjugate.

DEFINITION 4.3. *The operator $A : \mathcal{D}(A) \to \mathcal{H}$ is* **symmetric** *if*

$$\langle Af, g\rangle = \langle f, Ag\rangle \tag{4.9}$$

for all $f, g \in \mathcal{D}(A)$. The operator A is **self-adjoint** *if $A^* = A$.*

One can check that an operator A is symmetric if and only if A^* is an extension of A, i.e. $\mathcal{D}(A^*) \supset \mathcal{D}(A)$ and $A^* f = Af$ for all $f \in \mathcal{D}(A)$. It is usually easy to verify that an operator is symmetric, harder that it is self-adjoint. The example of the laplacian with boundary conditions is discussed in Exercise 4.2.

The spectrum of operators in infinite-dimensional Hilbert spaces is more complicated than that of matrices, which consists of a finite number of eigenvalues. The complex plane is now the disjoint union of the resolvent set, the point spectrum, the continuous spectrum, and the residual spectrum:

- The **resolvent set** $\rho(A)$ of the operator A is the set of complex numbers λ such that $A - \lambda\mathbb{1}$ is one-to-one, $\operatorname{ran}(A - \lambda\mathbb{1})$ is dense in $\mathcal{H}$, and the inverse operator $(A - \lambda\mathbb{1})^{-1} : \operatorname{ran}(A - \lambda\mathbb{1}) \to \mathcal{D}(A)$ is bounded.
- The **point spectrum** $\sigma_{\mathrm{p}}(A)$ is the set of complex numbers λ such that $A - \lambda\mathbb{1}$ is not one-to-one. That is, there exists $f \in \mathcal{D}(A) \setminus \{0\}$ such that $(A - \lambda\mathbb{1})f = 0$, or $Af = \lambda f$.
- The **continuous spectrum** $\sigma_{\mathrm{c}}(A)$ is the set of complex numbers λ such that $A - \lambda\mathbb{1}$ is one-to-one, $\operatorname{ran}(A - \lambda\mathbb{1})$ is dense, but $(A - \lambda\mathbb{1})^{-1}$ is unbounded.
- The **residual spectrum** $\sigma_{\mathrm{r}}(A)$ is the set of complex numbers λ such that $A - \lambda\mathbb{1}$ is one-to-one and $\operatorname{ran}(A - \lambda\mathbb{1})$ is not dense.

We accept without proof the next theorem about the spectrum of self-adjoint operators.

THEOREM 4.1. *If A is a self-adjoint operator, then $\sigma_{\mathrm{r}}(A) = \emptyset$, and $\sigma_{\mathrm{p}}(A) \cup \sigma_{\mathrm{c}}(A)$ is a closed subset in $\mathbb{R}$.*

3. Construction of the evolution operator

We now construct the operator $U_t = \mathrm{e}^{-\mathrm{i}tH}$ when H is a self-adjoint operator (possibly unbounded). This is not immediate and we first do it for bounded operators.

PROPOSITION 4.2. *Let A be a bounded operator on the Hilbert space $\mathcal{H}$ and $t \in \mathbb{C}$. Then*

(a) *The sequence of operators $\left(\sum_{k=0}^{n} \frac{(-\mathrm{i}t)^k}{k!} A^k\right)_{n\geq 0}$ is Cauchy.*

Since the space of bounded operators on $\mathcal{H}$ is complete (it is a Banach space), the following operator is well-defined for all fixed $t \in \mathbb{C}$:

$$U_t = \sum_{k\geq 0} \frac{(-\mathrm{i}t)^k}{k!} A^k.$$

(b) *For all $s, t \in \mathbb{C}$, we have $U_s U_t = U_{s+t}$.*
(c) *For all $t \in \mathbb{C}$, $\frac{\mathrm{d}}{\mathrm{d}t} U_t = -\mathrm{i}AU_t$.*

We assume in addition that A is self-adjoint. Then

(d) *For $t \in \mathbb{R}$, we have $U_t^* = U_{-t}$.*
(e) *U_t is a unitary operator for all $t \in \mathbb{R}$.*

PROOF. The proof is exactly as in the case of finite-dimensional Hilbert spaces. For (a), observe that, if $m < n$,

$$\sum_{k=0}^{n} \frac{(-\mathrm{i}t)^k}{k!} A^k - \sum_{k=0}^{m} \frac{(-\mathrm{i}t)^k}{k!} A^k = \sum_{k=m+1}^{n} \frac{(-\mathrm{i}t)^k}{k!} A^k. \tag{4.10}$$

The norm of the left side is then less than $\sum_{k\geq m+1} \frac{|t|^k}{k!} \|A\|^k$, which is as small as we want by taking m, n large. This proves (a).

For (b), we have

$$U_s U_t = \sum_{k,\ell\geq 0} \frac{(-\mathrm{i}s)^k}{k!} \frac{(-\mathrm{i}t)^\ell}{\ell!} A^k A^\ell = \sum_{n\geq 0} \frac{1}{n!} A^n \underbrace{\sum_{m=0}^{n} \frac{n!}{(n-m)!m!} (-\mathrm{i}s)^{n-m} (-\mathrm{i}t)^m}_{=(-\mathrm{i}(s+t))^n} = U_{s+t}. \tag{4.11}$$

(c) is easy because all series below converge uniformly in ε around 0.

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} U_t &= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} (U_{t+\varepsilon} - U_t) \\ &= \lim_{\varepsilon\to 0} \sum_{n\geq 0} \frac{(-\mathrm{i}(t+\varepsilon))^n - (-\mathrm{i}t)^n}{n!\varepsilon} A^n \\ &= -\mathrm{i} \sum_{n\geq 0} \frac{(-\mathrm{i}t)^{n-1}}{(n-1)!} A^n \\ &= -\mathrm{i}AU_t. \end{aligned} \tag{4.12}$$

(d) follows from the Taylor series; and (e) follows from (b) and (d) since $U_t^* U_t = U_{-t} U_t = U_0 = \mathbb{1}$. □

LEMMA 4.3. *Let A be a self-adjoint operator. For all $\delta \in \mathbb{R} \setminus \{0\}$, we have*

(a) $\big\|(A+\mathrm{i}\delta)^{-1}\big\| \leq \frac{1}{|\delta|}$.

(b) $\big\|A(A+\mathrm{i}\delta)^{-1}\big\| \leq 1$.

PROOF. We have

$$\|(A+\mathrm{i}\delta)^{-1}\| = \sup_{\substack{f\in\mathrm{ran}\,(A+\mathrm{i}\delta)\\ f\neq 0}} \frac{\|(A+\mathrm{i}\delta)^{-1}f\|}{\|f\|} = \sup_{\substack{g\in\mathcal{D}(A)\\ g\neq 0}} \frac{\|g\|}{\|(A+\mathrm{i}\delta)g\|}. \tag{4.13}$$

For any $g \in \mathcal{D}(A)$, we have

$$\begin{aligned}\|(A+\mathrm{i}\delta)g\|^2 &= \langle (A+\mathrm{i}\delta)g, (A+\mathrm{i}\delta)g\rangle \\ &= \|Ag\|^2 - \mathrm{i}\delta\langle g, Ag\rangle + \mathrm{i}\delta\langle Ag, g\rangle + \delta^2\|g\|^2 \\ &\geq \delta^2\|g\|^2.\end{aligned} \tag{4.14}$$

The claim (a) follows. For (b), we repeat the steps above to get

$$\|A(A+\mathrm{i}\delta)^{-1}\| = \sup_{\substack{f\in\mathrm{ran}\,(A+\mathrm{i}\delta)\\ f\neq 0}} \frac{\|A(A+\mathrm{i}\delta)^{-1}f\|}{\|f\|} = \sup_{\substack{g\in\mathcal{D}(A)\\ g\neq 0}} \frac{\|Ag\|}{\|(A+\mathrm{i}\delta)g\|}. \tag{4.15}$$

From (4.14) we also get $\|(A+\mathrm{i}\delta)g\| \geq \|Ag\|$ which leads to the claim (b). □

Let H a self-adjoint operator and $\lambda \in \mathbb{R}$. The **resolvent operator** of H is defined as

$$R_\lambda = (H - \mathrm{i}\lambda)^{-1}. \tag{4.16}$$

This allows to introduce **Yosida's approximation** of H as

$$H_\lambda = -\mathrm{i}\lambda H R_\lambda. \tag{4.17}$$

Notice that H_λ is bounded by Lemma 4.3 (b).

LEMMA 4.4. *For all $f \in \mathcal{D}(H)$, we have*

$$\lim_{\lambda\to\infty} H_\lambda f = Hf.$$

PROOF. We start by proving that $\lim_{\lambda\to\infty} HR_\lambda f = 0$ for all $f \in \mathcal{H}$. It is enough to prove it for $f \in \mathcal{D}(H)$, the general case follows from a continuity argument. We have

$$\lim_{\lambda\to\infty} \|HR_\lambda f\| = \lim_{\lambda\to\infty} \|R_\lambda(Hf)\| \leq \lim_{\lambda\to\infty} \frac{\|Hf\|}{|\lambda|} = 0. \tag{4.18}$$

We used Lemma 4.3 (a). From $\mathbb{1} = (H - \mathrm{i}\lambda)R_\lambda$ we get the identity

$$HR_\lambda = \mathbb{1} + \mathrm{i}\lambda R_\lambda. \tag{4.19}$$

Then, using Eqs (4.17) and (4.19), and the result we have just proved, we get

$$\lim_{\lambda\to\infty} \|H_\lambda f - Hf\| = \lim_{\lambda\to\infty} \|(-\mathrm{i}\lambda R_\lambda - \mathbb{1})Hf\| = \lim_{\lambda\to\infty} \|HR_\lambda Hf\| = 0. \tag{4.20}$$

□

Since H_λ is bounded, we can use Proposition 4.2 and define

$$U_t^{(\lambda)} = \mathrm{e}^{-\mathrm{i}tH_\lambda}. \tag{4.21}$$

We now check that this operator is bounded (since H_λ is not necessarily symmetric, this necessitates a proof). Moreover, its norm is uniform in λ.

LEMMA 4.5. *For all $\lambda \in \mathbb{R}$ and $t \geq 0$ we have $\|U_t^{(\lambda)}\| \leq 1$.*

PROOF. Using the identity (4.19) we have $H_\lambda = -\mathrm{i}\lambda + \lambda^2 R_\lambda$. Then

$$\|U_t^{(\lambda)}\| = \|\,\mathrm{e}^{-t\lambda}\,\mathrm{e}^{-\mathrm{i}t\lambda^2 R_\lambda}\,\| \leq \mathrm{e}^{-t\lambda} \sum_{n\geq 0} \frac{(t\lambda^2)^n}{n!} \underbrace{\|R_\lambda\|^n}_{\leq 1/\lambda^n} \leq 1. \tag{4.22}$$

We used Lemma 4.3 (a). □

THEOREM 4.6. *As $\lambda \to \infty$, $U_t^{(\lambda)}$ converges strongly[a] to a unitary operator U_t that satisfies $U_s U_t = U_{s+t}$ for all $s, t \in \mathbb{R}$, and $\frac{\mathrm{d}}{\mathrm{d}t}U_t = -\mathrm{i}HU_t = -\mathrm{i}U_t U$.*

[a] That is, $\lim_{\lambda\to\infty} \|(U_t^{(\lambda)} - U_t)f\| = 0$ for all $f \in \mathcal{H}$.

PROOF. We have

$$U_t^{(\lambda)} f - U_t^{(\mu)} f = -\int_0^t \frac{\mathrm{d}}{\mathrm{d}s} U_{t-s}^{(\lambda)} U_s^{(\mu)} f \,\mathrm{d}s. \tag{4.23}$$

Then, since $[H_\lambda, U_t^{(\mu)}] = 0$,[1]

$$\begin{aligned}\Big\|\frac{\mathrm{d}}{\mathrm{d}s} U_{t-s}^{(\lambda)} U_s^{(\mu)} f\Big\| &= \big\|U_{t-s}^{(\lambda)}(H_\mu - H_\lambda)U_s^{(\mu)} f\big\| \\ &= \big\|U_{t-s}^{(\lambda)} U_s^{(\mu)}(H_\mu - H_\lambda) f\big\| \\ &\leq \big\|(H_\mu - H_\lambda) f\big\|.\end{aligned} \tag{4.24}$$

It follows that

$$\big\|U_t^{(\lambda)} f - U_t^{(\mu)} f\big\| \leq t\big\|H_\mu f - H_\lambda f\big\| \to 0 \tag{4.25}$$

[1] This is a consequence of $[R_\lambda, R_\mu] = 0$, which in turn follows from the *resolvent identity* $R_\lambda - R_\mu = (\lambda - \mu)R_\lambda R_\mu$.

as $\lambda, \mu \to \infty$, for all $f \in \mathcal{D}(H)$. A continuity argument shows that $(U_t^{(\lambda)} - U_t^{(\mu)})f \to 0$ for all $f \in \mathcal{H}$. Then U_t exists, and it is bounded by the uniform boundedness theorem.[2]

The group property is immediate since it holds for $U_t^{(\lambda)}$. Finally, we have the identity

$$\frac{1}{h}\big(U_{t+h}^{(\lambda)} f - U_t^{(\lambda)} f\big) = -\frac{\mathrm{i}}{h}\int_t^{t+h} U_s^{(\lambda)} H f \mathrm{d}s. \tag{4.26}$$

Taking the limits $\lambda \to \infty$ then $h \to 0$, we get $\frac{\mathrm{d}}{\mathrm{d}t}U_t f = -\mathrm{i}HU_t f$ for all $f \in \mathcal{D}(H)$. □

The converse of Theorem 4.6 is Stone's theorem, which we accept without proof.

THEOREM 4.7 (Stone). *Let U_t be a strongly continuous one-parameter group of unitary operators on $\mathcal{H}$.*[a] *Then there exists a (densely-defined) self-adjoint operator A such that*

$$U_t = \mathrm{e}^{-\mathrm{i}tA}.$$

[a] That is, $U_s U_t = U_{s+t}$ and $\lim_{\varepsilon\to 0}\|U_{t+\varepsilon}f - U_t f\| = 0$ for all $f \in \mathcal{H}$.

4. Construction of self-adjoint operators

The goal now is to identify the correct domain so that the Schrödinger operator $H = -\Delta + V$ is self-adjoint. We first consider the laplacian Δ alone. Then we shall see a sufficient criterion that applies straightforwardly to bounded potentials and to other cases such as the hydrogen atom.

PROPOSITION 4.8. *Let $\mathcal{D}(\Delta) = H^2(\mathbb{R}^d)$. Then Δ is self-adjoint.*

PROOF. We work in the Fourier space where the laplacian is the multiplication operator $\widehat{f} \mapsto \|k\|^2\widehat{f}$. Recall the characterisation (4.7) of the Sobolev space H^2. The laplacian is clearly symmetric, so Δ^* is an extension of Δ. We need to show that if $f \in \mathcal{D}(\Delta^*)$, then $f \in H^2(\mathbb{R}^d) = \mathcal{D}(\Delta)$.

By definition of the domain of the adjoint, there exists $\widehat{g} \in L^2(\mathbb{R}^d)$ such that

$$\int_{\mathbb{R}^d} \|k\|^2\overline{\widehat{h}(k)}\widehat{f}(k)\mathrm{d}k = \int_{\mathbb{R}^d} \overline{\widehat{h}(k)}\widehat{g}(k)\mathrm{d}k \tag{4.27}$$

for all $\widehat{h}$ such that $\|k\|^2\widehat{h} \in L^2(\mathbb{R}^d)$. This implies that $\widehat{g}(k) = \|k\|^2\widehat{f}(k)$ almost everywhere. Since $\widehat{g} \in L^2(\mathbb{R}^d)$, we get that $f \in \mathcal{D}(\Delta)$. □

[2] Uniform boundedness or Banach-Steinhaus theorem: If (A_n) are bounded operators such that $\sup_n \|A_n f\| < \infty$ for all $f \in \mathcal{H}$, then $\sup_n \|A_n\| < \infty$.

Although unbounded operators are not continuous, they may be *closed*, which is a form of continuity. An operator is **closed** if its graph $\big\{(f, Af) : f \in \mathcal{D}(A)\big\}$ is closed in $\mathcal{H} \times \mathcal{H}$. Equivalently, a closed operator has the property that if $f_n \to f$ and $Af_n \to g$, then $f \in \mathcal{D}(A)$ and $Af = g$. It can be proved that all adjoint operators are closed.

PROPOSITION 4.9. *A closed symmetric operator A is self-adjoint if and only if there exists $\lambda \in \mathbb{C}$ such that $\lambda \in \rho(A)$ and $\overline{\lambda} \in \rho(A)$.*

PROOF. If A is self-adjoint, the resolvent set contains all the complex plane except some of the real line, so the "only if" direction is obvious. Conversely, if A is symmetric, its adjoint A^* is an extension. We need to show that $\mathcal{D}(A^*) \subset \mathcal{D}(A)$.

If $\lambda \in \rho(A)$, we have that $\operatorname{ran}(A - \lambda)$ is dense and $(A - \lambda)^{-1}$ exists and is bounded. Since A is closed, one can check that $\operatorname{ran}(A - \lambda) = \mathcal{H}$. Then for any $f \in \mathcal{D}(A^*)$, there exists $g \in \mathcal{D}(A)$ such that

$$(A^* - \lambda)f = (A - \lambda)g. \tag{4.28}$$

But $(A - \lambda)g = (A^* - \lambda)g$, so that

$$(A^* - \lambda)(f - g) = 0. \tag{4.29}$$

There is a theorem that states that $\mathcal{H} = \overline{\operatorname{ran} A} \oplus \ker A^*$ for all densely-defined A. Since $\overline{\operatorname{ran}(A - \overline{\lambda})} = \mathcal{H}$, we have $\ker(A^* - \lambda) = \{0\}$, so that $f = g$. Then $f \in \mathcal{D}(A)$. □

We can now prove a criterion that gives self-adjoint operators.

THEOREM 4.10. *Let A be self-adjoint and B be closed and symmetric, with $\mathcal{D}(A) \subset \mathcal{D}(B)$. Assume that there exist $\varepsilon < 1$ and $c \geq 0$ such that for all $f \in \mathcal{D}(A)$,*

$$\|Bf\| \leq \varepsilon\|Af\| + c\|f\|.$$

Then $T = A + B$, with domain $\mathcal{D}(T) = \mathcal{D}(A)$, is self-adjoint.

In particular, $A + B$ is self-adjoint when A is self-adjoint and B is bounded (and symmetric). See Exercise 4.4 for a non-trivial application that includes the Coulomb potential in $d = 3$.

PROOF. We show that $\mathrm{i}\lambda \in \rho(T)$ with $\lambda \in \mathbb{R}$, $|\lambda|$ large enough. The theorem then follows from Proposition 4.9. We have

$$T + \mathrm{i}\lambda = A + B + \mathrm{i}\lambda = \big(\mathbb{1} + B(A + \mathrm{i}\lambda)^{-1}\big)(A + \mathrm{i}\lambda). \tag{4.30}$$

We would like to invert $\mathbb{1} + B(A + \mathrm{i}\lambda)^{-1}$. We check that the norm of $B(A+\mathrm{i}\lambda)^{-1}$ is less than 1, so we can define the inverse using a geometric series. We have

$$\begin{aligned} \big\|B(A+\mathrm{i}\lambda)^{-1}f\big\| &\leq \varepsilon\big\|A(A+\mathrm{i}\lambda)^{-1}f\big\| + c\big\|(A+\mathrm{i}\lambda)^{-1}f\big\| \\ &\leq \varepsilon\|f\| + \frac{c}{|\lambda|}\|f\|. \end{aligned} \tag{4.31}$$

We used Lemma 4.3. We can choose λ so that $\varepsilon + \frac{c}{|\lambda|} < 1$. Then $T + \mathrm{i}\lambda$ has a bounded inverse, so that $\mathrm{i}\lambda \in \rho(T)$. □

Exercise 4.1. *Are the following functions in $L^2(\mathbb{R})$ weakly differentiable? If yes, find their weak derivative.*

(a) $f(x) = \mathrm{e}^{-|x|}$.
(b) $f(x) = |x|^{-1/3}$.
(c) $f(x) = \mathrm{e}^{-[x(1-x)]^{-1}}$ *for* $x \in (0,1)$, $f(x) = 0$ *for* $x \notin (0,1)$.

Exercise 4.2. *Let $\mathcal{H} = L^2([0,1])$ and $A_k = \frac{\mathrm{d}^2}{\mathrm{d}x^2}$, $k = 1,2,3,4$, with the following domains:*

- $\mathcal{D}(A_1) = \big\{f \in C^2([0,1]) : f(0) = f(1) = 0\big\}$.
- $\mathcal{D}(A_2) = C^2([0,1])$.
- $\mathcal{D}(A_3) = \big\{f \in H^2([0,1]) : f(0) = f(1) = 0\big\}$.
- $\mathcal{D}(A_4) = H^2([0,1])$.

Are these operators symmetric or self-adjoint? (It can be shown that $\mathcal{D}(A_3^*) = \mathcal{D}(A_3)$; you can assume it here.)

Exercise 4.3. *Check that the operator norms of the bounded operators A, B satisfy*

(a) $\|AB\| \leq \|A\|\,\|B\|$,
(b) $\|A^*\| = \|A\|$.

Exercise 4.4. *Let $d \leq 3$ and $V : \mathbb{R}^d \to \mathbb{R}$ a function satisfying $\int_{|V|>m} |V(x)|^2 < \infty$ for some $m \geq 0$. Prove that the Schrödinger operator $H = -\Delta + V$, with domain $\mathcal{D}(H) = H^2(\mathbb{R}^d)$, is self-adjoint.*

Check that the Coulomb potential, $V(x) = -1/\|x\|$ in dimension $d = 3$, satisfies the condition.

Hint: Use Theorem 4.10, the Sobolev inequality $\|f\|_\infty \leq C\|f\|_{H^2}$, and the simple inequality

$$\|Vf\|^2 \leq m^2\|f\|^2 + \|f\|_\infty^2 \int_{|V|>m} V^2(x)\mathrm{d}x.$$

CHAPTER 5

Two explicit examples

Exact solutions are useful and much of our intuition on Schrödinger operators is based on them. We seek the spectrum and the eigenvalues of $H = -\Delta + V$ for specific V. There are three particularly interesting cases:

- The infinite square well.
- The harmonic oscillator.
- The Coulomb potential.

We consider the first two examples in this chapter, and we will discuss the Coulomb potential in the next chapter.

It should be clear that $-\Delta = P^2 \geq 0$, that is, the kinetic energy is nonnegative.

1. Infinite square well

Formally, we choose

$$V(x) = \begin{cases} 0 & \text{if } 0 \leq x_j \leq L, j = 1, \dots, d; \\ \infty & \text{otherwise.} \end{cases} \tag{5.1}$$

More precisely, we consider the hamiltonian $H = -\Delta$ on $L^2([0, L]^d)$ with Dirichlet boundary conditions. This means that the domain is

$$\mathcal{D}(H) = \big\{ f \in H^2([0, L]^d) : f(x) = 0 \text{ if } x_j = 0, L \text{ for some } j = 1, \dots, d \big\}. \tag{5.2}$$

The quantum particle is trapped in the box $[0, L]^d$. In contrast to the classical situation, the possible energies form a discrete set.

We start with $d = 1$. The eigenvalue equation is

$$-\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi(x) = \lambda\psi(x). \tag{5.3}$$

The solutions are $\psi_n(x) = \sin \frac{\pi n x}{L}$, $\lambda_n = \frac{\pi^2 n^2}{L^2}$, for all $n \in \mathbb{N}$. The functions $\{\psi_n\}$ form an orthogonal basis of $L^2([0, L])$. Consequently, these are the only solutions and the spectrum of H is pure point, with

$$\sigma_{\mathrm{p}}(H) = \big\{ \tfrac{\pi^2 n^2}{L^2} : n \in \mathbb{N} \big\}. \tag{5.4}$$

The case of general space dimension d follows immediately. It is easy to check that these are eigenfunctions for all $n = (n_1, \dots, n_d) \in \mathbb{N}^d$:

$$\psi_{n_1,\dots,n_d}(x) = \prod_{j=1}^{d} \sin \frac{\pi n_j x_j}{L}, \tag{5.5}$$

with eigenvalues

$$\lambda_{n_1,\dots,n_d} = \frac{\pi^2}{L^2} \sum_{j=1}^{d} n_j^2. \tag{5.6}$$

These functions also form an orthogonal basis, since $L^2([0, L]^d) = \otimes_{j=1}^{d} L^2([0, L])$.

One can also consider the case of the finite square well, where

$$V(x) = \begin{cases} -V_0 & \text{if } 0 \leq x_j \leq L, j = 1, \dots, d; \\ 0 & \text{otherwise.} \end{cases} \tag{5.7}$$

Here, we assume that $V_0 \geq 0$. The hamiltonian is $H = -\Delta + V$ with domain $\mathcal{D}(H) = H^2(\mathbb{R}^d)$. One can check that $\sigma_{\mathrm{c}}(H) = [0, \infty)$ and that $\sigma_{\mathrm{p}}(H) \in (-V_0, 0)$. Details can be found in most textbooks, see e.g. [**3**].

2. The harmonic oscillator

We again start with $d = 1$. The Hilbert space is $\mathcal{H} = L^2(\mathbb{R}^d)$ and the hamiltonian is

$$H = -\Delta + x^2. \tag{5.8}$$

We do not discuss the domain explicitly, just noting that there is a unique choice that makes H self-adjoint. Before investigating its properties, let us consider general properties of such operators.

DEFINITION 5.1. *Let A be a self-adjoint operator on $\mathcal{H}$ and λ a complex number. A* **Weyl sequence** *for (A, λ) is a sequence (f_n) in $\mathcal{D}(A)$ such that*

(i) $\|f_n\| = 1$ *for all n.*
(ii) $\|(A - \lambda)f_n\| \to 0$ *as $n \to \infty$ (so f_n is close to be an eigenvector).*
(iii) $f_n \rightharpoonup 0$ *as $n \to \infty$. That is, $\langle f_n, g\rangle \to 0$ for all $g \in \mathcal{H}$. Then f_n does not converge to an eigenvector.*

THEOREM 5.1 (Weyl). *Let A be a self-adjoint operator on $\mathcal{H}$.*
(a) *If $\lambda \in \sigma_{\mathrm{c}}(A)$, there exists a Weyl sequence for (A, λ).*
(b) *If there exists a Weyl sequence for (A, λ), then $\lambda \in \sigma_{\mathrm{p}}(A) \cup \sigma_{\mathrm{c}}(A)$.*

PROOF. We start with (a). If $\lambda \in \sigma_{\mathrm{c}}(A)$, we have that $(A-\lambda)^{-1}$ is unbounded so there exists (g_n) in $\mathrm{ran}(A - \lambda)$ with $\|g_n\| = 1$ and $\|(A - \lambda)^{-1} g_n\| \to \infty$. Let

$$f_n = \frac{(A - \lambda)^{-1} g_n}{\|(A - \lambda)^{-1} g_n\|}. \tag{5.9}$$

By the definition of the continuous spectrum we have that $(A-\lambda)^{-1}$ is densely-defined, and for any h in its domain, we have

$$\langle f_n, h\rangle = \frac{1}{\|(A-\lambda)^{-1}g_n\|}\langle g_n, (A-\lambda)^{-1}h\rangle \to 0; \tag{5.10}$$

we used Cauchy-Schwarz to bound the last inner product. A continuity argument shows that $\langle f_n, h\rangle \to 0$ for all $h \in \mathcal{H}$.

For (b), we show that if (f_n) is a Weyl sequence and $\lambda \notin \sigma_{\mathrm{p}}(A)$, then $\lambda \in \sigma_{\mathrm{c}}(A)$. If $\lambda \notin \sigma_{\mathrm{p}}(A)$, then $(A-\lambda)^{-1}$ exists as an operator $\operatorname{ran}(A-\lambda) \to \mathcal{H}$. Let

$$g_n = \frac{(A-\lambda)f_n}{\|(A-\lambda)f_n\|}. \tag{5.11}$$

Then

$$1 = \|f_n\| = \|(A-\lambda)f_n\|\,\|(A-\lambda)^{-1}g_n\|, \tag{5.12}$$

so that $\|(A-\lambda)^{-1}g_n\| \to \infty$ as $n \to \infty$. Then $(A-\lambda)^{-1}$ is unbounded and $\lambda \in \sigma_{\mathrm{c}}(A)$. □

It is a fact that for any V bounded below, there exists a unique self-adjoint operator $H = -\Delta + V$ that is densely-defined. When V is bounded, we have see that $\mathcal{D}(H) = H^2(\mathbb{R}^d)$. There does not seem to be a nice definition in the general case, although $\mathcal{D}(H) = \{f \in L^2 : \|k\|^2\widehat{f} \in L^2 \text{ and } Vf \in L^2\}$ is a natural guess.

THEOREM 5.2. *Let V be a continuous function on $\mathbb{R}^d$ such that $V(x) \to \infty$ as $\|x\| \to \infty$. Then the spectrum of the self-adjoint operator $H = -\Delta + V$ consists of isolated eigenvalues $\lambda_1 < \lambda_2 < \dots$ such that $\lambda_n \to \infty$ as $n \to \infty$.*

Although we skip the proof of this theorem, we can reflect on why it is true. We know that its spectrum is given by $\sigma_{\mathrm{p}}(H) \cup \sigma_{\mathrm{c}}(H)$. If $\lambda \in \sigma_{\mathrm{c}}$, there exists a Weyl sequence (f_n) that converges weakly to 0. Then f_n must "converge to Dirac", or have faster and faster oscillations, or disappear at infinity. (Or a combination of the three.) In the first two cases, $\|\Delta f_n\| \to \infty$; in the third case, $\|Vf_n\| \to \infty$. It is impossible for $\|(-\Delta + V - \lambda)f_n\|$ to stay bounded, let alone to vanish. The proof that $\lambda_n \to \infty$ follows from the min-max theorem.

We now return to the harmonic oscillator. The Hilbert space is $L^2(\mathbb{R}^d)$ and the hamiltonian is

$$H = \tfrac{1}{2}(-\Delta + x^2). \tag{5.13}$$

We know from Theorem 5.2 that the spectrum consists of isolated eigenvalues. We rewrite the hamiltonian using the **creation** and **annihilation operators** a^* and a:

$$\begin{aligned} a^* &= \tfrac{1}{\sqrt{2}}(X - \mathrm{i}P), \\ a &= \tfrac{1}{\sqrt{2}}(X + \mathrm{i}P). \end{aligned} \tag{5.14}$$

We also define the operator $N = a^*a$. The hamiltonian (5.13) can then be written as

$$H = a^*a + \tfrac{1}{2} = N + \tfrac{1}{2}. \tag{5.15}$$

It is clear that $N \geq 0$, so $H \geq \frac{1}{2}$.

PROPOSITION 5.3. *The spectrum of N is pure point and $\sigma_{\mathrm{p}}(N) = \{0, 1, 2, 3, \dots\}$; each eigenvalue has multiplicity 1.*

PROOF. We know from Theorem 5.2 that the spectrum of H (and therefore N) consists of isolated eigenvalues. If $af = 0$, it is clear that $Nf = 0$; conversely, since $\langle f, Nf\rangle = \|af\|^2$, we see that $Nf = 0$ implies that $af = 0$. Consider the function

$$f_0(x) = \pi^{-1/4}\,\mathrm{e}^{-\frac{1}{2}x^2}\,. \tag{5.16}$$

One easily checks that $\|f_0\| = 1$, and that

$$af_0 = \tfrac{1}{\sqrt{2}}(X - \mathrm{i}P)f_0 = \tfrac{1}{\sqrt{2}\pi^{1/4}}\big(x + \tfrac{\mathrm{d}}{\mathrm{d}x}\big)\,\mathrm{e}^{-\frac{1}{2}x^2} = 0. \tag{5.17}$$

Then f_0 is eigenvector of N with eigenvalue 0. It follows from the theory of ODEs that f_0 is the unique solution of the equation $f_0' = -xf_0$ (with a condition such as $f_0(0) = \pi^{-1/4}$). The eigenvalue 0 has then multiplicity 1.

Next, let

$$f_n = c_n(a^*)^n f_0. \tag{5.18}$$

It is not hard to check that f_n are nonzero Schwartz functions, so they are in L^2 indeed; the constant c_n that normalise them is calculated in Exercise 5.2. Then

$$\begin{aligned} Nf_n &= c_n N(a^*)^n f_0 \\ &\overset{(a)}{=} c_n a^*(N+1)(a^*)^{n-1} f_0 \\ &= \dots = c_n(a^*)^n(N+n)f_0 \\ &\overset{(b)}{=} nf_n. \end{aligned} \tag{5.19}$$

We used (a) the relation $Na^* = a^*(N+1)$, see Exercise 5.1, and (b) $Nf_0 = 0$. Then f_n is eigenvector of N with eigenvalue n. This shows that σ_{p} contains all nonnegative integers.

Now suppose that $Nf = \lambda f$ with $\lambda \neq 0$. Then af is an eigenvector of N with eigenvalue $\lambda - 1$:

$$Naf = a(N-1)f = (\lambda - 1)af. \tag{5.20}$$

We used an identity of Exercise 5.1. Since $\|af\|^2 = \langle f, Nf\rangle = \lambda\|f\|^2$, we see that af is nonzero. We observe that if λ is a non-integer eigenvalue, then $\lambda-1, \lambda-2, \dots$ are all eigenvalues. Then N has negative eigenvalues, which is impossible since $N \geq 0$. We conclude that all eigenvalues are nonnegative integers.

In order to check that eigenvalues have multiplicity 1, observe that if f_n, g_n are orthogonal eigenvectors with eigenvalue n, then af_n, ag_n are eigenvectors with eigenvalue $n-1$, and they are also orthogonal:

$$\langle af_n, ag_n\rangle = \langle f_n, a^*ag_n\rangle = n\langle f_n, g_n\rangle = 0. \tag{5.21}$$

Iterating, we get that $a^n f_n, a^n g_n$ are orthogonal eigenvectors with eigenvalue 0. But we know that 0 has multiplicity 1, contradiction. Then all eigenvalues have multiplicity 1. □

We have proved:

THEOREM 5.4. *The spectrum of the self-adjoint operator* $H = \frac{1}{2}(-\Delta+x^2)$ *is*

$$\sigma(H) = \sigma_{\mathrm{p}}(H) = \left\{\tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \dots\right\}.$$

All eigenvalues have multiplicity 1.

The eigenvectors of the harmonic oscillator are related to the **Hermite polynomials** $H_n(x)$ that satisfy

$$H_n(x) = (-1)^n \,\mathrm{e}^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \,\mathrm{e}^{-x^2} \;=\; \mathrm{e}^{\frac{1}{2}x^2} \left(x - \frac{\mathrm{d}}{\mathrm{d}x}\right)^n \mathrm{e}^{-\frac{1}{2}x^2}. \tag{5.22}$$

EXERCISE 5.1. *The creation and annihilation operators defined in Eq. (5.14).*

(i) *Check that* $[a, a^*] = 1$. *Precisely, show that for any* f, g *in a dense set in* $L^2(\mathbb{R})$, *we have*

$$\langle f, a\, a^* g\rangle - \langle f, a^* a g\rangle = \langle f, g\rangle.$$

(ii) *Let* $N = a^*a$ *and check that*

$$Na = a(N-1),$$
$$Na^* = a^*(N+1).$$

EXERCISE 5.2. *Calculate the constant* c_n *that makes the functions* f_n *in (5.18) normalised,* $\|f_n\| = 1$.

CHAPTER 6

The hydrogen atom

The greatest triumph of quantum mechanics is arguably its description of the energy levels of the electron in the hydrogen atom. The problem is mathematically interesting and nontrivial. More importantly, its resolution allowed for considerable advances in molecular chemistry and condensed matter physics.

It was observed in the 19th century that elements are characterised by their *spectrum*: The light emitted by excited atoms has only a limited of possible wavelengths. The mathematician Balmer, confident that there was mathematical beauty in all things, searched for a formula that gives the four rays of energy of visible light of hydrogen. His formula was extended by Rydberg. The possible energies are given by $\frac{me^4}{2\hbar^2}(\frac{1}{n^2}-\frac{1}{m^2})$ where $n<m$ are integers. This prompted physicists to look for further rays, which they found, thus confirming this intriguing formula.

The explanation took a while but it was eventually understood that, at rest, the electron of the hydrogen atom occupies the "ground state", i.e. the state of lowest energy of the corresponding Schrödinger operator. When receiving light, the electron may absorb a photon and reach an eigenstate with higher eigenvalue. This is an unstable state and the electron eventually relaxes to an eigenstate of lower energy, emitting a photon with precisely the energy difference. The corresponding wavelength is given by the Planck–Einstein relation $\varepsilon = h\nu = hc/\lambda$. We shall see that the point spectrum of the hamiltonian of the hydrogen atom consists of all values of the form $-\frac{me^4}{2\hbar^2}\frac{1}{n^2}$, with $n = 1, 2, 3, ...$ Balmer's formula corresponds to possible differences.

1. Stability of matter

The system consists of one proton and one electron. The mass of the electron is about 10^4 times that of the electron, so it is effectively infinite — that is, the proton is a fixed classical particle and the electron is the sole quantum particle. Energy involves kinetic energy given by the laplacian, and interaction energy given by the Coulomb potential. We neglect physical constants. The Hilbert space for the electron is $\mathcal{H} = L^2(\mathbb{R}^3)$ and the hamiltonian is

$$H = -\Delta - \frac{1}{\|x\|}. \tag{6.1}$$

Here, $\|x\| = (x_1^2 + x_2^2 + x_3^2)^{1/2}$ and $1/\|x\|$ acts as a multiplication operator. It is not hard to check that H is unbounded; we take $\mathcal{D}(H) = H^2(\mathbb{R}^3)$ so the operator is self-adjoint (see Exercise 4.4). The hamiltonian is a Schrödinger operator.

An immediate question is why should H be bounded below? That is, $\langle f, Hf\rangle$ cannot be arbitrarily negative. This is the simplest question of the more general problem of stability of matter. There are several ways to prove a lower bound for H and we consider two of them. One is an operator inequality for $-\Delta$. The other is a Sobolev inequality.

PROPOSITION 6.1. *Let $\mathcal{H} = L^2(\mathbb{R}^d)$ with $d \geq 3$. Then*

$$-\Delta \geq \frac{(d-2)^2}{4}\frac{1}{\|x\|^2}.$$

In other words, we have

$$\int_{\mathbb{R}^d} \|\nabla f(x)\|^2 \mathrm{d}x \geq \frac{(d-2)^2}{4}\int_{\mathbb{R}^d}\frac{1}{\|x\|^2}|f(x)|^2\mathrm{d}x \tag{6.2}$$

for all f in a dense set (we can take the set $H^1(\mathbb{R}^d)$). Here, the square of the norm of the gradient is $\|\nabla f(x)\|^2 = \sum_{i=1}^d |\frac{\partial f}{\partial x_i}(x)|^2$.

PROOF. In this proof, f is a smooth function $\mathbb{R}^d \to \mathbb{R}$. Recalling that $P_j = -\mathrm{i}\frac{\partial}{\partial x_j}$ and noting that $\frac{\partial}{\partial x_j}\frac{1}{\|x\|} = -\frac{x_j}{\|x\|^3}$, we can check that

$$\begin{aligned} [P_j, \tfrac{1}{\|x\|}] &= \mathrm{i}\tfrac{x_j}{\|x\|^3}, \\ \tfrac{1}{\|x\|^2} &= \tfrac{\mathrm{i}}{d}\sum_{j=1}^d [\tfrac{1}{\|x\|}P_j\tfrac{1}{\|x\|}, X_j]. \end{aligned} \tag{6.3}$$

It follows from the second equation that

$$\langle f, \tfrac{1}{\|x\|^2} f\rangle = -\tfrac{2}{d}\sum_{j=1}^d \mathrm{Im}\,\langle \tfrac{1}{\|x\|}P_j\tfrac{1}{\|x\|}f, X_j f\rangle. \tag{6.4}$$

We now use the first identity in (6.3) to get

$$\langle \tfrac{1}{\|x\|}P_j\tfrac{1}{\|x\|}f, X_j f\rangle = \langle P_j f, \tfrac{x_j}{\|x\|^2}f\rangle - \mathrm{i}\langle f, \tfrac{x_j^2}{\|x\|^4}f\rangle. \tag{6.5}$$

We also used the fact that $\frac{1}{\|x\|}$ is a symmetric operator. Summing over $j = 1, \ldots, d$, we obtain

$$(d-2)\langle f, \tfrac{1}{\|x\|^2}f\rangle = -2\mathrm{Im}\sum_{j=1}^d \langle P_j f, \tfrac{x_j}{\|x\|^2}f\rangle. \tag{6.6}$$

Then

$$\begin{aligned}\tfrac{(d-2)^2}{4}\big|\langle f, \tfrac{1}{\|x\|^2} f\rangle\big|^2 &= \Big|\mathrm{Im}\sum_{j=1}^d \langle P_j f, \tfrac{x_j}{\|x\|^2} f\rangle\Big|^2 \\ &\le \Big|\sum_{j=1}^d \langle P_j f, \tfrac{x_j}{\|x\|^2} f\rangle\Big|^2 \\ &\overset{(\star)}{\le} \Big(\sum_{j=1}^d \langle P_j f, P_j f\rangle\Big)\Big(\sum_{j=1}^d \langle \tfrac{x_j}{\|x\|^2} f, \tfrac{x_j}{\|x\|^2} f\rangle\Big) \\ &= \langle f, -\Delta f\rangle \, \langle f, \tfrac{1}{\|x\|^2} f\rangle.\end{aligned} \tag{6.7}$$

The $(\star)$ inequality is the Cauchy-Schwarz inequality for an L^2 inner product involving vectors of d functions. The claim follows immediately. □

From Proposition 6.1, we immediately obtain a lower bound for the hamiltonian (6.1), namely

$$\langle f, Hf\rangle \ge \langle f, (\tfrac{1}{4\|x\|^2} - \tfrac{1}{\|x\|}) f\rangle \ge -1. \tag{6.8}$$

Indeed, the bracket above is always greater than -1. As we see later, the lowest eigenvalue of H is $-\frac{1}{4}$, so this result is rather good.

For the second method, observe that

$$\langle f, Hf\rangle = \int_{\mathbb{R}^d} \Big(\|\nabla f(x)\|^2 + V|f(x)|^2\Big)\mathrm{d}x = \|\nabla f\|^2 + \langle f, Vf\rangle. \tag{6.9}$$

Here, $\|\nabla f(x)\|^2$ has the same meaning as in (6.2). In order to get a lower bound, we use a Sobolev inequality.

THEOREM 6.2 (Sobolev inequality for gradient). *Let $d \ge 3$ and let $q = \frac{2d}{d-2}$. Then for any $f \in H^1(\mathbb{R}^d)$, we have*

$$\|\nabla f\|^2 \ge S_d \|f\|_{L^q}^2$$

where $S_d = \frac{d(d-2)}{4}|\mathbb{S}^d|^{2/d} = \frac{d(d-2)}{4} 2^{2/d}\pi^{1+1/d}\Gamma(\frac{d+1}{2})^{-2/d}$.

The proof of this theorem can be found in Lieb and Loss' Analysis textbook, see [**5**, Theorem 8.3]. Notice that for $d = 3$ we have $q = 6$ and $S_3 = 3(\frac{\pi}{2})^{4/3}$. We use it to get a lower bound for the energy spectrum of hydrogen. With H from (6.1), we have

$$\begin{aligned}\langle f, Hf\rangle &\ge 3(\tfrac{\pi}{2})^{4/3}\|f\|_{L^6}^2 - \int_{\mathbb{R}^3} \tfrac{1}{\|x\|}|f(x)|^2\mathrm{d}x \\ &= 3(\tfrac{\pi}{2})^{4/3}\|\rho\|_{L^3} - \int_{\mathbb{R}^3} \tfrac{1}{\|x\|}\rho(x)\mathrm{d}x,\end{aligned} \tag{6.10}$$

with $\rho(x) = |f(x)|^2$. In Exercise 6.1, you can check that the last line is always greater than $-1/3$. This gives a better bound than the previous method, close to the optimal result $-1/4$.

2. Spherical coordinates

The hydrogen hamiltonian is obviously spherically symmetric and it is natural to use spherical coordinates. We write $(x, y, z) \in \mathbb{R}^3$ as $(r, \theta, \varphi) \in [0, \infty) \times [0, \pi] \times [0, 2\pi)$ so that

$$\begin{aligned} x &= r\sin\theta\cos\varphi, \\ y &= r\sin\theta\sin\varphi, \\ z &= r\cos\theta. \end{aligned} \tag{6.11}$$

See the illustration in Fig. 6.1. The laplacian in spherical coordinates involves

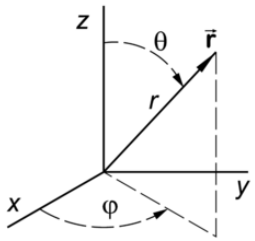

FIGURE 6.1. Spherical coordinates (Wikipedia).

the **radial laplacian**

$$\Delta_{\mathrm{rad}} = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} \tag{6.12}$$

and the **spherical laplacian**, or **Laplace-Beltrami operator**

$$\Delta_{\mathrm{sph}} = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}. \tag{6.13}$$

LEMMA 6.3. *We have*

$$\Delta = \Delta_{\mathrm{rad}} + \tfrac{1}{r^2}\Delta_{\mathrm{sph}}.$$

The proof is left as an exercise. The Laplace-Beltrami is an operator in $L^2(\mathbb{S}^2)$. We accept without proof its main properties. It is self-adjoint with

domain $\mathcal{D}(\Delta_{\rm sph}) = H^2(\mathbb{S}^2)$ and its spectrum is pure point. Its eigenvectors are the **spherical harmonics**

$$Y_\ell^m(\theta,\varphi) = c_{\ell,m} P_\ell^{|m|}(\cos\theta)\, \mathrm{e}^{\mathrm{i}m\varphi}, \tag{6.14}$$

where $\ell = 0, 1, 2, \dots$ and $m \in \{-\ell, \dots, \ell\}$. $c_{\ell,m}$ is a normalisation constant and P_ℓ^m is the **Legendre polynomial**

$$P_\ell^m(u) = \frac{(1-u^2)^{m/2}}{2^\ell \ell!} \frac{\mathrm{d}^{\ell+m}}{\mathrm{d}u^{\ell+m}} (u^2-1)^\ell. \tag{6.15}$$

The functions $\{Y_\ell^m\}$ form an orthonormal basis of $L^2(\mathbb{S}^2)$ and they satisfy

$$-\Delta_{\rm sph} Y_\ell^m = \ell(\ell+1) Y_\ell^m. \tag{6.16}$$

3. The radial equation and its solution

The eigenvalue equation in spherical coordinates is given by

$$\Big(-\Delta_{\rm rad} - \tfrac{1}{r^2}\Delta_{\rm sph} - \tfrac{1}{r}\Big) f(r,\theta,\varphi) = E f(r,\theta,\varphi). \tag{6.17}$$

We look for functions of the form $f(r,\theta,\varphi) = R(r)Y(\theta,\varphi)$. (This is the method of separation of variables.) Then $Y(\theta,\varphi)$ must be a spherical harmonic Y_ℓ^m with eigenvalue $\ell(\ell+1)$, and we get the differential equation for the radial coordinate:

$$\Big(-\Delta_{\rm rad} + \tfrac{\ell(\ell+1)}{r^2} - \tfrac{1}{r}\Big) R(r) = E R(r). \tag{6.18}$$

For r large the equation is approximately $-\frac{\mathrm{d}^2}{\mathrm{d}r^2} R(r) = E R(r)$ whose solutions behave like $\mathrm{e}^{\mathrm{i}\sqrt{E}r}$ if $E > 0$ and $\mathrm{e}^{\pm\sqrt{-E}r}$ if $E < 0$. For $E > 0$ one can construct a Weyl sequence (f_n) of normalised functions that are proportional to $\mathrm{e}^{\mathrm{i}\sqrt{E}\,r}$ on a large domain and that "escape to infinity" such that

$$\|(-\Delta - \tfrac{1}{r} - E) f_n\| \to 0 \tag{6.19}$$

as $n \to \infty$. Thus any $E > 0$ belongs to the spectrum of H. It is possible to show that $\sigma_{\rm c}(H) = [0, \infty)$.

Negative eigenvalues are possible a priori, and the corresponding eigenvectors should behave like $\mathrm{e}^{-\sqrt{-E}\,r}$ for large r. It is useful to introduce $s = 2\sqrt{-E}\,r$ and the function $F(s)$ by

$$R(r) = 2\sqrt{-E} s^\ell\, \mathrm{e}^{-\frac{1}{2}s}\, F(s). \tag{6.20}$$

Further, let $n = 1/2\sqrt{-E}$. The radial eigenvalue equation (6.18) becomes

$$\Big[s^2 \frac{\mathrm{d}^2}{\mathrm{d}s^2} + (2\ell + 2 - s)\frac{\mathrm{d}}{\mathrm{d}s} - (\ell + 1 - n)\Big] F(s) = 0. \tag{6.21}$$

It can be checked that only one solution is finite at the origin, the others have singularity $s^{-(2\ell+1)}$. Assume that F has the Taylor expansion

$$F(s) = \sum_{k \geq 0} a_k \tfrac{s^k}{k!}. \tag{6.22}$$

the coefficients can be found recursively, namely,

$$a_k = \frac{(\ell - n + k)(\ell - n + k - 1) \dots (\ell - n + 1)}{(2\ell + 1 + k) \dots (2\ell + 2)}. \tag{6.23}$$

If $n \in \{\ell+1, \ell+2, \dots\}$, we get that $a_k = 0$ for $k \geq n - \ell$ and $F(s)$ is a polynomial. Let us denote it $F_{n,\ell}$. Otherwise we have $a_k \sim k^{-n-\ell}$ for large k and $F(s) \sim \mathrm{e}^s$. Then $R(s) \sim \mathrm{e}^{\frac{1}{2}s}$ and this cannot be an L^2 function. It turns out that we have found all eigenvectors.

4. Summary: the spectrum of the hydrogen atom

The spectrum of the hamiltonian $H = -\Delta - \frac{1}{\|x\|}$ in $L^2(\mathbb{R}^3)$ consists of continuous spectrum $\sigma_{\rm c}(H) = [0, \infty)$, and point spectrum, $\sigma_{\rm p}(H) = \{-\frac{1}{4n^2} : n \in \mathbb{N}\}$. The eigenvectors corresponding to the eigenvalue $-\frac{1}{4n^2}$ are $f_{n,\ell,m}$ with $\ell = 0, 1, \dots, n-1$ and $m = -\ell, \dots, \ell$, where

$$f_{n,\ell,m}(r,\theta,\varphi) = R_{n,\ell}(r) Y_\ell^m(\theta,\varphi). \tag{6.24}$$

The functions Y_ℓ^m are the spherical harmonics; the functions $R_{n,\ell}$ are given by

$$R_{n,\ell}(r) = \tfrac{1}{n}\big(\tfrac{r}{n}\big)^\ell\, \mathrm{e}^{-r/2n}\, F_{n,\ell}\big(\tfrac{r}{n}\big), \tag{6.25}$$

where $F_{n,\ell}$ is a polynomial.

In the physics and chemistry literature, the eigenvalues are called "energy levels". The index n is used identically. But the index $\ell = 0, 1, \dots$ is replaced by the letters s, p, d, f, g, ... The levels of lithium ($Z = 3$) are similar and are illustrated in Fig. 6.2.

In a hydrogen atom at rest, the lone electron occupies the eigenstate with the lowest eigenvalue. If we neglect the interactions between electrons, the situation for bigger atoms is similar, all electrons occupy the states with lowest energy. But we must take into account:

- The **Pauli exclusion principle** states that two electrons cannot be in the same eigenstate, because they are fermions.
- The **spin** of the electron; this is an internal degree of freedom and the main consequence is that each eigenvector has multiplicity 2.

Then the electronic structure is obtained by populating the energy levels with lower energy, with 2 electrons in 1s; 2 electrons in 2s and 6 electrons in 2p; 2 electrons in 3s, 6 electrons in 3p, and 10 electrons in 3d; etc..., until the total number of Z electrons has been considered.

One could expect that the presence of electron-electron interactions makes this description naive, but it stays relevant in the "central potential approximation". The effect of interactions is to modify the energy of the levels, so 2s becomes smaller than 2p, etc... Since this approximation is at the centre of our understanding of the periodic table of elements, it is worth discussing it.

The full hamiltonian for the Z electrons around a nucleus with atomic number Z is

$$\begin{aligned} H &= \sum_{j=1}^{Z} \big(P_j^2 - \tfrac{Z}{\|x_j\|}\big) + \sum_{1\le i,j\le Z} \tfrac{1}{\|x_i - x_j\|} \\ &\approx \sum_{j=1}^{Z} \big(P_j^2 + V(\|x_j\|). \end{aligned} \tag{6.26}$$

Here, P_j, x_j are 3-component vectors that represent the jth electrons. As we see, the approximation consists in replacing the many-body interactions by an effective one-body potential. This is justified by the fact that electrons spread in a spherically uniform way, and by Newton's theorem stating that the resulting potential is the same as if all charges are located at the centre. Accordingly, the effective potential V should behave as

$$V(r) \approx \begin{cases} -\frac{Z}{r} & \text{if } r \text{ is small,} \\ -\frac{1}{r} & \text{if } r \text{ is large.} \end{cases} \tag{6.27}$$

The eigenvalue equation for the radial coordinates becomes

$$\big(-\Delta + \tfrac{\ell(\ell+1)}{r^2} + V(r)\big)R(r) = ER(r). \tag{6.28}$$

The effective potential is quite similar to the correct potential for the hydrogen atom and the theory essentially applies. We still get solutions indexed by n, ℓ, m with $n \in \mathbb{N}$; $\ell = 0, \dots, n-1$; $m = -\ell, \dots, \ell$. But the energy levels (eigenvalues) depend on both n and ℓ, not only on n. The order of succession of the levels is essentially independent of Z and is

1s, 2s, 2p, 3s, 3p, [4s, 3d], 4p, [5s, 4d], 5p, ...

where $[\cdot,\cdot]$ indicates that the levels have almost identical energy. The situation for lithium is shown in Fig. 6.2.

From these considerations follow the basic principles of chemistry! This accounts for the intriguing regularity of the periodic table of elements (it was discovered empirically by Mendeleev in 1869) and this explains chemical bonds and the formation of molecules, at least qualitatively.

EXERCISE 6.1. *Check that the last line in Eq. (6.10) is greater than or equal to $-1/3$ for all $\rho \in L^1(\mathbb{R}^3)$ such that $\int \rho = 1$ and $\int \frac{1}{\|x\|}\rho < \infty$.*

Hint: Use Hölder inequality to show that for all nonnegative ρ, we have $27(\frac{\pi}{2})^4 \int \rho^3 \ge (\int (\frac{1}{\|x\|} - \frac{1}{3})\rho)^3$. This inequality is sharp, the optimiser is $\rho(x) = \mathrm{const}(\frac{1}{\|x\|} - \frac{1}{3})_+^{1/2}$.

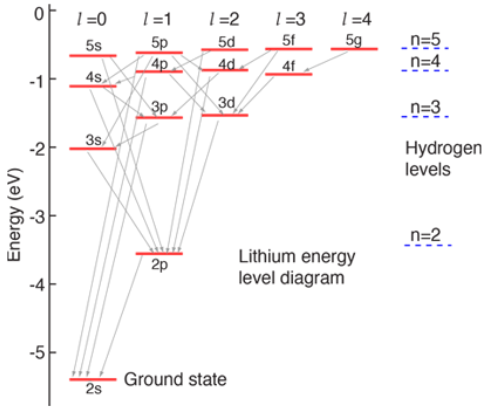

FIGURE 6.2. Energy levels of lithium; they are similar to those of hydrogen (the level 1s is outside the displayed range). (Source: internet.)

EXERCISE 6.2. *Prove Lemma 6.3. That is, show that for every twice-differentiable function $f : \mathbb{R}^3 \to \mathbb{C}$, we have*

$$\Delta f(x,y,z) = (\Delta_{\mathrm{rad}} + \tfrac{1}{r^2}\Delta_{\mathrm{sph}})f(r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta).$$

Here, (x,y,z) is given by (6.11).

EXERCISE 6.3. *Check that*

(i) $\langle Y_\ell^m, Y_{\ell'}^{m'}\rangle = \int_0^\pi \sin\theta \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\varphi\, \overline{Y_\ell^m(\theta,\varphi)}\, Y_{\ell'}^{m'}(\theta,\varphi) = \delta_{\ell,\ell'}\delta_{m,m'}$.

(ii) $-\Delta_{\mathrm{sph}} Y_\ell^m = \ell(\ell+1)Y_\ell^m$.

EXERCISE 6.4. *Write explicitly the eigenvector of the hydrogen hamiltonian with lowest eigenvalue. Check that it satisfies the eigenvalue equation.*

EXERCISE 6.5. *Prove that the positive real axis belongs to the spectrum of the hydrogen hamiltonian.*

CHAPTER 7

The Feynman-Kac formula

Motivated in part by Maupertuis' principle of least action, Feynman described the evolution of quantum particles using path integrals. While it allows to do correct calculations, the many mathematical difficulties are still unsolved. A similar description of the corresponding "parabolic" problem where the operator $\mathrm{e}^{-\mathrm{i}tH}$ is replaced by e^{-tH}, turns out to be mathematically possible, interesting, with useful applications to statistical mechanics, probability theory, and stochastic analysis. We describe it here and use it to prove some properties of the hydrogen atom.

1. Integral kernels and the Perron-Frobenius theorem

In $\mathbb{C}^n$ an operator A is represented by a matrix $(a_{i,j})_{1\leq i,j\leq n}$. Given a vector $f = (f(i))_{i=1}^n \in \mathbb{C}^n$, we have the relations

$$\begin{aligned} f(i) &= \langle e_i, f\rangle, \\ a_{i,j} &= \langle e_i, Ae_j\rangle, \\ (Af)(i) &= \sum_{j=1}^n a_{i,j} f(j), \end{aligned} \tag{7.1}$$

where $\{e_i\}_{i=1}^n$ is the usual orthonormal basis of $\mathbb{C}^n$. There are two natural generalisations of these relations to spaces of functions such as $L^2(\mathbb{R})$.

The first generalisation deals with separable infinitely-dimensional Hilbert space $\mathcal{H}$. Let $\{e_i\}_{i\geq 1}$ denote a countable orthonormal basis. If A is a bounded operator in $\mathcal{H}$, let

$$a_{i,j} = \langle e_i, Ae_j\rangle. \tag{7.2}$$

We then have a generalisation of the last relation in (7.1)

$$\langle e_i, Af\rangle = \sum_{j\geq 1} a_{i,j}\langle e_j, f\rangle. \tag{7.3}$$

The finite sum has become a series; it is absolutely convergent, as can be seen using Cauchy-Schwarz inequality.

The second generalisation of the relations (7.1) is to replace $f(i)$, $i = 1,\dots,n$, by a function $f(x)$, $x \in \mathbb{R}$. And by replacing the finite sum $\sum a_{i,j}f(j)$ by an

integral. That is, let $a(x,y)$ be a function $\mathbb{R}\times\mathbb{R}\to\mathbb{C}$ and define the operator A by

$$(Af)(x) = \int_{\mathbb{R}} a(x,y)f(y)\mathrm{d}y. \tag{7.4}$$

Such an operator is called an **integral operator** and the function $a(x,y)$ is its **integral kernel**. Under some conditions, see Exercises 7.1 and 7.2, one can show that A is a bounded operator.

We now consider a generalisation of the Perron-Frobenius theorem for matrices.

THEOREM 7.1 (Perron-Frobenius). *Let A be a self-adjoint integral operator with positive integral kernel, $a(x,y) > 0$ for all $x, y \in \mathbb{R}$. Assume that $\|A\| < \infty$ and that $\|A\|$ is an eigenvalue. Then the eigenvalue $\|A\|$ has multiplicity 1 and its eigenfunction can be taken to be positive, $f(x) > 0$ for all $x \in \mathbb{R}$.*

PROOF. Let $f \in L^2(\mathbb{R})$ a function such that $\int a(x,y)f(y)\mathrm{d}y = \|A\|f(x)$; then $\operatorname{Re} f$ (or $\operatorname{Im} f$) is also eigenfunction, since $a(x,y) \in \mathbb{R}$. This shows there exist real eigenfunctions, and there remains to see that they can be taken to be positive. It is clear that

$$(A|f|)(x) \geq |(Af)(x)|; \tag{7.5}$$

then

$$\langle |f|, A|f|\rangle \geq \langle |f|, |Af|\rangle = \|A\|\langle |f|, |f|\rangle = \|A\|\,\|f\|^2. \tag{7.6}$$

On the other hand, we have by Cauchy Schwarz

$$\langle |f|, A|f|\rangle \leq \|f\|\,\|A\,|f|\,\| \leq \|A\|\,\|f\|^2. \tag{7.7}$$

All above inequalities are therefore identities, so that $|f| \sim A|f|$ and $|f|$ is also an eigenfunction with eigenvalue $\|A\|$. Further, $|f|$ is positive almost everywhere since it is not identically zero and

$$|f(x)| = \frac{1}{\|A\|}\int_{\mathbb{R}} a(x,y)|f(y)|\mathrm{d}y. \tag{7.8}$$

The inequalities-identities above also imply that $A|f|(x) = |Af(x)|$ so that

$$\int_{\mathbb{R}} a(x,y)|f(y)|\mathrm{d}y = \Big|\int_{\mathbb{R}} a(x,y)f(y)\mathrm{d}y\Big|. \tag{7.9}$$

Then $f(y) = \mathrm{e}^{\mathrm{i}\theta}\,|f(y)|$ for some $\theta \in \mathbb{R}$ that does not depend on y. It follows that all eigenfunctions with eigenvalue $\|A\|$ can be taken to be positive almost everywhere; there cannot be two such functions that orthogonal, so the multiplicity of $\|A\|$ is 1. □

2. Wiener measure and Feynman-Kac formula

Given an arbitrary self-adjoint operator H on a Hilbert space $\mathcal{H}$, we have seen how to construct the operator $\mathrm{e}^{-\mathrm{i}tH}$. The latter operator is unitary, it satisfies the group property, and it is indeed related to H by the relation $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{-\mathrm{i}tH} = -\mathrm{i}H\,\mathrm{e}^{-\mathrm{i}tH}$. The same method allows to construct the operator e^{-tH} for $t \geq 0$ provided that H is bounded below.

THEOREM 7.2. *Assume that H is self-adjoint and that $E_0 > -\infty$, where*

$$E_0 \doteq \inf_{\substack{f\in\mathcal{D}(H)\\ \|f\|=1}} \langle f, Hf\rangle.$$

Then for every $t \geq 0$ there exists a bounded operator e^{-tH} that satisfies:

(i) *Strong continuity:* $\lim_{s\to t} \|\,\mathrm{e}^{-sH} f - \mathrm{e}^{-tH} f\| = 0$ *for all* $f \in \mathcal{H}$.
(ii) *Boundedness:* $\|\,\mathrm{e}^{-tH}\,\| \leq \mathrm{e}^{-tE_0}$.
(iii) *Semi-group property:* $\mathrm{e}^{-sH}\,\mathrm{e}^{-tH} = \mathrm{e}^{-(s+t)H}$ *for all* $s, t \geq 0$.
(iv) *Derivative involves H:* $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{-tH} = -H\,\mathrm{e}^{-tH} = -\,\mathrm{e}^{-tH}\,H$.

The proof of this theorem can be done using the Yosida's approximation method as for the evolution operator — using $(H - \lambda)$ with $\lambda < E_0$ instead of $(H - i\lambda)$; this is left as an exercise. Next we consider the operator $-\frac{\sigma^2}{2}\Delta$ in $L^2(\mathbb{R}^d)$, where $\sigma > 0$ is a parameter. This operator is nonnegative, so $\mathrm{e}^{t\frac{\sigma^2}{2}\Delta}$ exists and is bounded. Let g_t denote the gaussian function with mean 0 and variance σ^2:

$$g_t(x) = (2\pi\sigma^2 t)^{-d/2}\,\mathrm{e}^{-\frac{x^2}{2\sigma^2 t}}. \tag{7.10}$$

It is normalised so that $\int g_t = 1$.

PROPOSITION 7.3. *The operator $\mathrm{e}^{t\frac{\sigma^2}{2}\Delta}$ is an integral operator with integral kernel $g_t(x-y)$. That is, we have for all $f \in L^2(\mathbb{R}^d)$ that*

$$\big(\mathrm{e}^{t\frac{\sigma^2}{2}\Delta} f\big)(x) = \int_{\mathbb{R}^d} g_t(x-y) f(y)\mathrm{d}y.$$

PROOF. We already know that $\mathrm{e}^{t\frac{\sigma^2}{2}\Delta}$, and the integral operator with kernel $g_t(x-y)$, are bounded (hence continuous). It is enough to check the identity for a dense set of functions such as Schwartz functions. Taking the Fourier transform, the identity becomes

$$(2\pi)^{-d/2}\int_{\mathbb{R}^d} \mathrm{d}x\,\mathrm{e}^{-\mathrm{i}kx}\big(\mathrm{e}^{t\frac{\sigma^2}{2}\Delta} f\big)(x) = (2\pi)^{d/2}\widehat{g}_t(k)\widehat{f}(k) \tag{7.11}$$

with $\widehat{g}_t(k) = (2\pi)^{-d/2}\,\mathrm{e}^{-t\frac{\sigma^2}{2}k^2}$. Both sides of (7.11) are equal in the limit $t \to 0+$:

$$\lim_{t\to 0+}(2\pi)^{-d/2}\int_{\mathbb{R}^d}\mathrm{d}x\,\mathrm{e}^{-\mathrm{i}kx}\big(\mathrm{e}^{t\frac{\sigma^2}{2}\Delta} f\big)(x) = \widehat{f}(k) = (2\pi)^{d/2}\lim_{t\to 0+}\widehat{g}_t(k)\widehat{f}(k). \tag{7.12}$$

Next we check that both sides of (7.11) satisfy the same differential equation. The right side gives

$$\frac{\mathrm{d}}{\mathrm{d}t}(2\pi)^{d/2}\widehat{g}_t(k)\widehat{f}(k) = -\tfrac{\sigma^2}{2}k^2(2\pi)^{d/2}\widehat{g}_t(k)\widehat{f}(k). \tag{7.13}$$

The left side gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(2\pi)^{-d/2}\int_{\mathbb{R}^d}\mathrm{d}x\,\mathrm{e}^{-\mathrm{i}kx}\big(\mathrm{e}^{t\frac{\sigma^2}{2}\Delta} f\big)(x) &= (2\pi)^{-d/2}\int_{\mathbb{R}^d}\mathrm{d}x\,\mathrm{e}^{-\mathrm{i}kx}\big(\tfrac{\sigma^2}{2}\Delta\,\mathrm{e}^{t\frac{\sigma^2}{2}\Delta} f\big)(x)\\ &= -\tfrac{\sigma^2}{2}k^2(2\pi)^{-d/2}\int_{\mathbb{R}^d}\mathrm{d}x\,\mathrm{e}^{-\mathrm{i}kx}\big(\mathrm{e}^{t\frac{\sigma^2}{2}\Delta} f\big)(x).\end{aligned} \tag{7.14}$$

where the last identity follows from two integrations by parts. This proves that the identity (7.11) holds for all t and all k. One gets the claim with an inverse Fourier transform. □

We now describe the Wiener measure for paths in $\mathbb{R}^d$. Let C_x be the set of continuous paths $\omega : [0,\infty) \to \mathbb{R}^d$ such that $\omega(0) = x$. Let Σ be the smallest σ-algebra that contains all sets of the form

$$\big\{\omega \in C_x : \omega(t_1) \in I_1, \dots, \omega(t_n) \in I_n\big\}, \tag{7.15}$$

for arbitrary $n \in \mathbb{N}$, arbitrary $0 < t_1 < \cdots < t_n$, and arbitrary open sets $I_1, \dots, I_n \subset \mathbb{R}^d$. The Wiener measure can be characterised by the integral of a dense set of functions on paths. Given $n \in \mathbb{N}$ and $0 < t_1 < \cdots < t_n$, and given a continuous function $f : \mathbb{R}^{dn} \to \mathbb{R}$, let F be the function on C_x defined by

$$F(\omega) = f\big(\omega(t_1), \dots, \omega(t_n)\big). \tag{7.16}$$

THEOREM 7.4 (Wiener measure). *There exists a unique probability measure μ on (C_x, Σ) such that the Lebesgue integral of a function F as above is given by*

$$\int_{C_x} F(\omega)\mu(\mathrm{d}\omega) = \int_{\mathbb{R}^d}\mathrm{d}x_1 \cdots \int_{\mathbb{R}^d}\mathrm{d}x_n\, g_{t_1}(x - x_1)g_{t_2-t_1}(x_1 - x_2)\dots g_{t_n-t_{n-1}}(x_{n-1} - x_n) f(x_1, \dots, x_n).$$

The paths ω are also called **Brownian paths** by probabilists. This theorem is a consequence of the Riesz-Markov theorem, that states that positive linear functionals on spaces of continuous functions are represented by measures.

A useful property of the Wiener measure is that it concentrates on paths that are more than continuous: they are Hölder continuous.

DEFINITION 7.1. *Let $\alpha \in [0,1]$. A path $\omega : [0,t] \to \mathbb{R}^d$ is* **Hölder continuous** *with parameter α if there exists a constant K such that*

$$\|\omega(s) - \omega(u)\| < K|s-u|^\alpha$$

for all $s, u \in [0,t]$.

We let $H_x^\alpha([0,t])$ denote the subset of C_x whose paths are Hölder continuous with parameter α in the interval $[0,t]$. The next theorem is not that hard to prove.

THEOREM 7.5. *With respect to the Wiener measure, the set $H_x^\alpha([0,t])$ has probability 1 for all $x \in \mathbb{R}^d$, all $t > 0$, and all $\alpha < \frac{1}{2}$. The set $H_x^\alpha([0,t])$ has probability 0 for all $x \in \mathbb{R}^d$, all $t > 0$, and all $\alpha \geq \frac{1}{2}$.*

It follows from the last claim that Brownian paths are nowhere differentiable. A useful consequence of this theorem is that, if V is a function $\mathbb{R}^d \to \mathbb{R}$ with sufficient regularity, we can define the integral along a Wiener path by

$$\int_0^t V(\omega(s))\mathrm{d}s = \lim_{n\to\infty} \frac{t}{n} \sum_{j=1}^n V\big(\omega(j\tfrac{t}{n})\big). \tag{7.17}$$

THEOREM 7.6 (Feynman-Kac formula). *Let $H = -\frac{\sigma^2}{2}\Delta + V$ where V is continuous and bounded below. Then for any $f \in L^2(\mathbb{R})$, we have*

$$\big(\mathrm{e}^{-tH} f\big)(x) = \int_{C_x} \exp\Big\{-\int_0^t V\big(\omega(s)\big)\mathrm{d}s\Big\} f\big(\omega(t)\big)\mu(\mathrm{d}\omega).$$

We need the Trotter product formula; the proof of Theorem 7.6 can be found immediately afterwards.

THEOREM 7.7 (Trotter product formula). *Assume that $A, B, A+B$ are self-adjoint operators that are bounded below, and that the domain $\mathcal{D}(A)\cap \mathcal{D}(B)$ is dense in $\mathcal{H}$. Then for all $f \in \mathcal{H}$, we have*

$$\mathrm{e}^{-(A+B)} f = \lim_{n\to\infty} \Big(\mathrm{e}^{-\frac{1}{n}A}\, \mathrm{e}^{-\frac{1}{n}B}\Big)^n f.$$

PROOF. It is enough to prove the claim for $f \in \mathcal{D}(A)\cap\mathcal{D}(B)$. Writing $S_n = \mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B}$ and $T_n = \mathrm{e}^{-\frac{1}{n}(A+B)}$, and using telescopic sums, we have

$$\begin{aligned}\Big(\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B}\Big)^n - \mathrm{e}^{-(A+B)} &= S_n^n - S_n^{n-1}T_n + S_n^{n-1}T_n - S_n^{n-2}T_n^2 + \cdots - T_n \\ &= \sum_{k=1}^n \Big(\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B}\Big)^{n-k}\Big[\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B} - \mathrm{e}^{-\frac{1}{n}(A+B)}\Big]\mathrm{e}^{-\frac{k-1}{n}(A+B)}.\end{aligned} \tag{7.18}$$

Since $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{-tT} f\big|_{t=0} = -Tf$ for $f \in \mathcal{D}(T)$, there exists a sequence $(f_n^{(T)})$ such that $\|f_n^{(T)}\| \to 0$ as $n \to \infty$, and

$$\mathrm{e}^{-\frac{1}{n}T} f = f - \tfrac{1}{n}Tf + \tfrac{1}{n}f_n^{(T)}. \tag{7.19}$$

Then

$$\begin{aligned}\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B} f &= \mathrm{e}^{-\frac{1}{n}A}\big(f - \tfrac{1}{n}Bf + \tfrac{1}{n}f_n^{(B)}\big) \\ &= \mathrm{e}^{-\frac{1}{n}A} f - \tfrac{1}{n}\mathrm{e}^{-\frac{1}{n}A}Bf + \tfrac{1}{n}\mathrm{e}^{-\frac{1}{n}A} f_n^{(B)} \\ &= f - \tfrac{1}{n}Af + \tfrac{1}{n}f_n^{(A)} - \tfrac{1}{n}Bf + \tfrac{1}{n}h_n + \tfrac{1}{n}\mathrm{e}^{-\frac{1}{n}A} f_n^{(B)},\end{aligned} \tag{7.20}$$

where $h_n = (1 - \mathrm{e}^{-\frac{1}{n}A})Bf$ is such that $\|h_n\| \to 0$ as $n \to \infty$. Then

$$\big\|n\big(\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B} - \mathrm{e}^{-\frac{1}{n}(A+B)}\big)f\big\| \leq \big\|f_n^{(A)} + h_n + \mathrm{e}^{-\frac{1}{n}A} f_n^{(B)} - f_n^{(A+B)}\big\| \to 0 \tag{7.21}$$

as $n \to \infty$. It follows from the uniform boundedness theorem that the operator $n\big(\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B} - \mathrm{e}^{-\frac{1}{n}(A+B)}\big)$ converges to 0 uniformly on the set $\{\mathrm{e}^{-s(A+B)} f : s \in [0,1]\}$. Then

$$\big\|\big(\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B} - \mathrm{e}^{-\frac{1}{n}(A+B)}\big)\,\mathrm{e}^{-\frac{k-1}{n}(A+B)} f\big\| = o(\tfrac{1}{n}) \tag{7.22}$$

with $o(\frac{1}{n})$ independent of k. Finally

$$\big\|\big(\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B}\big)^n f - \mathrm{e}^{-(A+B)} f\big\| \leq \sum_{k=1}^n \big\|\big(\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B}\big)^{n-k}\big\|\, o(\tfrac{1}{n}). \tag{7.23}$$

We have $\big\|\big(\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B}\big)^{n-k}\big\| \leq \|\,\mathrm{e}^{-\frac{1}{n}A}\,\|^{n-k}\|\,\mathrm{e}^{-\frac{1}{n}B}\,\|^{n-k}$ which is less than a constant by Theorem 7.2 (ii). □

We now prove the Feynman-Kac formula.

PROOF OF THEOREM 7.6. Without loss of generality, we can assume that $\sigma^2 = 1$ and $V \geq 0$. Let

$$F_n(\omega) = \exp\Big\{-\frac{t}{n}\sum_{j=1}^n V\big(\omega(j\tfrac{t}{n})\big)\Big\}. \tag{7.24}$$

Since $V(\omega(s))$ is continuous as a function of s, the Riemann integral exists for all $\omega \in H^\alpha_x([0,t])$ (with $\alpha > 0$). Next,

$$\begin{aligned}
\int_{C_x} \mathrm{e}^{-\int_0^t V(\omega(s))\mathrm{d}s} f(\omega(t))\mu(\mathrm{d}\omega) &= \int_{C_x} \lim_{n\to\infty} F_n(\omega) f(\omega(t))\mu(\mathrm{d}\omega) \\
&\overset{(a)}{=} \lim_{n\to\infty} \int_{C_x} F_n(\omega) f(\omega(t))\mu(\mathrm{d}\omega) \\
&\overset{(b)}{=} \lim_{n\to\infty} \int_{\mathbb{R}^d} \mathrm{d}x_1 \cdots \int_{\mathbb{R}^d} \mathrm{d}x_n \, g_{\frac{t}{n}}(x - x_1) \dots g_{\frac{t}{n}}(x_{n-1} - x_n) \, \mathrm{e}^{-\frac{t}{n}\sum_{j=1}^n V(x_j)} f(x_n) \\
&\overset{(c)}{=} \lim_{n\to\infty} \Big(\underbrace{\mathrm{e}^{\frac{t}{2n}\Delta} \, \mathrm{e}^{-\frac{t}{n}V} \dots \mathrm{e}^{\frac{t}{2n}\Delta} \, \mathrm{e}^{-\frac{t}{n}V}}_{n \text{ times}} f \Big)(x) \\
&\overset{(d)}{=} \big(\mathrm{e}^{-t(-\frac{1}{2}\Delta + V)} f \big)(x).
\end{aligned} \tag{7.25}$$

The exchange of integral and limit in (a) is justified by the dominated convergence theorem, using $F_n(\omega) \le 1$ for all n and ω; the identity (b) uses the definition of the Wiener measure; we used Proposition 7.3 in (c) to replace gaussian integral kernels by the exponentials of laplacians; finally, the identity (d) is the Trotter product formula, Theorem 7.7. □

Using the same ideas, we can write the operator $\mathrm{e}^{-t(-\frac{1}{2}\Delta+V)}$ as an integral operator whose kernel can be expressed using the Wiener measure for Brownian bridges. Let $C^{(t)}_{x,y}$ be the set of continuous paths $\omega : [0,t] \to \mathbb{R}^d$ such that $\omega(0) = x$ and $\omega(t) = y$. We equip $C^{(t)}_{x,y}$ with the σ-algebra generated by the sets

$$\big\{ \omega \in C^{(t)}_{x,y} : \omega(t_1) \in I_1, \dots, \omega(t_n) \in I_n \big\} \tag{7.26}$$

for arbitrary $n \in \mathbb{N}$, arbitrary $0 < t_1 < \cdots < t_n < t$, and arbitrary open sets $I_1, \dots, I_n \subset \mathbb{R}^d$. As before, we consider functions of the form

$$F(\omega) = f\big(\omega(t_1), \dots, \omega(t_n)\big), \tag{7.27}$$

for some $n \in \mathbb{N}$, some times $0 < t_1 < \cdots < t_n < t$, and some continuous function $f : \mathbb{R}^{dn} \to \mathbb{R}$. The construction of the measure on bridges is then similar to Theorem 7.4.

THEOREM 7.8 (Wiener measure for Brownian bridges). *There exists a unique measure μ on $C^{(t)}_{x,y}$ such that the Lebesgue integral of a function F as above is given by*

$$\int_{C^{(t)}_{x,y}} F(\omega)\mu(\mathrm{d}\omega) = \int_{\mathbb{R}^d} \mathrm{d}x_1 \cdots \int_{\mathbb{R}^d} \mathrm{d}x_n \, g_{t_1}(x - x_1) g_{t_2 - t_1}(x_1 - x_2) \dots g_{t - t_n}(x_n - y) f(x_1, \dots, x_n).$$

The proof of this theorem also follows from the Riesz-Markov theorem. Notice that μ here is not a probability measure, because $\mu(C^{(t)}_{x,y}) = g_t(x-y)$ which is not equal to 1 (check this!). We now use the measure on Brownian bridges to write the integral kernel of the exponential of a Schrödinger operator.

THEOREM 7.9 (Feynman-Kac formula for integral kernel). *Let $H = -\frac{\sigma^2}{2}\Delta + V$ where V is continuous and bounded below. The operator e^{-tH} is an integral operator with kernel*

$$k(x,y) = \int_{C^{(t)}_{x,y}} \exp\Big\{ -\int_0^t V\big(\omega(s)\big)\mathrm{d}s \Big\} \mu(\mathrm{d}\omega).$$

PROOF. We need to show that

$$\big(\mathrm{e}^{-tH} f \big)(x) = \int_{\mathbb{R}^d} k(x,y) f(y) \mathrm{d}y \tag{7.28}$$

for f in a dense set of functions. By dominated convergence and Theorem 7.8, we have

$$\begin{aligned}
\int_{\mathbb{R}^d} k(x,y) f(y) \mathrm{d}y = \lim_{n\to\infty} \int_{\mathbb{R}^d} \mathrm{d}y f(y) \int_{\mathbb{R}^d} \mathrm{d}x_1 \cdots \int_{\mathbb{R}^d} \mathrm{d}x_{n-1} \, g_{\frac{t}{n}}(x - x_1) \\
\dots g_{\frac{t}{n}}(x_{n-1} - y) \, \mathrm{e}^{-\frac{t}{n}\sum_{j=1}^n V(x_j)}.
\end{aligned} \tag{7.29}$$

This is the same expression as in the middle line of Eq. (7.25), which is equal to $(\mathrm{e}^{-tH} f)(x)$. □

An important consequence of the Feynman-Kac formula is that the integral kernel of e^{-tH}, where H is a Schrödinger operator, is positive.

3. Application: Ground state of the hydrogen atom

We have only discussed partial proofs of the spectrum of the hydrogen atom. We have proved that $\{-\frac{1}{4n^2}\}_{n\in\mathbb{N}}$ are eigenvalues, and that $[0,\infty)$ belongs to the spectrum (we have not proved that $[0,\infty)$ is in the continuous spectrum). We now use the Feynman-Kac theory to prove that $-\frac{1}{4}$ is in the bottom of the spectrum (in other words, $(-\infty, -\frac{1}{4})$ belongs to the resolvent set).

Theorems 7.6 and 7.9 do not apply to the operator $H = -\frac{1}{2}\Delta - \frac{1}{\|x\|}$ since $V(x) = -\frac{1}{\|x\|}$ is not bounded below. But we can use a limiting argument. Let

$$V_n(x) = \begin{cases} -\frac{1}{\|x\|} & \text{if } \|x\| > \frac{1}{n}, \\ -n & \text{otherwise,} \end{cases} \tag{7.30}$$

so that V_n is continuous and bounded below; let $H_n = -\frac{1}{2}\Delta + V_n$. By Theorem 7.9 the integral kernel of e^{-tH_n} is

$$k_n(x,y) = \int_{C^{(t)}_{x,y}} \exp\Big\{-\int_0^t V_n\big(\omega(s)\big)\mathrm{d}s\Big\}\mu(\mathrm{d}\omega). \tag{7.31}$$

The limit $n\to\infty$ of $\mathrm{e}^{-\int_0^t V_n(\omega(s))}$ exists (it can be $+\infty$). Then $k_n(x,y)$ converges pointwise by monotone convergence to a function $k(x,y)$. What remains to be checked is that $k(x,y)$ is the integral kernel of e^{-tH}; in other words, k does not depend on our choice of the regularisation of V.

PROPOSITION 7.10. *For every $f\in L^2(\mathbb{R}^d)$, we have*

$$\big(\mathrm{e}^{-tH}f\big)(x) = \int_{\mathbb{R}^d} k(x,y)f(y)\mathrm{d}y.$$

PROOF. Observe that $H_nf(x) = Hf(x) + (\frac{1}{\|x\|}-n)1_{\|x\|<\frac{1}{n}}f(x)$. Then $H_nf\to Hf$ for all $f\in\mathcal{D}(H)$. We now check that $\mathrm{e}^{-tH_n}f\to\mathrm{e}^{-tH}f$ as $n\to\infty$, for every $t>0$ and $f\in L^2(\mathbb{R}^d)$. For $f\in\mathcal{D}(H)$,

$$\begin{aligned}\big\|\big(\mathrm{e}^{-tH_n}-\mathrm{e}^{-tH}\big)f\big\| &= \Big\|\int_0^t \mathrm{d}s\frac{\mathrm{d}}{\mathrm{d}s}\,\mathrm{e}^{-sH_n-(t-s)H}f\Big\| \\ &\le \int_0^t \mathrm{d}s\big\|\,\mathrm{e}^{-sH_n-(t-s)H}(H_n-H)f\big\| \\ &\le t\,\mathrm{e}^{-tE_0}\,\|(H_n-H)f\|,\end{aligned} \tag{7.32}$$

which indeed vanishes in the limit $n\to\infty$. A continuity argument extends this to all $f\in L^2(\mathbb{R}^d)$. Then

$$\begin{aligned}\big(\mathrm{e}^{-tH}f\big)(x) &= \lim_{n\to\infty}\big(\mathrm{e}^{-tH_n}f\big)(x) \\ &= \lim_{n\to\infty}\int_{\mathbb{R}^d} k_n(x,y)f(y)\mathrm{d}y \\ &= \int_{\mathbb{R}^d} k(x,y)f(y)\mathrm{d}y.\end{aligned} \tag{7.33}$$

The latter identity holds for every positive or negative f by monotone convergence. Then it extends to all real f using $f=f_+-f_-$, and to all complex f using $f=\mathrm{Re}\,f+\mathrm{i}\,\mathrm{Im}\,f$. This also implies that $k(x,y)$ is finite almost everywhere since e^{-tH} is bounded. □

We now take $H=-\Delta-\frac{1}{\|x\|}$ (i.e. $\sigma^2=2$) so we have the same operator as in Chapter 6. We prove that $-1/4$ is the bottom of the spectrum of H.

PROPOSITION 7.11.
(a) $\mathrm{e}^{1/4}$ *is an eigenvalue of* e^{-H}.
(b) $\mathrm{e}^{1/4} = \|\,\mathrm{e}^{-H}\,\|$.
(c) $-\frac{1}{4} = \inf\sigma(H)$.

In order to prove this proposition we need to show that $\|\,\mathrm{e}^{-H}\,\|$ belongs to the spectrum of e^{-H}. This follows from the fact that $\mathrm{e}^{-H}\ge 0$, and the following lemma.

LEMMA 7.12. *Let A be a bounded self-adjoint operator and let*

$$m = \inf_{f:\|f\|=1}\langle f,Af\rangle, \qquad M = \sup_{f:\|f\|=1}\langle f,Af\rangle.$$

Then
(a) $\|A\| = \sup_{f:\|f\|=1}|\langle f,Af\rangle|$.
(b) $\|A\| = \max(|m|,|M|)$.
(c) $m, M\in\sigma(A)$.

PROOF. Let $a = \sup_{f:\|f\|=1}|\langle f,Af\rangle|$; We have $a\le\|A\|$ by the Cauchy-Schwarz inequality. Next, assuming that $A\neq 0$, we can assume that $Af\neq 0$ in the supremum, and we introduce $g=\|Af\|^{1/2}f$ and $h=\|Af\|^{-1/2}Af$. We have

$$\langle g+h,A(g+h)\rangle - \langle g-h,A(g-h)\rangle = 2\langle f,A^2f\rangle + 2\langle Af,Af\rangle = 4\|Af\|^2. \tag{7.34}$$

We also have

$$\begin{aligned}\big|\langle g+h,A(g+h)\rangle - \langle g-h,A(g-h)\rangle\big| &\le \big|\langle g+h,A(g+h)\rangle\big| + \big|g-h,A(g-h)\rangle\big| \\ &\le a\big(\|g+h\|^2+\|g-h\|^2\big) \\ &\overset{(a)}{=} 2a\big(\|g\|^2+\|h\|^2\big) \\ &= 4a\|Af\|.\end{aligned} \tag{7.35}$$

Here, $\overset{(a)}{=}$ is the parallelogram identity. Combining (7.34) and (7.35), we get $\|Af\|\le a$, thus completing the proof of (a).

(b) follows immediately from (a).

For (c), we can assume without loss of generality that $0\le m\le M$. Let $(f_n)\subset\mathcal{H}$ such that $\|f_n\|=1$ and $\langle f_n,Af_n\rangle\to M=\|A\|$. We have

$$\begin{aligned}0\le\|Af_n-Mf_n\|^2 &= \langle Af_n-Mf_n,Af_n-Mf_n\rangle \\ &= \|Af_n\|^2 - 2M\langle f_n,Af_n\rangle + M^2 \\ &\le \|A\|^2 - 2M\langle f_n,Af_n\rangle + M^2 \to 0\end{aligned} \tag{7.36}$$

as $n \to \infty$. Then $(A - M)$ cannot have a bounded inverse, since

$$\|(A-M)^{-1}\| = \sup_{g\neq 0} \frac{\|(A-M)^{-1}g\|}{\|g\|} \geq \frac{\|f_n\|}{\|(A-M)f_n\|}. \tag{7.37}$$

For the lower bound we chose $g = (A - M)f_n$ (if $g = 0$, then $M \in \sigma_{\mathrm{p}}(A)$ and we are done). The right side diverges as $n \to \infty$. Then M does not belong to the resolvent set, it must belong to the spectrum. □

Proof of Proposition 7.11. Since $f(x) = \mathrm{e}^{-\frac{1}{2}\|x\|}$ is eigenfunction of H with eigenvalue $-\frac{1}{4}$, we have that f is eigenfunction of e^{-H} with eigenvalue $\mathrm{e}^{1/4}$ by Exercise 7.5 (i). This proves (a).

For (b), we first show that $\sigma_{\mathrm{c}}(\mathrm{e}^{-H}) \subset [0, 1]$. If $\lambda \in \sigma_{\mathrm{c}}(\mathrm{e}^{-H})$, there exists a Weyl sequence (g_n) such that $\|(\mathrm{e}^{-H} - \lambda)g_n\| \to 0$. Then $\|\mathrm{e}^{-H} g_n\| \to \lambda$. But

$$\begin{aligned}
\langle \mathrm{e}^{-H} g_n, \mathrm{e}^{-H} g_n\rangle &= \int_{\mathbb{R}^d} \mathrm{d}x \Big| \int_{\mathbb{R}^d} \mathrm{d}y\, k(x,y) g_n(y)\Big|^2 \\
&= \underbrace{\int_{\|x\|<R} \mathrm{d}x \Big| \underbrace{\int_{\mathbb{R}^d} \mathrm{d}y\, k(x,y) g_n(y)}_{\to 0 \,\forall x \text{ (Weyl)}}\Big|^2}_{\to 0 \text{ by dominated convergence}} + \int_{\|x\|>R} \mathrm{d}x \Big| \int_{\mathbb{R}^d} \mathrm{d}y\, k(x,y) g_n(y)\Big|^2 \\
&\leq C_n(R) + \mathrm{e}^{c_R} \underbrace{\langle \mathrm{e}^{\Delta} |g_n|, \mathrm{e}^{\Delta} |g_n|\rangle}_{\leq 1},
\end{aligned} \tag{7.38}$$

where $C_n(R) \to 0$ as $n \to \infty$ and $c_R \to 0$ as $R \to \infty$. For the second term in the middle line, we used that $k(x, y)$ is approximately gaussian for large $\|x\|$. Then $\lambda^2 \leq 1$ and $\sigma_{\mathrm{c}}(\mathrm{e}^{-H}) \in [0, 1]$.

It follows from Lemma 7.12 that $\|\mathrm{e}^{-H}\|$ belongs to the spectrum of e^{-H}. Since $\|\mathrm{e}^{-H}\| \geq \mathrm{e}^{1/4} > 1$, we get that $\|\mathrm{e}^{-H}\|$ belongs to the point spectrum. The corresponding eigenfunction can be taken to be positive by Theorem 7.1. It cannot be orthogonal to $\mathrm{e}^{-\frac{1}{2}\|x\|}$, so it must in fact be equal to this function. Then $\|\mathrm{e}^{-H}\| = \mathrm{e}^{1/4}$, and we get (b).

Together with Exercise 7.5 (ii), the result above implies that $\sigma_{\mathrm{c}}(H) \subset [0, \infty)$, so the bottom of the spectrum is an eigenvalue. It cannot be smaller than $-1/4$ since it would violate (b). This concludes the proof of (c). □

Exercise 7.1. *Assume that the integral kernel $a(x, y)$ is an $L^2(\mathbb{R} \times \mathbb{R})$ function. Show that the corresponding integral operator is bounded; more precisely, show that*

$$\|A\|^2 \leq \int_{\mathbb{R}\times\mathbb{R}} |a(x,y)|^2 \mathrm{d}x\mathrm{d}y. \tag{7.39}$$

The upper bound is in fact the square of the "Hilbert-Schmidt norm". If it is finite, the operator is compact. This is too restrictive for our purpose.

Exercise 7.2. *Assume for simplicity that $a(x, y)$ is continuous and let*

$$\begin{aligned}
R &= \sup_{x\in\mathbb{R}} \int_{\mathbb{R}} |a(x,y)| \mathrm{d}y, \\
\mathscr{R} &= \sup_{y\in\mathbb{R}} \int_{\mathbb{R}} |a(x,y)| \mathrm{d}x.
\end{aligned} \tag{7.40}$$

Show that the norm of the corresponding integral operator A satisfies

$$\|A\| \leq \sqrt{\mathscr{R} R}. \tag{7.41}$$

Hint: Use Jensen's inequality.

Exercise 7.3. *Assume that A is an integral operator with continuous integral kernel $a(x, y)$. Show that*

(i) *We have $a(y, x) = \overline{a(x, y)}$ for all $x, y \in \mathbb{R}$ if and only if A is symmetric.*
(ii) *If $\mathrm{Tr}\,|A| < \infty$, show that $\mathrm{Tr}\, A = \int_{\mathbb{R}} a(x, x)\mathrm{d}x$.*

Exercise 7.4. *Re-prove Theorem 7.7 in the simpler situation where A and B are bounded operators.*

Exercise 7.5. *Let H be a self-adjoint operator that is bounded below. Prove that*

(i) $\{\mathrm{e}^{-\lambda} : \lambda \in \sigma_{\mathrm{p}}(H)\} \subset \sigma_{\mathrm{p}}(\mathrm{e}^{-H})$.
(ii) $\{\mathrm{e}^{-\lambda} : \lambda \in \sigma_{\mathrm{c}}(H)\} \subset \sigma_{\mathrm{p}}(\mathrm{e}^{-H}) \cup \sigma_{\mathrm{p}}(\mathrm{e}^{-H})$.

APPENDIX A

Solutions to some exercises

EXERCISE 3.2: It is useful to start by establishing the formula for the integral of a gaussian. For $a, b \in \mathbb{C}$ with $\mathrm{Re}\, b > 0$,

$$\begin{aligned}\int_{\mathbb{R}} \mathrm{e}^{ax-bx^2}\,\mathrm{d}x &= \mathrm{e}^{a^2/4b}\int_{-\infty}^{\infty}\mathrm{e}^{-b(x-\frac{a}{2b})^2}\,\mathrm{d}x = \mathrm{e}^{a^2/4b}\int_{-\infty-\frac{a}{2b}}^{\infty-\frac{a}{2b}}\mathrm{e}^{-bx^2}\,\mathrm{d}x \\ &= \mathrm{e}^{a^2/4b}\int_{-\infty}^{\infty}\mathrm{e}^{-bx^2}\,\mathrm{d}x = \sqrt{\frac{\pi}{b}}\,\mathrm{e}^{a^2/4b}.\end{aligned} \tag{A.1}$$

We used complex analysis to change the path of the integral back to the real line.

Let us now solve the exercise, assuming that ψ_0 is a Schwartz function (an L^1 function should be ok). We regularise the integral in order to use Fubini theorem to exchange integrals.

$$\begin{aligned}\frac{1}{\sqrt{\pi \mathrm{i} t}}\int_{\mathbb{R}}\mathrm{e}^{\mathrm{i}(x-y)^2/4t}\,\psi_0(y)\mathrm{d}y &= \lim_{\varepsilon\to 0+}\frac{1}{\sqrt{\pi \mathrm{i} t}}\int \mathrm{d}y\,\mathrm{e}^{-(\varepsilon-\mathrm{i}/4t)(x-y)^2}\,\frac{1}{\sqrt{2\pi}}\int \mathrm{d}k\,\mathrm{e}^{\mathrm{i}ky}\,\widehat{\psi_0}(k) \\ &= \frac{1}{\sqrt{2\pi^2 \mathrm{i} t}}\lim_{\varepsilon\to 0+}\int \mathrm{d}k\,\mathrm{e}^{\mathrm{i}kx}\,\widehat{\psi_0}(k)\int \mathrm{d}y\,\mathrm{e}^{-\mathrm{i}k(x-y)}\;\mathrm{e}^{-(\varepsilon-\mathrm{i}/4t)(x-y)^2}.\end{aligned} \tag{A.2}$$

The last integral gives $\sqrt{\frac{\pi t}{-\mathrm{i}}}\,\mathrm{e}^{-tk^2}$ in the limit $\varepsilon \to 0+$, using (A.1). We get Eq. (3.19).

EXERCISE 3.3: Using Exercise 3.2, we have

$$\psi(x,t) = \frac{1}{\sqrt{4\pi \mathrm{i} t}}\int_{\mathbb{R}}\mathrm{e}^{\mathrm{i}(x-y)^2/4t}\left(\frac{2}{\pi\sigma^2}\right)^{1/4}\mathrm{e}^{\mathrm{i}v_0(y-x_0)}\;\mathrm{e}^{-(x-y)^2/\sigma^2}\,\mathrm{d}y. \tag{A.3}$$

One can make the change of variables $z = y - x$, rearrange the integrant, and use the formula (A.1) for gaussian integrals to get the solution.

EXERCISE 3.5: The answer is $g(x) = \mathrm{e}^{-\mathrm{i}k_0(x+x_0)}\,f(x+x_0)$. Then (i) is obvious; (ii) follows from the change of variables $x - x_0 \mapsto y$; and (iii) follows from $\widehat{g}(k) = \mathrm{e}^{\mathrm{i}kx_0}\,\widehat{f}(k+k_0)$, and the change of variables $k - k_0 \mapsto k'$.

EXERCISE 4.4: We use the hints. We have

$$\begin{aligned}\|Vf\|^2 &= \int |V(x)|^2|f(x)|^2\mathrm{d}x \\ &\le m^2\int_{|V|\le m}|f(x)|^2\mathrm{d}x + \|f\|_\infty^2\int_{|V|>m}|V(x)|^2\mathrm{d}x \\ &\le m^2\|f\|^2 + C^2\Big(\int_{|V|>m}|V|^2\Big)\,\big(\|f\|^2+\|\Delta f\|^2\big).\end{aligned} \tag{A.4}$$

Since $\sqrt{a+b} \le \sqrt{a}+\sqrt{b}$, we find

$$\|Vf\| \le C\sqrt{\int_{|V|>m}|V|^2}\,\|\Delta f\| + \sqrt{m^2 + C^2\int_{|V|>m}|V|^2}\,\|f\|. \tag{A.5}$$

The first constant can be made arbitrarily small by taking m large enough. This shows that the condition of Theorem 4.10 is fulfilled.

It is easy to apply this to the Coulomb potential in $d = 3$. Indeed, $\int_{\|x\|\le 1}\frac{1}{\|x\|^2}\mathrm{d}x$ is finite, as can be seen by writing this integral using spherical coordinates.

EXERCISE 6.5: For every $E \ge 0$, one can find a Weyl sequence by choosing functions $f_n(x)$ that are zero outside the domain $n < x_i < 2n$, $i = 1, 2, 3$, and that are proportional to $\mathrm{e}^{\mathrm{i}\sqrt{E}\,x_1}$ for $n + 1 < x_i < 2n - 1$. The functions are smooth in the boundary region. Then $\| - \frac{1}{\|x\|}f_n\| \sim \frac{1}{n}$ and $\|(-\Delta - E)f_n\| \sim \frac{1}{\sqrt{n}}$.

EXERCISE 7.2: Starting with the definition of the norm of an integral operator, and introducing $r(x) = \int |a(x,y)|\mathrm{d}y$, we have

$$\begin{aligned}\|A\|^2 &= \sup_{f:\int|f|^2=1}\int_{\mathbb{R}}\mathrm{d}x\left|\int_{\mathbb{R}}\mathrm{d}y\,a(x,y)f(y)\right|^2 \\ &\le \sup_f\int \mathrm{d}x\,r(x)^2\Big(\int \mathrm{d}y\,\frac{|a(x,y)|}{r(x)}|f(y)|\Big)^2 \\ &\le \sup_f\int \mathrm{d}x\,r(x)^2\int \mathrm{d}y\,\frac{|a(x,y)|}{r(x)}|f(y)|^2.\end{aligned} \tag{A.6}$$

The last inequality is Jensen. We now bound $r(x) \le R$ and after interchanging the integrals we get

$$\|A\|^2 \le R\sup_f\int \mathrm{d}y\,|f(y)|^2\int \mathrm{d}x|a(x,y)|. \tag{A.7}$$

The integral over x is less than $\mathfrak{R}$ and the integral over y gives 1.

EXERCISE 7.5: If $Hf = \lambda f$, we have $(H-\alpha)^{-1}f = (\lambda-\alpha)^{-1}f$ since the latter is the unique function which $(H-\alpha)$ maps to f. Then f is eigenfunction of Yosida's approximation $H_\alpha = -\alpha(H-\alpha)^{-1}H$ with eigenvalue $\frac{-\alpha\lambda}{\lambda-\alpha} = \lambda(1-\frac{\lambda}{\alpha})^{-1}$. It follows that

$$\mathrm{e}^{-tH}f = \lim_{\alpha\to-\infty}\mathrm{e}^{-tH_\alpha}f = \lim_{\alpha\to-\infty}\mathrm{e}^{-t\lambda(1-\frac{\lambda}{\alpha})^{-1}}f = \mathrm{e}^{-t\lambda}f. \tag{A.8}$$

This proves (i).

For (ii), let (f_n) be a Weyl sequence for (H, λ). We check that it is also a Weyl sequence for $(\,\mathrm{e}^{-H}\,,\,\mathrm{e}^{-\lambda}\,)$. We have

$$\begin{aligned}\big\|(\,\mathrm{e}^{-H}\,-\,\mathrm{e}^{-\lambda}\,)f_n\big\| &= \Big\|\int_0^1 \mathrm{d}t \frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{e}^{-tH-(1-t)\lambda}\,f_n\Big\| \\ &\leq \int_0^1 \mathrm{d}t \big\|\,\mathrm{e}^{-tH-(1-t)\lambda}\,(H-\lambda)f_n\big\| \\ &\leq \sup_{t\in[0,1]} \big\|\,\mathrm{e}^{-tH-(1-t)\lambda}\,\big\|\,\|(H-\lambda)f_n\|.\end{aligned} \tag{A.9}$$

The supremum over t is bounded and the last norm vanishes in the limit $n \to \infty$. Then (f_n) is a Weyl sequence for $(\,\mathrm{e}^{-H}\,,\,\mathrm{e}^{-\alpha}\,)$ and $\mathrm{e}^{-\alpha} \in \sigma_{\mathrm{c}}(\,\mathrm{e}^{-H}\,)$ by Theorem 5.1.

Bibliography

[1] G.B. Folland, *Real Analysis*, Wiley Interscience (1999)

[2] V. Gelfreich, *MA4A7 Quantum Mechanics: Basic Principles and Probabilistic Methods*, lecture notes available to members of the University of Warwick (2018)

[3] S.J. Gustafson, I.M. Sigal, *Mathematical Concepts of Quantum Mechanics*, Springer (2003)

[4] E. Kreyszig, *Introductory Functional Analysis with Applications*, Wiley (1989)

[5] E.H. Lieb, M. Loss, *Analysis*, Amer. Math. Soc. (2001)

[6] A. Messiah, *Quantum Mechanics*, Dover (1999)

Index