AN INTRODUCTION TO QUANTUM MECHANICS FOR MATHEMATICIANS

VEDRAN SOHINGER, AUTUMN 2022

Contents

1. **Introduction**

1.1. **Origins of quantum mechanics.** Up until the early 20th century, physics was primarily described by classical theory. According to this point of view, the physical world consisted of *matter* and *radiation (waves)*. The evolution of the possible states of the physical system is described by a first-order differential equation. This did not show to be sufficient to describe all relevant physical phenomena.

The starting point of quantum mechanics was Max Planck's idea in 1900 that electromagnetic radiation is emitted and absorbed in discrete amounts, which he called *quanta*. In 1905 Albert Einstein extended this theory to claim that electromagnetic radiation consists of discrete localised quanta of energy, which he called *photons*. This was in contrast to previous beliefs that light was thought to consist of waves of electromagnetic fields. For his contribution (the *photoelectric effect*), Albert Einstein obtained the Nobel Prize for Physics in 1921. In 1924 Louis de Broglie presented his thesis of the *wave–particle duality*, according to which all matter displays the wave–particle duality of photons. The latter was confirmed by the *Davisson-Germer experiment* in 1927. This experiment concerned the scattering of electrons off a crystal. It can be abstracted as the *double-slit experiment* outlined in Section 1.2 below.

The formal framework of Quantum Mechanics was developed in the 1920s by Werner Heisenberg and Erwin Schrödinger. Many fundamental contributions were later made by Paul Dirac. The mathematical framework of the study of Quantum Mechanics was subsequently pioneered by John von Neumann. This theory motivated in great part the development of Functional Analysis. For a more detailed discussion on the origin of quantum mechanics, we refer the reader to [**?**, **?**]

1.2. **The double-slit experiment.** In this section, we recall the setup of *Young's double-slit experiment*[1]. The point of this experiment was to note that the behaviour of individual electrons is intrinsically random. Moreover, the randomness propagates according to the laws of wave mechanics. Let us explain this more precisely.

We consider a beam of electrons which is fired at a plate with two narrow slits. On the other side of the plate is a screen. See Figure 1.

[1]Thomas Young (1773-1829) was a British scientist, who in addition to the double-slit experiment made fundamental contributions to the decypherment of Egyptian hieroglyphs!

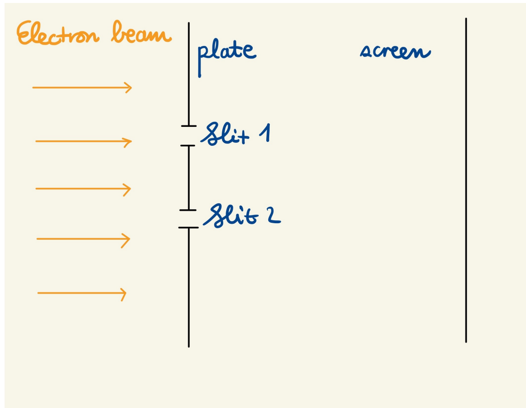

FIGURE 1. The setup of the double-slit experiment.

We first fire the beam with one of the slits blocked. In each case, we record the intensity distributions P_1, P_2 respectively on the screen as in Figures 2 and 3.

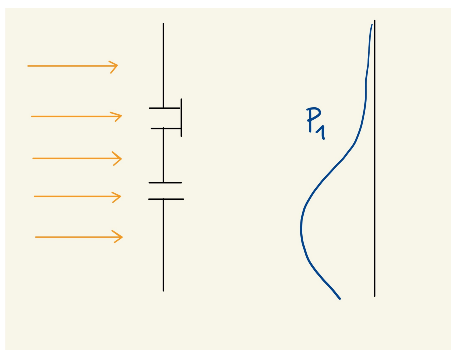

FIGURE 2. The first slit is blocked.

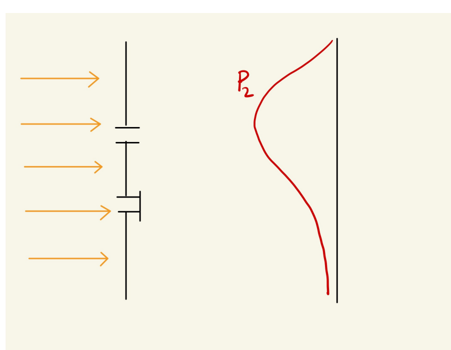

FIGURE 3. The second slit is blocked.

When both slits are open, one records an intensity distribution of the following type.

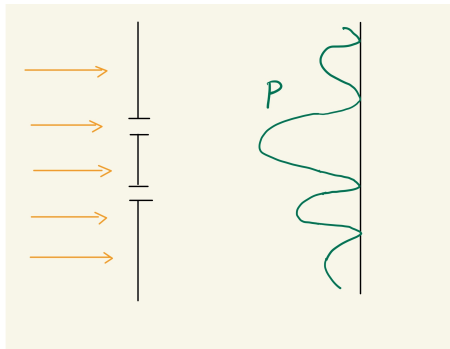

FIGURE 4. Both slits are open.

The surprising revelation was that $P \neq P_1 + P_2$!
We are presented with the following conclusions.

(1) It is impossible to predict exactly where a given electron will hit the screen. We can only determine the *probability distribution of the locations*.
(2) There is *interference*. The analogous situation is if two waves Φ_1 and Φ_2, represented by complex numbers corresponding to the amplitude and phase are combined, their intensity is $|\Phi_1 + \Phi_2|^2 \neq |\Phi_1|^2 + |\Phi_2|^2$. (Note that in this analogy $P_j = |\Phi_j|^2$).

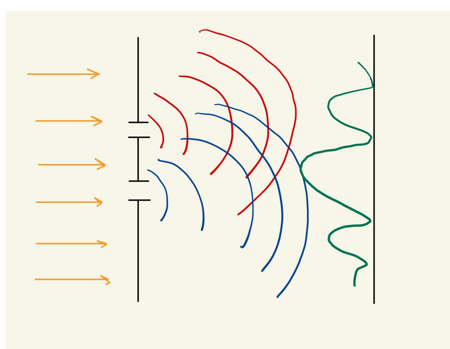

FIGURE 5. Wave interference.

Conclusion.

(1) Matter behaves in a random way.
(2) Matter exhibits wave-like properties.

Remark 1.1. *A more detailed discussion of the double-slit experiment and related experiments is given in* [**?**, Chapter 1].

In our module, we will set up a mathematical framework to study the objects and phenomena that were outlined above.

Remark 1.2. *The material in Sections 1.1 and 1.2 is not examinable.*

1.3. **Hamiltonian formulation of classical mechanics.** In this section, we will briefly summarise the setup of Classical Mechanics. As we will see later, Quantum Mechanics provides a structure which is a natural generalisation. Our discussion follows that of [**?**, Section 4.7] and [**?**, Section 1.2].

Our starting point is the *principle of minimal action*, by which solutions of physical equations minimise action functionals, defined in Definition 1.3 below.

Definition 1.3 (Action functional)**.** *Let $X \subset \mathbb{R}^d$ be an open set and let $L \in C^2(X \times \mathbb{R}^d; \mathbb{R})$ (i.e. it is a twice continuously differentiable function on $X \times \mathbb{R}^d$ taking values in $\mathbb{R}$). For $T > 0$, the **action functional** $S : C^1([0,T]; X) \to \mathbb{R}$ is defined to be the integral*

$$S(\varphi) := \int_0^T L(\varphi(t), \dot{\varphi}(t))\, \mathrm{d}t\,,$$

where $\dot{\varphi} \equiv \frac{\mathrm{d}}{\mathrm{d}t}\varphi$. Here, X is called the position space and L is called the Lagrangian.

In what follows, we take $X = \mathbb{R}^d$ and, unless otherwise specified, we consider

$$L(x,v) = \frac{1}{2}mv^2 - V(x)\,, \tag{1.1}$$

where $m > 0$ and $V \in C^2(\mathbb{R}^d)$. Here, m is referred to as the mass.

Let us us fix $T > 0$. With the Lagrangian L given as in (1.1), we compute the action as

$$S(\varphi) = \int_0^T \Big(\frac{m}{2}\,|\dot{\varphi}|^2 - V(\varphi)\Big)\, \mathrm{d}t\,. \tag{1.2}$$

We fix $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ and we consider (1.2) acting on

$$\mathcal{P}_{\mathbf{a},\mathbf{b}} := \big\{\varphi \in C^1([0,T]; \mathbb{R}^d)\,,\ \varphi(0) = \mathbf{a}\,,\ \varphi(T) = \mathbf{b}\big\}\,. \tag{1.3}$$

We want to minimize S over all $\varphi \in \mathcal{P}_{\mathbf{a},\mathbf{b}}$. This leads to solving the *Euler-Lagrange equation*

$$S'(\varphi) = 0\,. \tag{1.4}$$

In (1.4), $S'(\varphi)$ denotes the variational derivative of S at φ. Let us explain this precisely. Given $\varphi \in \mathcal{P}_{\mathbf{a},\mathbf{b}}$, $\xi \in \mathcal{P}_{\mathbf{0},\mathbf{0}}$ (with notation as in (1.3)), and $\lambda \in \mathbb{R}$, we define $\varphi_\lambda^\xi := \varphi + \lambda\xi \in \mathcal{P}_{\mathbf{a},\mathbf{b}}$. We then *define* $S'(\varphi) : [0,T] \to \mathbb{R}^d$ by duality. Namely, it is the unique map from $[0,T]$ to $\mathbb{R}^d$ satisfying

$$\int_0^T \langle S'(\varphi)(t), \xi(t)\rangle\, \mathrm{d}t = \frac{\mathrm{d}}{\mathrm{d}\lambda} S(\varphi_\lambda^\xi)\big|_{\lambda=0} \quad \forall \xi \in \mathcal{P}_{\mathbf{0},\mathbf{0}}\,. \tag{1.5}$$

In (1.5) and in the remainder of the section, we denote by $\langle\cdot,\cdot\rangle$ the inner product on $\mathbb{R}^d$. At a minimiser, we want

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} S(\varphi^\xi_\lambda)\big|_{\lambda=0} = 0 \quad \forall \xi \in \mathcal{P}_{\mathbf{0},\mathbf{0}}$$

and so we indeed obtain (1.4).

Let us introduce some notation that will simplify the discussion in the sequel. We denote by $\partial_\varphi L$ and $\partial_{\dot\varphi} L$ the derivative of L in the variables $\varphi, \dot\varphi$ respectively. Recalling (1.2), by the chain rule, we compute[2] for $\xi \in \mathcal{P}_{\mathbf{0},\mathbf{0}}$

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} S(\varphi^\xi_\lambda)\big|_{\lambda=0} = \int_0^T \Big(\partial_{\dot\varphi} L(\varphi,\dot\varphi)\,\dot\xi(t) + \partial_\varphi L(\varphi,\dot\varphi)\,\xi(t)\Big)\,\mathrm{d}t\,. \tag{1.6}$$

We integrate by parts in the first term on the right-hand side of (1.6) and recall that by construction $\xi(0) = \xi(T) = \mathbf{0}$. Therefore, we can rewrite the right-hand side of (1.6) as

$$\int_0^T \Big(-\partial_t\big[\partial_{\dot\varphi} L(\varphi,\dot\varphi)\big] + \partial_\varphi L(\varphi,\dot\varphi)\Big)\,\xi(t)\,\mathrm{d}t\,. \tag{1.7}$$

By (1.6)–(1.7), it follows that we can rewrite (1.4) as a pointwise equation in t.

$$-\partial_t\big[\partial_{\dot\varphi} L(\varphi,\dot\varphi)\big] + \partial_\varphi L(\varphi,\dot\varphi) = 0\,. \tag{1.8}$$

For the explicit choice of Lagrangian (1.1), (1.8) becomes the *Newton equation of Classical Mechanics*

$$m\ddot\varphi = -\nabla V(\varphi)\,. \tag{1.9}$$

Definition 1.4 (Energy of a path)**.** *Given* $\varphi \in C^1([0,T];\mathbb{R}^d)$*, we define its energy* $\mathcal{E}(\varphi) : [0,T] \to \mathbb{R}$ *by*

$$\mathcal{E}(\varphi)(t) := \frac{\partial L}{\partial \dot\varphi}\big(\varphi(t),\dot\varphi(t)\big)\,\dot\varphi(t) - L(\varphi(t),\dot\varphi(t)) \equiv \frac{\partial L}{\partial\dot\varphi}\,\dot\varphi - L\,. \tag{1.10}$$

Lemma 1.5 (Conservation of energy)**.** *If* φ *solves the Euler-Lagrange equation* (1.4) *(and hence* (1.8)*), then the energy functional from Definition 1.4 is conserved in time.*

Proof. We compute, using the Leibniz rule and chain rule in (1.10)

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\bigg(\frac{\partial L}{\partial\dot\varphi}\,\dot\varphi - L\bigg) &= \bigg(\partial_t\big[\partial_{\dot\varphi} L(\varphi,\dot\varphi)\big]\dot\varphi + \partial_{\dot\varphi} L(\varphi,\dot\varphi)\,\ddot\varphi - \partial_\varphi L(\varphi,\dot\varphi)\,\dot\varphi - \partial_{\dot\varphi} L(\varphi,\dot\varphi)\,\ddot\varphi\bigg) \\ &= \bigg(\partial_t\big[\partial_{\dot\varphi} L(\varphi,\dot\varphi)\big] - \partial_\varphi L(\varphi,\dot\varphi)\bigg)\,\dot\varphi = 0\,.\end{aligned} \tag{1.11}$$

For the last equality in (1.11), we used (1.8). ∎

[2]Throughout we interpret the relevant expressions, such as $\partial_{\dot\varphi} L(\varphi,\dot\varphi)\,\dot\xi(t) \equiv \langle \partial_{\dot\varphi} L(\varphi,\dot\varphi), \dot\xi(t)\rangle$ below etc. as appropriate inner products and we omit the explicit mention of the inner product to simplify notation. We write $\mathbf{a}, \mathbf{b}$, and $\mathbf{0}$ to emphasise that these are vectors in $\mathbb{R}^d$. We do not emphasise this elsewhere for simplicity of notation. Alternatively, one can consider the case $d = 1$ in which case all of the quantities are scalar-valued.

We now pass to a new set of variables $(x, v) \mapsto (x, p) \in \mathbb{R}^d \times \mathbb{R}^d$. Here p is a function of x and v given by

$$p = \partial_v L(x, v)\,. \tag{1.12}$$

Remark 1.6. *We note that, a priori p given by* (1.12) *belongs to the dual of $\mathbb{R}^d$, which by the Riesz representation theorem is canonically isomorphic to $\mathbb{R}^d$ via the mapping $v \mapsto \langle v, \cdot\rangle$. This is more relevant in the case when v takes values in a general vector space, as is analysed in* [**?**, Section 4.7]*, but we will not consider this in our module.*

Assumption: In what follows, we assume that, given x and p (1.12) has a unique solution for v. This is automatically satisfied for L as in (1.1), since we have that $p = mv$. In general, it can be shown that (1.12) has a unique solution for v provided that L is strictly convex in the second variable[3].

With the above assumption in mind, we can express the energy from Definition 1.4 in the new variables (x, p) as follows

$$H(x, p) := \big[\langle p, v\rangle - L(x, v)\big]\Big|_{v,\, \partial_v L(x,v) = p}\,. \tag{1.13}$$

Definition 1.7 (The Hamiltonian)**.** *The quantity in* (1.13) *is called the Hamilton function or the Hamiltonian.*

We note the following fundamental result, whose proof we omit. See [**?**, Theorem 4.8] for details.

Theorem 1.8. *Suppose that $L(x, v)$ and $H(x, p)$ are related by* (1.13)*. Then* (1.8) *is equivalent to* ***Hamilton's equations****, which are given by*

$$\dot{x} = \partial_p H(x, p)\,, \qquad \dot{p} = -\partial_x H(x, p)\,. \tag{1.14}$$

We apply (1.13) to the classical mechanics Lagrangian (1.1), we obtain the classical Hamiltonian

$$H(x, p) = \frac{1}{2m}|p|^2 + V(x)\,. \tag{1.15}$$

From (1.15) and Theorem 1.8, we in turn obtain Hamilton's equations of the form

$$\dot{x} = \frac{1}{m}\,p \qquad \dot{p} = -\partial_x V(x)\,, \tag{1.16}$$

which are equivalent to the Newton equations (1.9).

Let us now give a systematic interpretation of this setup that will be useful later in the module. In the above discussion, we interpret $x \equiv x(t)$ as being the ***position*** and $p \equiv p(t)$ as being the ***momentum*** of a classical particle at time t. We call $(x(t), p(t))$ the ***state*** of the classical particle at time t.

In what follows, we fix $H \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d)$ and $M \subset \mathbb{R}^d \times \mathbb{R}^d$ open such that the flow of (1.14) on M is global and maps M to itself, i.e. $(x(0), p(0)) \in M \Rightarrow (x(t), p(t)) \in M$

[3] In other words, for all $\theta \in (0, 1)$, and for all $x, v, v' \in \mathbb{R}^d$, we have $L(x, \theta v + (1-\theta)v') < \theta L(x, v) + (1-\theta)L(x, v')$. The latter holds provided that $\partial_v^2 L(x, v) > 0$ for all $x, v \in \mathbb{R}^d$.

for all t. We refer to M as the ***state space***. We denote by $G_t : M \to M$ the map induced by the flow of (1.14), i.e. $G_t(x_0, p_0) = (x(t), p(t))$ if $x(0) = x_0, p(0) = p_0$.

Instead of looking at the trajectory $(x(t), p(t))$ in the state space, we consider the evolution of smooth functions on M.

Definition 1.9 (Observables on M)**.** *We define $\mathcal{A} = C^\infty(M)$ to be the set of observables on M.*

We note that $\mathcal{A}$ is an algebra under the operation of addition and multiplication of functions. With G_t defined as above and with $\mathcal{A}$ as in Definition 1.9, the ***(classical) evolution operator*** is defined as

$$U_t : \mathcal{A} \to \mathcal{A}\,, \qquad U_t f := f \circ G_t\,. \tag{1.17}$$

Lemma 1.10 (Properties of the evolution operator)**.** *The evolution operator* (1.17) *satisfies the following group structure properties.*

(i) $U_0 = \mathbf{1}$*, where* $\mathbf{1} \equiv \mathbf{1}_{\mathcal{A}}$ *denotes the identity operator on* $\mathcal{A}$*.*
(ii) $U_s\, U_t = U_{s+t}$ *for all* $s, t \in \mathbb{R}$*.*
(iii) $U_t^{-1} = U_{-t}$ *for all* $t \in \mathbb{R}$*.*

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

Given $f \in \mathcal{A}$ and U_t as in (1.17), we let $f_t := U_t f$.

Lemma 1.11. *With f_t defined as above, we have*

$$\frac{\mathrm{d}}{\mathrm{d}t}\, f_t = \{H, f_t\}\,, \tag{1.18}$$

where $\{\cdot,\cdot\} : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ is given by

$$\{f, g\} := \sum_{i=1}^{d} \left(\frac{\partial f}{\partial p_i}\, \frac{\partial g}{\partial x_i} - \frac{\partial f}{\partial x_i}\, \frac{\partial g}{\partial p_i} \right). \tag{1.19}$$

We refer to (1.19) *as the* ***Poisson bracket on*** $\mathcal{A}$*.*

Proof. Let $(x, p) \in M$ be given. We denote by $(x(t), p(t))$ the solution of (1.14) with initial data (x, p). We compute, using the chain rule

$$\frac{\mathrm{d}}{\mathrm{d}t}\, f_t(x,p) = \frac{\mathrm{d}}{\mathrm{d}t}\, f(x(t), p(t)) = \sum_{i=1}^{d} \left[\frac{\partial f}{\partial x_i}(x(t), p(t))\, \dot{x}_i(t) + \frac{\partial f}{\partial p_i}(x(t), p(t))\, \dot{p}_i(t) \right].$$

Since $(x(t), p(t))$ solves (1.14), the above expression is

$$= \sum_{i=1}^{d} \left[\frac{\partial H}{\partial p_i}\, \frac{\partial f}{\partial x_i} - \frac{\partial H}{\partial x_i}\, \frac{\partial f}{\partial p_i} \right] (x(t), p(t)) = \{H, f_t\}(x, p)\,.$$

We deduce the claimed identity. □

Remark 1.12. *In some places in the literature, one considers* (1.19) *with the opposite sign and then the right-hand side of* (1.18) *gets replaced by* $\{f_t, H\}$. *This is a matter of conventions and will not play a significant role in the module.*

1.4. **Exercises for Section 1.**

Exercise 1.1. *Prove Lemma 1.10.*

Exercise 1.2. *Show that the Poisson bracket defined in* (1.19) *satisfies the following properties for all* $f, g, h \in \mathcal{A}$

(i) ***Antisymmetry:*** $\{f, g\} = -\{g, f\}$.
(i) ***Distributivity:*** $\{f + g, h\} = \{f, h\} + \{g, h\}$.
(iii) ***Leibniz rule***[4] ***:*** $\{fg, h\} = \{f, h\} g + f \{g, h\}$.
(iv) ***Jacobi identity:*** $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$.

Exercise 1.3. *Using* (1.14)*, show directly that energy is conserved under the trajectories of the Hamilton equations, i.e.* $\frac{\mathrm{d}}{\mathrm{d}t} H(x(t), p(t)) = 0$.

[4] The terminology here comes by thinking of $\{\cdot, h\}$ as a derivative

2. **Finite-dimensional quantum systems**

When analysing quantum systems, we are led to study a complex Hilbert space $\mathcal{H}$ and linear operators on this space. We will explain this in full detail in forthcoming chapters. The most common space to study is $\mathcal{H} = L^2(\mathbb{R}^3)$, which is *infinite-dimensional*. The operators that one typically studies are *not bounded* on H. For these reasons, some care will be required to rigorously set up the analysis. In this section, we first study the simpler setting of finite-dimensional Hilbert spaces. In other words, we consider $\mathcal{H} = \mathbb{C}^n$. Some of the facts from linear algebra will be recalled without proof (for proofs one can consult the text [**?**]).

2.1. **The Hilbert structure and operators on $\mathbb{C}^n$/linear algebra review.**

We recall some notions from linear algebra. In what follows, we sometimes denote elements of $\mathbb{C}^n$ as $\varphi = (\varphi_1, \dots, \varphi_n), \psi = (\psi_1, \dots, \psi_n)$, where $\varphi_i, \psi_i \in \mathbb{C}$ for $i = 1, \dots, n$.

- Given $\varphi, \psi \in \mathbb{C}^n$, we define their ***inner product*** by

  $$\langle \varphi, \psi \rangle := \sum_{i=1}^{n} \bar{\varphi}_i \, \psi_i \, .$$

  Note: In our module, we will be using the convention that complex inner products are *linear* in the second variable and *conjugate linear* in the first variable. In this convention, linear functionals look nice: $f(x) \equiv \langle f, x \rangle$.
- We say that $\varphi \in \mathbb{C}^n$ is ***perpendicular*** to $\psi \in \mathbb{C}^n$ and we write $\varphi \perp \psi$ if $\langle \varphi, \psi \rangle = 0$.
- The ***norm*** of a vector $\varphi \in \mathbb{C}^n$ is $\|\varphi\| = \sqrt{\langle \varphi, \varphi \rangle}$. We recall that a norm has the following properties.
  - (i) *Positive-definiteness:* $\|\varphi\| \geq 0$ for all $\varphi \in \mathbb{C}^n$ with equality if and only $\varphi = 0$.
  - (ii) *Triangle inequality:* $\|\varphi + \psi\| \leq \|\varphi\| + \|\psi\|$ for all $\varphi, \psi \in \mathbb{C}^n$.
  - (iii) *Absolute homogeneity:* $\|c\varphi\| = |c| \|\varphi\|$ for all $\varphi \in \mathbb{C}^n$ and $c \in \mathbb{C}$.
- We recall the ***Cauchy-Schwarz inequality*** which states that for all $\varphi, \psi \in \mathbb{C}^n$ we have $|\langle \varphi, \psi \rangle| \leq \|\varphi\| \, \|\psi\|$. Equality holds if and only if there exists $\lambda \in \mathbb{C}$ such that $\varphi = \lambda \psi$ or $\psi = \lambda \varphi$ (which we write as $\varphi \,||\, \psi$).
- A collection of vectors $\{e_1, \dots, e_n\}$ in $\mathbb{C}^n$ is an ***orthonormal basis*** of $\mathbb{C}^n$ if $\langle e_i, e_j \rangle = \delta_{i,j}$ for all $i, j = 1, \dots, n$. Here $\delta_{i,j}$ denotes the Kronecker delta function, which takes the value 1 if $i = j$ and 0 if $i \neq j$.

  We often work in the *canonical basis*, where e_i is the vector whose i-th slot is given by $\delta_{i,j}$.
- Any vector $\varphi \in \mathbb{C}^n$ can be written as

  $$\varphi = \sum_{i=1}^{n} \langle e_i, \varphi \rangle e_i \, . \tag{2.1}$$

Moreover, (2.1) is the unique way to represent φ in the form $\sum_{i=1}^n b_i e_i$ for coefficients $b_i \in \mathbb{C}$.
(Note that in our convention, we take $b_i = \langle e_i, \varphi\rangle$.)

- We call an ***operator on*** $\mathbb{C}^n$ a linear map $A : \mathbb{C}^n \to \mathbb{C}^n$.
(In our module, all operators will be assumed to be linear except if explicitly stated otherwise.)
- Given A, an operator on $\mathbb{C}^n$, we define its ***operator norm*** as
$$\|A\| := \sup_{\varphi \neq 0} \frac{\|A\varphi\|}{\|\varphi\|} . \tag{2.2}$$
See Exercise 2.1 below.
- The adjoint of A is the operator A^* such that for all $\varphi, \psi \in \mathbb{C}^n$ we have
$$\langle A^* \varphi, \psi\rangle = \langle \varphi, A\psi\rangle . \tag{2.3}$$
This operator always exists is uniquely determined by (2.3).
- Given an orthonormal basis $\{e_i\}$ of $\mathbb{C}^n$, we recall that the matrix of an operator T in the basis $\{e_i\}$ is the $n \times n$ matrix whose ij-entry is given by $(T)_{ij} = \langle e_i, Te_j\rangle$. Note that we have for all $\varphi, \psi \in \mathbb{C}^n$
$$\langle \psi, T\varphi\rangle = \sum_{i,j=1}^n \langle \psi, e_i\rangle \langle e_j, \varphi\rangle \langle e_i, Te_j\rangle . \tag{2.4}$$
Given A an operator on $\mathbb{C}^n$, and $i, j \in \{1, \ldots, n\}$, we have by (2.3)
$$(A^*)_{ij} = \overline{A_{ji}} .$$
- We have
$$(AB)^* = B^* A^* . \tag{2.5}$$
- An operator A on $\mathbb{C}^n$ is ***self-adjoint (or Hermitian)*** if $A = A^*$.
- We fix an orthonormal basis $\{e_i\}$ of $\mathbb{C}^n$ (e.g. the canonical basis). We define the ***trace*** of the operator A on $\mathbb{C}^n$ to be
$$\mathrm{Tr} A := \sum_{i=1}^n A_{ii} .$$
Note that $\mathrm{Tr} A = \sum_{i=1}^n \langle e_i, Ae_i\rangle$. From the latter equality, it can be shown that $\mathrm{Tr} A$ is independent of the choice of orthonormal basis $\{e_i\}$, see Exercise 2.2.
- With $\{e_i\}$ as above, we define the ***determinant*** of an operator A on $\mathbb{C}^n$ by
$$\det A = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_{i=1}^n A_{i\sigma(i)} . \tag{2.6}$$
Here S_n denotes the set of all permutations on n elements and $\mathrm{sign} : S_n \to \{1, -1\}$ denotes the sign of a permutation. It is shown in linear algebra that (2.6)

is independent of the choice of $\{e_i\}$. Moreover, one has for all A, B operators on $\mathbb{C}^n$

$$\det(AB) = \det(A)\det(B)\,.$$

- A is ***invertible*** if there exists an operator A^{-1} such that $AA^{-1} = A^{-1}A = \mathbf{1}$, where $\mathbf{1}$ is the identity operator on $\mathbb{C}^n$. The inverse is unique and A is invertible if and only if $\det(A) \neq 0$.
- An operator U is ***unitary*** if $UU^* = U^*U = \mathbf{1}$. One then has $\|U\varphi\| = \|\varphi\|$ for all $\varphi \in \mathbb{C}^n$. Furthermore, it can be shown that there exists A self-adjoint such that $U = \mathrm{e}^{\mathrm{i}A}$.
- If $\lambda \in \mathbb{C}$ is such that there exists $\varphi \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ with $A\varphi = \lambda\varphi$, we say that λ is an ***eigenvalue*** of A and that φ is an ***eigenvector*** of A (corresponding to the eigenvalue λ).
  We note the following properties of eigenvalues and eigenvectors.
  1. An operator A on $\mathbb{C}^n$ has at most n distinct eigenvalues.
  2. The eigenvalues of a self-adjoint operator are all real.
  3. The eigenvalues of a unitary operator all have modulus 1.
  4. If A is ***normal***, meaning that $AA^* = A^*A$, then there exist eigenvectors of A that form an orthonormal basis of $\mathbb{C}^n$.
- An operator A is ***positive-definite***, which we denote by $A \geq 0$ if $\langle \varphi, A\varphi \rangle \geq 0$ for all $\varphi \in \mathbb{C}^n$. Positive-definite operators are necessarily self-adjoint, see Exercise 2.4. All of their eigenvalues are nonnegative.
- An operator P is a ***projection*** if $P^2 = P$. It is an ***orthogonal projection*** if in addition we have $P^* = P$.
- Given $\varphi \in \mathbb{C}^n$ with $\|\varphi\| = 1$, we define the ***orthogonal projector*** P_φ *onto the one-dimensional space spanned by* φ by

$$P_\varphi \psi := \langle \varphi, \psi \rangle \varphi\,. \tag{2.7}$$

  Note that P_φ is an orthogonal projection.
- Suppose that A is a normal operator. We know that there exists an orthonormal basis of eigenvectors $\{\varphi_1, \dots, \varphi_n\}$, where φ_i is an eigenvector corresponding to eigenvalue λ_i. We can then write

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

- Let $\mathcal{B}$ denote the vector space of operators on $\mathbb{C}^n$. We define the sesquilinear map $\langle \cdot, \cdot \rangle : \mathcal{B} \times \mathcal{B} \to \mathbb{C}$ by

$$\langle A, B \rangle := \mathrm{Tr} A^* B\,. \tag{2.8}$$

  Then $\langle \cdot, \cdot \rangle$ is an ***inner product*** on $\mathcal{B}$, see Exercise 2.5.

2.2. **Quantum states.** In the context of classical systems, we recall that the algebra of observables is given by smooth functions on the phase space M, recall Definition 1.9. In quantum systems, we replace smooth functions by self-adjoint operators. To this end, we give the following (slightly informal) definition.

Definition 2.1 (Quantum observables)**.** *In (finite-dimensional) Quantum Mechanics, the set of observables $\mathcal{A}$ consists of self-adjoint operators on a Hilbert space. In particular, in the finite-dimensional setting, $\mathcal{A}$ denotes the class of all self-adjoint operators on $\mathbb{C}^n$.*

We recall that $\mathcal{B}$ denotes the class of all (linear) operators on $\mathbb{C}^n$. Note that $\mathcal{B}$ is closed under composition, but $\mathcal{A}$ is not. Namely, for $A, B \in \mathcal{A}$, we have $(AB)^* = B^*A^* = BA$, which in general is not equal to AB. Therefore, we cannot conclude that $AB \in \mathcal{A}$.

One possibility to deal with the aforementioned problem is to define a different product $\bullet$ that is not 'composition of operators'. Namely, on $\mathcal{B}$, we define the product $\bullet$ by

$$A \bullet B := \frac{1}{2}(AB + BA)\,. \tag{2.9}$$

Let us note that for $A, B \in \mathcal{A}$, we have that $A \bullet B \in \mathcal{A}$. Namely, for $A, B \in \mathcal{A}$, we note that by (2.5)

$$(A \bullet B)^* = \frac{1}{2}\big((AB)^* + (BA)^*\big) = \frac{1}{2}(B^*A^* + A^*B^*) = \frac{1}{2}(BA + AB) = A \bullet B\,.$$

We will not use the product $\bullet$ much in our module. For an application, see Exercise 2.7 below. The ***commutator*** of $A, B \in \mathcal{B}$ is defined to be

$$[A, B] := AB - BA \in \mathcal{B}\,. \tag{2.10}$$

We note that in general it is not true that the commutator of two self-adjoint operators is self-adjoint; see Exercise 2.6 below. Instead, we consider the ***'quantum Poisson bracket'*** of $A, B \in \mathcal{B}$ given by

$$\{A, B\} = \mathrm{i}[A, B]\,. \tag{2.11}$$

(This should be compared with (1.19) above.)

Lemma 2.2. *For $A, B \in \mathcal{A}$ and $\{A, B\}$ given as in* (2.11)*, we have $\{A, B\} \in \mathcal{A}$.*

Proof. From (2.11) and (2.10), followed by (2.5), we compute

$$(\{A, B\})^* = -\mathrm{i}([A, B])^* = -\mathrm{i}(B^*A^* - A^*B^*) = \mathrm{i}(AB - BA) = \mathrm{i}[A, B] = \{A, B\}\,.$$

$\square$

Definition 2.3 (Quantum state)**.** *We define a **quantum state** to be a map $\omega : \mathcal{B} \to \mathbb{C}$ satisfying the following properties.*

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

(iii) *Normalisation:* $\omega(\mathbf{1}) = 1$ *(recall that* $\mathbf{1}$ *is the identity operator on* $\mathbb{C}^n$*).*

Remark 2.4. *We make a few comments on Definition 2.3.*

1. *Condition (ii) is equivalent to saying that* $\omega(B) \geq 0$ *for all positive-definite operators* B*, as they can be written of the form* $B = A^*A$ *for* $A \in \mathcal{B}$ *given by the 'square root of* B*'.*
2. *One can equivalently define a state to be a functional* $\tilde{\omega} : \mathcal{A} \to \mathbb{R}$ *which satisfies the following properties.*
   (i') *Linearity:* $\tilde{\omega}(\alpha A + \beta B) = \alpha\tilde{\omega}(A) + \beta\tilde{\omega}(B)$ *for all* $A, B \in \mathcal{A}$ *and* $\alpha, \beta \in \mathbb{R}$*.*
   (ii') *Positivity:* $\tilde{\omega}(A) \geq 0$ *for all positive-definite* A*.*
   (iii') *Normalisation:* $\tilde{\omega}(\mathbf{1}) = 1$*.*
   See Exercise 2.8.

Definition 2.5. *We say that an operator* ρ *on* $\mathbb{C}^n$ *is a* ***density operator*** *if it satisfies the following conditions.*

(i) ρ *is positive-definite (and hence self-adjoint by Exercise 2.4).*
(ii) $\mathrm{Tr}\rho = 1$*.*

Let ρ be a density operator as in Definition 2.5. Then, we note that $\omega : \mathcal{B} \to \mathbb{C}$ given by

$$\omega(A) := \mathrm{Tr}(\rho A) \tag{2.12}$$

defines a quantum state; see Exercise 2.9.

The converse statement holds.

Proposition 2.6 (Riesz representation theorem for quantum states)**.** *Let* ω *be a quantum state. Then there exists a unique density operator* ρ *as in Definition 2.5 such that* $\omega(A) = \mathrm{Tr}(\rho A)$ *for all* $A \in \mathcal{B}$*.*

Proof. Let $\omega : \mathcal{B} \to \mathbb{C}$ be a quantum state. The inner product (2.8) defined on $\mathcal{B}$ makes $\mathcal{B}$ into a Hilbert space (recall that we are working in finite dimensions). Therefore, by the general *Riesz representation theorem*, we deduce that there exists a unique operator $\rho \in \mathcal{B}$ such that for all $A \in \mathcal{B}$, we have

$$\omega(A) = \langle \rho, A \rangle = \mathrm{Tr}(\rho^* A)\,. \tag{2.13}$$

We now check that ρ is a density operator, i.e. it satisfies (i)–(ii) from Definition 2.5.

In order to check condition (i), let $\varphi \in \mathbb{C}^n$ with $\|\varphi\| = 1$ be given. We fix an orthonormal basis $\{e_1, \ldots, e_n\}$ of $\mathbb{C}^n$ (e.g. the canonical basis). We then use (2.4) with $T = \rho^*$ and $\psi = \varphi$ to write

$$\langle \varphi, \rho^* \varphi \rangle = \sum_{i,j=1}^{n} \langle \varphi, e_i \rangle \langle e_j, \varphi \rangle (\rho^*)_{ij}\,, \tag{2.14}$$

where in (2.14) $(\rho^*)_{ij} = \langle e_i, \rho^* e_j \rangle$ and $\langle \cdot, \cdot \rangle$ denotes the inner product on $\mathbb{C}^n$. We recall (2.7) and write $(P_\varphi)_{ji} = \langle e_j, \varphi \rangle \langle \varphi, e_i \rangle$ to deduce that the right-hand side of (2.14) equals

$$\sum_{i,j=1}^{n} (P_\varphi)_{ji}\,(\rho^*)_{ij} = \mathrm{Tr}(\rho^* P_\varphi) = \omega(P_\varphi)\,. \tag{2.15}$$

Recalling that $P_\varphi = P_\varphi^2$, we rewrite the right hand-side of (2.15) as

$$\omega(P_\varphi) = \omega(P_\varphi^2) \geq 0\,,$$

where for the last inequality, we used $P_\varphi^* = P_\varphi$ and condition (ii) from Definition 2.3. By scaling (recall that for now we considered $\|\varphi\| = 1$), we deduce that ρ is positive-definite, so condition (i) holds.

In order to check condition (ii), we compute

$$\mathrm{Tr}\rho = \mathrm{Tr}(\rho\mathbf{1}) = \mathrm{Tr}(\rho^*\mathbf{1}) = \omega(\mathbf{1}) = 1\,.$$

Here, we used the observation that ρ is self-adjoint (since it is positive-definite), (2.13) and condition (iii) from Definition 2.3. ☐

Remark 2.7. *The arguments in* (2.14)–(2.15) *show that for all $\varphi \in \mathbb{C}^n$ with $\|\varphi\| = 1$ and all $A \in \mathcal{B}$, we have*

$$\langle \varphi, A\varphi \rangle = \mathrm{Tr}(P_\varphi A)\,. \tag{2.16}$$

We now study the concept of *mixed* and *pure* quantum states. Before giving the precise definition, let us note the following elementary property.

Lemma 2.8 (Convexity of the set of quantum states)**.** *The set of quantum states is* ***convex****, i.e. given quantum states ω_1, ω_2 and $\alpha \in [0,1]$, their convex combination $\alpha\omega_1 + (1-\alpha)\omega_2$ is a quantum state.*

The proof of Lemma 2.8 is left as Exercise 2.10.

Definition 2.9 (Mixed and pure quantum states)**.** *A quantum state ω is said to be* ***mixed*** *if it can be written as a nontrivial convex combination of two distinct quantum states. More precisely, this means that there exist quantum states $\omega_1 \neq \omega_2$ and $\alpha \in (0,1)$ such that $\omega = \alpha\omega_1 + (1-\alpha)\omega_2$.*
A quantum state ω is said to be ***pure*** *if it is not mixed.*

We can reformulate Lemma 2.8 and Definition 2.9 as saying that pure quantum states are the *extremal points* of the convex set of quantum states.

Proposition 2.10. *A quantum state ω is pure if and only if there exists $\varphi \in \mathbb{C}^n$ with $\|\varphi\| = 1$ such that $\omega(A) = \mathrm{Tr}(P_\varphi A)$ for all $A \in \mathcal{B}$.*

Proof. Given ω a quantum state, let ρ be its associated density as in Proposition 2.6. The claim is equivalent to showing that ω is pure if and only if $\rho = P_\varphi$ for some $\varphi \in \mathbb{C}^n$ with $\|\varphi\| = 1$, i.e. if and only if ρ is an orthogonal projector (as in (2.7)).

By construction, ρ is self-adjoint and can therefore be written as

$$\rho = \sum_{i=1}^{n} \lambda_i P_{\varphi_i}\,, \tag{2.17}$$

where $(\varphi_i)_{i=1}^n$ is an orthonormal basis of $\mathbb{C}^n$ and $(\lambda_i)_{i=1}^n$ are elements of $[0,1]$ with the property that $\mathrm{Tr}\rho = \sum_{i=1}^n \lambda_i = 1$. By using (2.17), we deduce that

$$\omega = \sum_{i=1}^n \lambda_i \omega_i \,, \tag{2.18}$$

where for $i = 1, \ldots, n$, ω_i is the quantum state given by $\omega_i(\cdot) = \mathrm{Tr}(P_{\varphi_i} \cdot)$.

Suppose that ρ is not an orthogonal projector. Then (2.17) is not a trivial convex combination (in the sense that not all of the λ_i are zero or one). Substituting this into (2.18), we deduce that ω is mixed. Let us explain this step in a bit more detail. Without loss of generality, we can assume that $\lambda_1 \in (0,1)$. Therefore, we can rewrite (2.18) as

$$\omega = \lambda_1 \omega_1 + (1 - \lambda_1) \tilde{\omega}_2 \,, \tag{2.19}$$

where and

$$\tilde{\omega}_2 := \frac{1}{1-\lambda_1} \sum_{i=2}^n \lambda_i \omega_i \,.$$

We use Lemma 2.8 to deduce that $\tilde{\omega}_2$ is a quantum state. By construction, it is distinct from the quantum state ω_1. Hence, (2.19) indeed implies that ω is mixed.

Conversely, assume that $\rho = P_\varphi$ is an orthogonal projector for some $\varphi \in \mathbb{C}^n$ with $\|\varphi\| = 1$. We now show that the corresponding quantum state $\omega(\cdot) = \mathrm{Tr}(P_\varphi \cdot)$ is pure. Let us assume that

$$\omega = \alpha \omega_1 + (1-\alpha) \omega_2 \tag{2.20}$$

for some quantum states ω_1, ω_2 and for some $\alpha \in [0,1]$. We want to show that necessarily $\omega_1 = \omega_2 = \omega$. By Proposition 2.6, there exist density operators ρ_i such that $\omega_i = \mathrm{Tr}(\rho_i \cdot)$ for $i = 1, 2$. In particular, by (2.20) and the uniqueness in Proposition 2.6, we deduce that

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

We now use respectively the fact that ρ is a density operator, the fact that $P_\varphi^2 = P_\varphi$, (2.21), and linearity combined with (2.16) to deduce that

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

We note that for $j = 1, 2$, by using the Cauchy-Schwarz inequality, (2.2), and $\|\varphi\| = 1$, we have that

$$|\langle \varphi, \rho_j \varphi \rangle| \leq \|\varphi\| \, \|\rho_j \varphi\| \leq \|\rho_j\| \, \|\varphi\|^2 = \|\rho_j\| \,. \tag{2.23}$$

We note that for $j = 1, 2$, we have

$$\|\rho_j\| \leq 1 \,, \quad \text{with equality if and only if } \rho_j \text{ is an orthogonal projector.} \tag{2.24}$$

See Exercise 2.11 below.

Combining (2.22)–(2.24), we get that

$$1 = \alpha \langle \varphi, \rho_1 \varphi \rangle + (1-\alpha) \langle \varphi, \rho_2 \varphi \rangle \leq 1 \,. \tag{2.25}$$

Hence, equalities must hold in all steps in (2.23)–(2.25). Since $\alpha\langle\varphi, \rho_1\varphi\rangle + (1 - \alpha)\langle\varphi, \rho_2\varphi\rangle = 1$, (2.23)-(2.24) imply that for $j = 1, 2$, we have $\rho_j = P_{\varphi_j}$ for some $\varphi_j \in \mathbb{C}^n$ with $\|\varphi_j\| = 1$. Since $1 = \alpha\langle\varphi, P_{\varphi_1}\varphi\rangle + (1-\alpha)\langle\varphi, P_{\varphi_2}\varphi\rangle$, it follows that for $j = 1, 2$, we have $\langle\varphi, P_{\varphi_j}\varphi\rangle = 1$, which is true if and only if $P_{\varphi_1} = P_{\varphi_2} = P_\varphi$. The claim now follows. $\square$

Note: **The proof of the converse claim in Proposition 2.10 is due to Peter Mühlbacher. The proof of Proposition 2.10 is not examinable.**

Let us now fix a quantum state ω. We interpret this as *'the quantum system being in the state ω'*, or that the *'expectation'* of an operator $A \in \mathcal{A}$ is $\omega(A)$. In other words, if the experiment is repeated many times, we would expect that the average of the outcomes should converge to $\omega(A)$.
(One should interpret the above discussion on a purely heuristic level).

Let us now explain the above setup more rigorously. Fix a quantum state ω. By Proposition 2.6, there exists a density matrix ρ, which is given by (2.17). We interpret this as projecting onto φ_i with probability λ_i. For an operator $A \in \mathcal{A}$, we define its ***expectation in the quantum state*** ω to be

$$\sum_{i=1}^{n} \lambda_i \langle\varphi_i, A\varphi_i\rangle = \sum_{i=1}^{n} \lambda_i \operatorname{Tr}(P_{\varphi_i} A) = \operatorname{Tr}(\rho A) = \omega(A)\,.$$

For the first equality, we used (2.16).

We quantify the fluctuations of $A \in \mathcal{A}$ with respect to its expectation in the quantum state ω by using the ***standard deviation*** σ_ω defined as

$$\sigma_\omega(A) := \sqrt{\omega(A^2) - (\omega(A))^2} = \sqrt{\omega\big((A - \omega(A)\mathbf{1})^2\big)}\,. \tag{2.26}$$

(Note that the expression in (2.26) is indeed well-defined).

Proposition 2.11 (Heisenberg Uncertainty Principle in finite dimensions)**.** *For any quantum state ω on $\mathcal{B}$ and any self-adjoint operators $A, B \in \mathcal{A}$, with $\sigma_\omega(\cdot)$ as in* (2.26)*, we have*

$$\sigma_\omega(A)\,\sigma_\omega(B) \geq \frac{1}{2}\left|\omega([A,B])\right|.$$

Proof. Given the quantum state ω, let ρ be the associated density given by Proposition 2.6. We have

$$\left|\omega([A,B])\right| \leq |\omega(AB)| + |\omega(BA)| = |\operatorname{Tr}(\rho AB)| + |\operatorname{Tr}(\rho BA)|\,. \tag{2.27}$$

Let us analyse the first term on the right-hand side of (2.27). We know that ρ is positive-definite. Hence, it has a square root $\rho^{1/2}$. We apply the result of Exercise 2.3 and Cauchy-Schwarz inequality for the inner product (2.8), as well as the assumption that A, B are self-adjoint to deduce that

$$\begin{aligned}\left|\mathrm{Tr}(\rho AB)\right|^2 = \left|\mathrm{Tr}(B\rho A)\right|^2 = \left|\mathrm{Tr}(B\rho^{1/2}\rho^{1/2}A)\right|^2 \\ \leq \mathrm{Tr}(B\rho^{1/2}\rho^{1/2}B)\,\mathrm{Tr}(A\rho^{1/2}\rho^{1/2}A) = \mathrm{Tr}(\rho B^2)\,\mathrm{Tr}(\rho A^2) = \omega(A^2)\,\omega(B^2)\,.\end{aligned} \tag{2.28}$$

Note that, in the second line of (2.28), we used $(B\rho^{1/2})^* = \rho^{1/2}B$ (since B is self-adjoint). Furthermore, we used Exercise 2.3 to obtain

$$\mathrm{Tr}(\rho^{1/2}BB\rho^{1/2}) = \mathrm{Tr}(B\rho^{1/2}\rho^{1/2}B) = \mathrm{Tr}(\rho B^2)\,.$$

The second factor is obtained analogously. Arguing as for (2.28), we get

$$\left|\mathrm{Tr}(\rho BA)\right|^2 \leq \omega(A^2)\,\omega(B^2)\,. \tag{2.29}$$

Combining (2.27)–(2.29), we deduce that

$$\omega(A^2)\,\omega(B^2) \geq \frac{1}{4}\left|\omega([A,B])\right|^2\,. \tag{2.30}$$

We now replace A, B with $A-\omega(A)\mathbf{1}, B-\omega(B)\mathbf{1}$ in (2.30). Note that the new operators are indeed self-adjoint. Furthermore, we have that

$$[A-\omega(A)\mathbf{1}, B-\omega(B)\mathbf{1}] = [A,B]\,. \tag{2.31}$$

Hence, the claim follows from (2.30) and (2.31) if we recall (2.26). □

For a slightly simpler proof in the special case of Proposition 2.11 when ω is a pure state, see Exercise 2.12 below.

2.3. **Quantum evolution.** In this section, we set up the equations that describe the dynamics of quantum systems. In our setup, the possible states (wave functions) are described as *unit vectors in a Hilbert space*[5]. In this section we take the space to be $\mathbb{C}^n$.

We now introduce the ***quantum evolution operator***.

Definition 2.12. *The evolution of a quantum system up to time $t \in \mathbb{R}$ is given by a map $U_t : \mathbb{C}^n \to \mathbb{C}^n$ which satisfies the following properties for all $s, t \in \mathbb{R}$.*

(i) *U_t is linear.*
(ii) *U_t is unitary.*
(iii) *$U_{s+t} = U_s U_t$ and $U_0 = \mathbf{1}$.*

Let us comment on Definition 2.12. One should start by comparing this to the classical evolution operator (1.17) and its properties given in Lemma 1.10. From point (ii), we know by definition of unitary operators that U_t is invertible and that $\|U_t(\varphi)\| =$

[5]In the sequel, we well refer to the unit vectors in Hilbert space as **wave functions**, and we will reserve the term 'state' for quantum states.

$\|\varphi\|$ for all $\varphi \in \mathbb{C}^n$. Let us elaborate a bit more on the latter identity. Since $U_t^* = (U_t)^{-1}$, we have that

$$\|U_t\varphi\|^2 = \langle U_t\varphi, U_t\varphi\rangle = \langle U_t^* U_t\varphi, \varphi\rangle = \langle \varphi, \varphi\rangle = \|\varphi\|^2 .$$

From point (iii), we obtain that $U_{-t} = (U_t)^{-1} = (U_t)^*$. We refer to (iii) as the *semigroup property*. We interpret it as saying that the laws of motion do not change in time. By unitarity of U_t for all t, there exists H_t self-adjoint such that $U_t = \mathrm{e}^{-\mathrm{i}tH_t}$. However, by (iii), the H_t it is plausible to expect that U_t is independent of t. We will make this assumption without proof for now. This is true under appropriate continuity assumptions in t, as we will see in a more systematic way later. [6]

In particular, we assume that there exists H self-adjoint such that

$$U_t = \mathrm{e}^{-\mathrm{i}tH} . \tag{2.32}$$

Given $t \in \mathbb{R}$ and $\varphi \in \mathbb{C}^n$ with $\|\varphi\| = 1$, we let

$$\varphi_t := U_t\varphi = \mathrm{e}^{-\mathrm{i}tH}\varphi . \tag{2.33}$$

This is interpreted as the evolution of the (wave function) φ to time t. From (2.32), we deduce that φ_t satisfies the **Schrödinger equation**

$$\mathrm{i}\,\frac{\mathrm{d}}{\mathrm{d}t}\,\varphi_t = H\varphi_t . \tag{2.34}$$

(See Exercise 2.13.)

Lemma 2.13. *Let $\varphi \in \mathbb{C}^n$ with $\|\varphi\| = 1$ and $t \in \mathbb{R}$ be given. Recalling the notation* (2.7)*, we have that*

$$U_t\, P_\varphi\, U_t^* = P_{U_t\varphi} . \tag{2.35}$$

Proof. We know from Definition 2.12 (ii) that $\|U_t\varphi\| = \|\varphi\| = 1$. Note that, since $U_t^*\, U_t = \mathbf{1}$, we have

$$(U_t\, P_\varphi\, U_t^*)\,(U_t\, P_\varphi\, U_t^*) = U_t\, P_\varphi\, U_t^*$$

Therefore $U_t\, P_\varphi\, U_t^*$ is a projection. Moreover $(U_t\, P_\varphi\, U_t^*)^* = U_t\, P_\varphi\, U_t^*$ so it is an orthogonal projection. We now show that it has rank 1 (and therefore it is an orthogonal projection onto a one-dimensional subspace of $\mathbb{C}^n$). Namely, we have

$$\operatorname{rank}(U_t\, P_\varphi\, U_t^*) \le \operatorname{rank} P_\varphi = 1 .$$

Since $U_t\, P_\varphi\, U_t^* \neq 0$, we deduce that its rank is equal to 1.

Finally, using $U_t^*\, U_t = \mathbf{1}$ again, we note that

$$U_t\, P_\varphi\, U_t^*\,(U_t\,\varphi) = U_t\, P_\varphi\,\varphi = U_t\varphi .$$

Therefore, $U_t\, P_\varphi\, U_t^*$ is the projection on the one-dimensional space spanned by $U_t\varphi$. The claim (2.35) now follows. ☐

[6] When $n = 1$, let us write $U_t = \mathrm{e}^{\mathrm{i}\theta(t)}$ for θ a real-valued function. One is then led to study θ which solve the Cauchy functional equation $\theta(t+s) = \theta(t) + \theta(s)$. If we assume that θ is continuous we can show $\theta(t) = ct$ for some constant c. Otherwise, there exist quite wildly behaved solutions of this equation, for details see this entry on Wolfram Mathworld .

We now want to define in a consistent way the notion of time evolution of a quantum state ω. Given ω, we consider its associated density operator ρ, given by Proposition 2.6. We recall that we can then write ρ as (2.17). By using linearity and (2.35) in (2.17), we have

$$\rho_t := U_t \, \rho \, U_t^* = \sum_{i=1}^n \lambda_i \, P_{U_t \varphi_i} \,. \tag{2.36}$$

In particular ρ_t is a density operator since by unitarity $\{U_t \varphi_1, \ldots, U_t \varphi_n\}$ is an orthonormal basis of $\mathbb{C}^n$.

The evolution of ω to time t is then defined as

$$\omega_t(A) := \mathrm{Tr}(\rho_t A) \,. \tag{2.37}$$

for ρ_t given as in (2.36). In light of our earlier discussion, we refer to (2.37) as the *expectation of $A \in \mathcal{B}$ in the state ω_t at time t.*

Using (2.36) and Exercise 2.3 (i), we can rewrite (2.37) as

$$\omega_t(A) = \mathrm{Tr}(U_t \, \rho \, U_t^* \, A) = \mathrm{Tr}(\rho \, U_t^* \, A \, U_t) = \omega(U_t^* \, A \, U_t) \,. \tag{2.38}$$

In light of (2.38), we can let the observable A evolve in time (instead of the state ω). We hence define the time-evolution A_t of A to time t by

$$A_t := U_t^* \, A \, U_t \,. \tag{2.39}$$

We can therefore write (2.38) as

$$\omega_t(A) = \omega(A_t) \,. \tag{2.40}$$

Combining (2.32) and (2.39), it follows that

$$A_t = \mathrm{e}^{\mathrm{i}tH} \, A \, \mathrm{e}^{-\mathrm{i}tH} \,. \tag{2.41}$$

Using Exercise 2.13 (iii), we have that

$$\frac{\mathrm{d}}{\mathrm{d}t} \, \mathrm{e}^{\pm \mathrm{i}tH} = \pm \mathrm{i} H \, \mathrm{e}^{\pm \mathrm{i}tH} \,. \tag{2.42}$$

Taking derivatives in t in (2.41) and using (2.42), we deduce that

$$\frac{\mathrm{d}}{\mathrm{d}t} \, A_t = \mathrm{i} \, [H, A_t] \,. \tag{2.43}$$

Recalling (2.11), we can rewrite (2.43) as

$$\frac{\mathrm{d}}{\mathrm{d}t} \, A_t = \{H, A_t\} \,. \tag{2.44}$$

This should be compared with (1.18) in Lemma 1.11 above[7]. One should note that in (2.44), one is studying the time-evolution of operators and in (1.18), one is studying the time evolution of smooth functions.

Throughout the sequel, we refer to H as the ***(Quantum) Hamiltonian***.

[7] We have set up our convention in (1.19) above so that the equations (1.18) and (2.44) formally look the same.

Definition 2.14 (Stationary quantum state)**.** *We say that a quantum state ω is **stationary** if it is invariant over time, i.e. if $\omega_t = \omega$ for all $t \in \mathbb{R}$.*

With notation as above, by using Proposition 2.6, it follows that $\omega_t = \omega$ is equivalent to $\rho_t = \rho$, which in turn is equivalent to $\frac{\mathrm{d}}{\mathrm{d}t}\rho_t = 0$. Recalling (2.35), (2.36), and arguing analogously as for (2.43), it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\rho_t = -\mathrm{i}[H, \rho_t]\,. \tag{2.45}$$

From (2.45), we deduce the following claim.

Proposition 2.15. *A quantum state ω is stationary if and only if its density operator commutes with the Hamiltonian H.*

Let us now analyse pure stationary states.

Proposition 2.16. *Let ω be a pure quantum state. Suppose that its density operator is given by $\rho = P_\varphi$ for $\varphi \in \mathbb{C}^n$ with $\|\varphi\| = 1$. Then ω is a stationary state if and only if φ is an eigenvector of the Hamiltonian H.*

Proof. Suppose that the quantum state ω is stationary. Then, by Proposition 2.15, the density operator $\rho = P_\varphi$ and the Hamiltonian H commute. Since $P_\varphi \varphi = \varphi$, we deduce that

$$H\varphi = HP_\varphi\varphi = P_\varphi H\varphi\,. \tag{2.46}$$

From (2.46), we deduce that $H\varphi = \lambda\varphi$ for some $\lambda \in \mathbb{C}$. Therefore, since $\varphi \neq 0$, we deduce that φ is an eigenvector of H.

Conversely, suppose that the density operator of ω is given by $\rho = P_\varphi$ for some eigenvector φ of H with $\|\varphi\| = 1$. Then, we know that

$$[H, P_\varphi] = 0\,. \tag{2.47}$$

In order to see (2.47), suppose that $H\varphi = \lambda\varphi$ and let $\psi \in \mathbb{C}^n$ be given. Note that by the self-adjointness of H, we have $\lambda \in \mathbb{R}$. We have that

$$HP_\varphi\psi = H\big(\langle\varphi, \psi\rangle\varphi\big) = \lambda\langle\varphi, \psi\rangle\,\varphi\,. \tag{2.48}$$

Also

$$P_\varphi H\psi = \langle\varphi, H\psi\rangle\varphi = \langle H\varphi, \psi\rangle\varphi = \langle\lambda\varphi, \psi\rangle\varphi = \lambda\langle\varphi, \psi\rangle\varphi\,. \tag{2.49}$$

In the proof of (2.49), we used the fact that H is self-adjoint and $\lambda \in \mathbb{R}$. We hence deduce (2.47) from (2.48) and (2.49). Therefore, by Proposition 2.15, it follows that ω is stationary. $\square$

We conclude this discussion with an observation about the evolution of eigenvectors of H.

Lemma 2.17. *Let $\varphi \in \mathbb{C}^n$ with $\|\varphi\| = 1$ be an eigenvector of the Hamiltonian H with eigenvalue λ. The following properties hold.*

(i) *Its evolution is given by $\varphi_t = \mathrm{e}^{-\mathrm{i}t\lambda}\varphi$.*

(ii) *The corresponding orthogonal projector is time-invariant, i.e.* $P_{\varphi_t} = P_\varphi$.

Proof. Claim (i) follows from (2.33) and Exercise 2.13 (iv) below. For claim (ii), we use (2.7), claim (i), and write for $\psi \in \mathbb{C}^n$

$$P_{\varphi_t}\psi = \langle \varphi_t, \psi\rangle \varphi_t = \langle \mathrm{e}^{-\mathrm{i}t\lambda}\varphi, \psi\rangle \mathrm{e}^{-\mathrm{i}t\lambda}\varphi = \langle \varphi, \psi\rangle\varphi = P_\varphi\psi\,.$$

Claim (ii) now follows. □

Remark 2.18 (Non-examinable remark on measurement in Quantum Mechanics)**.** *We give a brief remark on* ***measurement*** *in Quantum Mechanics. In the Copenhagen Interpretation, the physical system which is initially described by the vector* $\varphi \in \mathbb{C}^n$ *with* $\|\varphi\| = 1$*, evolves deterministically according to the Schrödinger equation* (2.34) *until an experiment is performed. In this framework, a measurement corresponds to a self-adjoint operator* A *on* $\mathbb{C}^n$*. Let us for simplicity assume that the spectrum of* A *consists of* n *distinct eigenvalues* $\lambda_1, \ldots, \lambda_n$*. We find a corresponding orthonormal basis of eigenvectors* $\{\psi_1, \ldots, \psi_n\}$.

The possible outcomes of the experiment are the eigenvalues λ_i*. These outcomes are random and the probability of measuring the eigenvalue* λ_i *is* $|\langle \varphi_t, \psi_i\rangle|^2$*. Note that since* $\|\varphi_t\| = 1$ *and since the* ψ_i *form an orthonormal basis, these numbers indeed add up to* 1*. Furthermore, after the experiment, the state of the system is given by the corresponding vector* ψ_i.

2.4. **Exercises for Section 2.**

Exercise 2.1 (Properties of the operator norm)**.** *This exercise is concerned with properties of the operator norm* (2.2)*.*

(i) *Show from the definition that* $\|\cdot\|$ *indeed defines a norm.*
(ii) *Show that for an operator* A *on* $\mathbb{C}^n$*, we have*

$$\|A\| = \sup_{\|\varphi\|=1} \|A\varphi\|\,.$$

(iii) *Given operators* A, B *on* $\mathbb{C}^n$*, show that* $\|AB\| \le \|A\|\|B\|$.
(iv) *Find operators* A, B *in* $\mathbb{C}^n$ *such that equality holds in (iii).*
(v) *Find operators* A, B *in* $\mathbb{C}^n$ *such that equality does not hold in (iii).*

Exercise 2.2 (Independence of the trace on the choice of orthonormal basis–finite dimensional case)**.** *Let* A *be an operator on* $\mathbb{C}^n$ *and let* $\{e_1, \ldots, e_n\}$ *and* $\{f_1, \ldots, f_n\}$ *be orthonormal bases of* $\mathbb{C}^n$*. Show that*

$$\sum_{i=1}^n \langle e_i, Ae_i\rangle = \sum_{j=1}^n \langle f_j, Af_j\rangle\,. \tag{2.50}$$

Hint for (2.50)*: For fixed* $j \in \{1, \ldots, n\}$*, note that* $Af_j = \sum_{i=1}^n \langle e_i, f_j\rangle\, Ae_i$*. Since all the sums are finite, we do not need to worry about interchanging orders of summation.*

Exercise 2.3 (Cyclicity of the trace–finite dimensional case)**.** *We collect several useful properties of the trace.*

(i) *Given operators A, B on $\mathbb{C}^n$ show that* $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

(ii) *Deduce that for operators A, B, C on $\mathbb{C}^n$, we have*

$$\mathrm{Tr}(ABC) = \mathrm{Tr}(CAB) = \mathrm{Tr}(BCA)\,. \tag{2.51}$$

The claim in (2.51) *is referred to as the 'cyclicity of the trace'.*

Exercise 2.4 (Positive-definite operators are self-adjoint)**.** *Suppose that A is a positive operator on $\mathbb{C}^n$. In this exercise, we outline the proof that A is self-adjoint.*

(i) *Let $\varphi \in \mathbb{C}^n$ be given. Show that $\langle A\varphi, \varphi\rangle = \langle \varphi, A\varphi\rangle$ and deduce that we have $\langle \varphi, (A - A^*)\varphi\rangle = 0$. We hence reduce the claim to showing that*

$$\langle \varphi, B\varphi\rangle = 0 \;\forall \varphi \in \mathbb{C}^n \Rightarrow \; B = 0\,. \tag{2.52}$$

(ii) *In order to prove* (2.52)*, observe that by assumption we have that $\forall \varphi, \psi \in \mathbb{C}^n$ and for all $z \in \mathbb{C}$ we have*

$$\langle(\varphi + z\psi), B(\varphi + z\psi)\rangle = 0\,. \tag{2.53}$$

Choosing appropriate values of z in (2.53)*, deduce that $\langle \varphi, B\psi\rangle = 0 \;\forall \varphi, \psi \in \mathbb{C}^n$ and deduce the claim.*

(iii) *Explain how the proof generalises to all operators A with the property that $\langle A\varphi, \varphi\rangle \in \mathbb{R}$ for all $\varphi \in \mathbb{C}^n$.*

Exercise 2.5 (Inner product on $\mathcal{B}$)**.** *Recall that $\mathcal{B}$ denotes the vector space of (linear) operators on $\mathbb{C}^n$. We recall that $\langle A, B\rangle := \mathrm{Tr}(A^*B)$.*

(i) *Show carefully that $\langle\cdot,\cdot\rangle$ defines an inner product on $\mathcal{B}$.*

(ii) *Is the norm that one obtains this way the same as the operator norm?*

(iii) *Write down the Cauchy-Schwarz inequality that one obtains from $\langle\cdot,\cdot\rangle$.*

(iv) *(Optional) Suppose that $A \in \mathcal{B}$ satisfies $A^*A = A^2$. Show that A is self-adjoint. HINT for (iv): It suffices to show that $\langle A - A^*, A - A^*\rangle = 0$ for the above inner product.*

Exercise 2.6. *Let $A, B \in \mathcal{A}$ be such that $[A, B] \in \mathcal{A}$. What can we say about A and B?*

Exercise 2.7 (Properties of the commutator)**.** *Recall that for $A, B \in \mathcal{B}$, the commutator $[A, B]$ is given by* (2.10) *and $A \bullet B$ is given by* (2.9)*. Given $A, B, C \in \mathcal{B}$ and $\alpha, \beta \in \mathbb{C}$ be given. The following properties hold.*

(i) *Antisymmetry: $[B, A] = -[A, B]$.*

(ii) *Bilinearity: $[\alpha A + \beta B, C] = \alpha[A, C] + \beta[B, C]$.*
(Note that we then immediately obtain linearity in the second component by (i)).

(iii) *Leibniz rule: $[A, B \bullet C] = [A, B] \bullet C + B \bullet [A, C]$.*

(iv) *Jacobi identity: $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$.*

Exercise 2.8 (Quantum states as real linear functionals)**.** *We now prove point 2. from Remark 2.4.*

(i) *Let $\tilde{\omega} : \mathcal{A} \to \mathbb{R}$ be a real linear functional. Show that there exists a unique extension of $\tilde{\omega}$ to a complex linear functional $\omega : \mathcal{B} \to \mathbb{C}$.*
In this exercise, it is helpful to note that every $A \in \mathcal{B}$ can be written as $A = B + \mathrm{i}C$ with $B, C \in \mathcal{A}$. In order to see this, we can take $B = (A + A^*)/2$ and $C = (A - A^*)/(2\mathrm{i})$. Is this representation unique?

(ii) *Use part (ii) to deduce point 2. from Remark 2.4.*

Exercise 2.9. *Show that ω given by* (2.12) *defines a quantum state. Here, you may use without proof the linear algebra fact that there exists $B \in \mathcal{B}$ such that $\rho = B^* B$ (this is needed to show point (ii) in Definition 2.3).*

Exercise 2.10. *Prove Lemma 2.8.*

Exercise 2.11. *Prove* (2.24)*.*
HINT: Use (2.17)*.*

Exercise 2.12 (Heisenberg Uncertainty Principle for pure states)**.** *Show Proposition 2.11 when ω is a pure state. More precisely, consider $\omega(\cdot) = \mathrm{Tr}(P_\varphi \cdot)$ for fixed $\varphi \in \mathbb{C}^n$ with $\|\varphi\| = 1$ and obtain the claim* (2.28) *in this special case by using just the Cauchy-Schwarz inequality on $\mathbb{C}^n$ (the proof now proceeds as before; this is the only step that simplifies slightly).*

Exercise 2.13 (Exponential of an operator)**.** *In this exercise, we rigorously obtain the construction of the exponential of an operator and we study its properties.*

(i) *Given $A \in \mathcal{B}$, show that the Taylor series*

$$\sum_{k=0}^{\infty} \frac{1}{k!} A^k \tag{2.54}$$

converges in the operator norm (2.2)*. We define the limit to be e^A.*
HINT: Since $\mathcal{B}$ with the operator norm is a Banach space, it suffices to show that the sequence

$$\left(\sum_{k=0}^{n} \frac{1}{k!} A^k \right)_n$$

is Cauchy with respect to the operator norm. In order to show this, it is helpful to recall Exercise 2.1 (iii).

(ii) *Show that $\mathrm{e}^A \, \mathrm{e}^B = \mathrm{e}^{A+B}$ for all $A, B \in \mathcal{B}$ such that $[A, B] = 0$.*
In particular, deduce that e^A is invertible with $(\mathrm{e}^A)^{-1} = \mathrm{e}^{-A}$.
HINT: For the main claim, it is helpful to note that for $A, B \in \mathcal{B}$ such that $[A, B] = 0$, we can write

$$(A + B)^n = \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k}$$

as in the standard binomial theorem.

(iii) *Show that* e^{tA} *satisfies the following differential equation.*

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{e}^{tA} = A\,\mathrm{e}^{tA}\,. \tag{2.55}$$

Show furthermore that the right-hand side of (2.55) *is equal to* $\mathrm{e}^{tA}A$.
HINT: For the last point, look at partial sums defining e^{tA} *(which involve powers of* A*) and note that they commute with* A*. Then take the limit.*

(iv) *Suppose that* A *is a normal operator with eigenvectors* $\{\lambda_i\}_{i=1}^n$ *and a corresponding orthonormal basis of eigenvectors* $\{\varphi_i\}_{i=1}^n$*. Then, we have*

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

3. **A quantum particle in the continuum**

When studying a quantum particle in the continuum, we are led to work in infinite-dimensional Hilbert spaces. In this section, we show how it is possible to extend the analysis from the finite-dimensional setting to this more challenging context. We first need to recall some notions from analysis.

In our framework, a quantum particle in the d-dimensional continuum will be given by $\varphi \in \mathcal{H} := L^2(\mathbb{R}^d)$ with $\|\varphi\|_{L^2} = 1$. We first review the precise definition of $L^2(\mathbb{R}^d)$. Let $\mathcal{L}^2(\mathbb{R}^d)$ denote the set of all complex-valued measurable functions $f : \mathbb{R}^d \to \mathbb{C}$ such that

$$\|f\|_{\mathcal{L}^2} := \left(\int |f|^2 \,\mathrm{d}x\right)^{1/2} < \infty\,. \tag{3.1}$$

Here $\mathrm{d}x$ denotes Lebesgue measure on $\mathbb{R}^d$. One can show (see e.g. [**?**, Lemma 0.31]) that $\|f\|_{\mathcal{L}^2} = 0$ if and only if $f = 0$ almost everywhere. Therefore (3.1) is not a norm on $\mathcal{L}^2$. Instead, one identifies functions that are equal almost everywhere (with respect to Lebesgue measure). Let $\mathcal{N}(\mathbb{R}^d) := \{f \in \mathcal{L}^2(\mathbb{R}^d)\,,\ f = 0 \text{ almost everywhere}\}$. We then let $L^2(\mathbb{R}^d)$ denote the corresponding quotient space.

$$L^2(\mathbb{R}^d) := \mathcal{L}^2(\mathbb{R}^d)/\mathcal{N}(\mathbb{R}^d)\,.$$

$L^2(\mathbb{R}^d)$ is an *infinite-dimensional vector space*. We define $\|\cdot\|_{L^2}$ by $\|f\|_{L^2} := \|f\|_{\mathcal{L}^2}$. This gives a well-defined norm on $L^2(\mathbb{R}^d)$ with respect to which $L^2(\mathbb{R}^d)$ is complete. Moreover, it is a Hilbert space with respect to the inner product

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

The set $C_c^\infty(\mathbb{R}^d)$ of smooth compactly supported functions is dense in $L^2(\mathbb{R}^d)$. A proof of all of the above facts can be found in [**?**, Section 0.6]. An analogous construction (without the inner product) gives us the space $L^p(\mathbb{R}^d)$ for $p \in [1,\infty)$. We also consider the space $L^\infty(\mathbb{R}^d)$. Here, the norm is given by

$$\|f\|_{L^\infty} := \mathrm{esssup}|f| \equiv \inf\bigl\{a\,, \mu(|f| > a) = 0\bigr\}\,. \tag{3.3}$$

In (3.3), μ denotes Lebesgue measure and $|f| > a$ denotes the set of all $x \in \mathbb{R}^d$ such that $|f(x)| > a$. Throughout, we say that a (linear) operator A mapping $L^2(\mathbb{R}^d)$ to the space of Lebesgue measurable functions on $\mathbb{R}^d$ is **bounded** if there exists $C > 0$ such that for all $f \in L^2(\mathbb{R}^d)$, we have $\|Af\|_{L^2} \leq C\|f\|_{L^2}$.

3.1. **The Fourier transform.**

We now give a self-contained summary of the important properties of the ***Fourier transform*** that we will use throughout the module. Some of this material might have been covered in other modules. We fix $d \in \mathbb{N}$ to be the spatial dimension.

Let $C^\infty(\mathbb{R}^d)$ denote the set of all smooth $f : \mathbb{R}^d \to \mathbb{C}$ which are smooth, i.e. which possess partial derivatives of any order. We call $\alpha \in \mathbb{N}_0^d$ a **multi-index** and $|\alpha| :=$

$\alpha_1 + \cdots + \alpha_d$ its **order**. We write $x^\alpha := x_1^{\alpha_1} \cdots x_d^{\alpha_d}$. If $\lambda \in \mathbb{C}$, we let

$$(\lambda x)^\alpha := \lambda^{|\alpha|} x_1^{\alpha_1} \cdots x_d^{\alpha_d} \,.$$

With these conventions, we write for $f \in C^\infty$

$$\partial_\alpha f := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} \,.$$

We define the class $\mathcal{S}(\mathbb{R}^d)$ of ***Schwartz functions*** by

$$\mathcal{S}(\mathbb{R}^d) := \Big\{ f \in C^\infty(\mathbb{R}^d) \,,\; \rho_{\alpha,\beta}(f) := \sup_x \left| x^\alpha \, (\partial^\beta f)(x) \right| < \infty \;\forall \alpha, \beta \in \mathbb{N}_0^d \Big\} \,. \tag{3.4}$$

Note that $\mathcal{S}(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$; see Exercise **??**.

Definition 3.1 (The Fourier transform on the Schwartz class)**.** *Given $f \in \mathcal{S}(\mathbb{R}^d)$, we define its Fourier transform as*[8]

$$\mathcal{F} f(p) \equiv \hat{f}(p) := \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(x) \, \mathrm{e}^{-\mathrm{i} p \cdot x} \, \mathrm{d}x \,. \tag{3.5}$$

Here $p \cdot x = \sum_{i=1}^d p_i x_i$ denotes the inner product on $\mathbb{R}^d$.
We refer to p as the ***momentum variable****.*

Lemma 3.2. *Let $f \in \mathcal{S}(\mathbb{R}^d)$ and $\alpha \in \mathbb{N}_0^d$ be given. Then the following identities hold*[9]

(i) $(\partial_\alpha f)\hat{\ }(p) = (\mathrm{i}p)^\alpha \, \hat{f}(p)$.
(ii) $(x^\alpha f(x))\hat{\ }(p) = \mathrm{i}^{|\alpha|} \, \partial_\alpha \hat{f}(p)$.

In particular, we deduce that the Fourier transform $\mathcal{F}$ maps $\mathcal{S}(\mathbb{R}^d)$ to itself.

Proof. We first prove (i). Let $1 \leq j \leq d$ be given. We can integrate by parts to compute

$$\begin{aligned}
\widehat{\left(\frac{\partial f}{\partial x_j}\right)}(p) &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \frac{\partial f}{\partial x_j}(x) \, \mathrm{e}^{-\mathrm{i} p \cdot x} \, \mathrm{d}x \\
&= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(x) \left(-\frac{\partial}{\partial x_j} \, \mathrm{e}^{-\mathrm{i} p \cdot x} \right) \mathrm{d}x \\
&= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{i} p_j \, f(x) \, \mathrm{e}^{-\mathrm{i} p \cdot x} \, \mathrm{d}x = \mathrm{i} \, p_j \, \hat{f}(p) \,.
\end{aligned}$$

By iterating this identity, we obtain (i).

[8]Some textbooks take a slightly different convention for the Fourier transform from the one used in (3.5). The other common convention is $\hat{f}(p) = \int f(x) \, \mathrm{e}^{-2\pi \mathrm{i} p \cdot x} \, \mathrm{d}x$. In our module, we will use (3.5).

[9]We interpret this as saying that, under the Fourier transform, multiplication and differentiation get exchanged.

We now prove (ii). Similarly as before, we have

$$\begin{aligned}(x_j f(x))^\frown(p) &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} x_j\, f(x)\, \mathrm{e}^{-\mathrm{i}p\cdot x}\, \mathrm{d}x \\ &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(x) \left(\mathrm{i}\frac{\partial}{\partial p_j} \mathrm{e}^{-\mathrm{i}p\cdot x}\right) \mathrm{d}x = \mathrm{i}\, \frac{\partial \widehat{f}}{\partial p_j}(p)\,.\end{aligned}$$

For the last equality, we can differentiate under the integral sign since $f \in \mathcal{S}(\mathbb{R}^d)$; see Remark 3.8 below. Iterating the above identity, we obtain claim (ii).

In order to deduce the claim that $\mathcal{F}$ maps $\mathcal{S}(\mathbb{R}^d)$ to itself, we first note that by (3.5), we have for all $g \in \mathcal{S}(\mathbb{R}^d)$.

$$\|\widehat{g}\|_{L^\infty} \leq \frac{1}{(2\pi)^{d/2}}\, \|g\|_{L^1} < \infty\,. \tag{3.6}$$

Here, we recall Exercise **??**.

If $f \in \mathcal{S}(\mathbb{R}^d)$, and $\alpha, \beta \in \mathbb{N}_0^d$, we have by (i) and (ii) that

$$p^\alpha\, (\partial_\beta \widehat{f}\,)(p) = \mathrm{i}^{-|\alpha|-|\beta|}\, [\partial_\alpha (x^\beta f)(x)]^\frown(p)\,,$$

which we see is bounded by taking $g = \partial_\alpha(x^\beta f) \in \mathcal{S}(\mathbb{R}^d)$ in (3.6). □

We give further properties of the Fourier transform in Exercise **??**.

An important feature of the Fourier transform is that it is a bijection $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$. In particular, we can compute its inverse.

Proposition 3.3 (The Fourier inversion formula)**.** *The Fourier transform $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ is a bijection, with inverse given by*

$$\mathcal{F}^{-1}(f)(x) \equiv \check{f}(x) := \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(p)\, \mathrm{e}^{\mathrm{i}p\cdot x}\, \mathrm{d}p\,. \tag{3.7}$$

In order to prove Proposition 3.3, we need a lemma.

Lemma 3.4 (The Fourier transform of a Gaussian)**.** *Let $z \in \mathbb{C}$ with $\operatorname{Re} z > 0$ be given. Then $\mathrm{e}^{-z|x|^2/2} \in \mathcal{S}(\mathbb{R}^d)$ and we have*

$$\mathcal{F}\Big(\mathrm{e}^{-\frac{z|x|^2}{2}}\Big)(p) = \frac{1}{z^{d/2}}\, \mathrm{e}^{-\frac{|p|^2}{2z}}\,,$$

where we interpret $z^{d/2}$ as $(\sqrt{z})^d$, and the branch cut of the square root is chosen along the negative real axis.

Proof of Lemma 3.4. The proof that $\mathrm{e}^{-z|x|^2/2} \in \mathcal{S}(\mathbb{R}^d)$ is left as Exercise **??** below. We note that the general claim follows from the special case $d = 1$. We now set $d = 1$ and define

$$g_z : \mathbb{R} \to \mathbb{C}\,, \qquad g_z(x) := \mathrm{e}^{-zx^2/2}\,. \tag{3.8}$$

The function g_z solves the initial value problem

$$\begin{cases} g_z'(x) + zxg_z(x) = 0 \\ g_z(0) = 1\,. \end{cases} \tag{3.9}$$

In (3.9), the prime denotes differentiation with respect to x. We take Fourier transforms in (3.9) and use Lemma 3.2 to deduce that

$$\mathrm{i}(p\widehat{g_z}(p) + z\widehat{g_z}'(p)) = 0\,. \tag{3.10}$$

In (3.10), the prime denotes differentiation with respect to p. We rewrite (3.10) as

$$\widehat{g_z}'(p) + \frac{p}{z}\,\widehat{g_z}(p) = 0\,, \tag{3.11}$$

which is equivalent to the equation in (3.9) with z replaced by $1/z$ (and x replaced by p). We now work backwards and solve (3.11) to obtain that

$$\widehat{g_z}(p) = c\,g_{1/z}(p)\,, \tag{3.12}$$

where $c \in \mathbb{C}$ is a suitably chosen constant. We determine c by setting $p = 0$ in (3.12) and recalling that (by (3.8)), we have $g_{1/z}(0) = 1$. Therefore,

$$c = \widehat{g_z}(0) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{-zx^2/2}\,\mathrm{d}x\,. \tag{3.13}$$

We want to show that $c = \frac{1}{\sqrt{z}}$. When $z > 0$, this is true by applying the change of variables $y = \sqrt{z}x$ in the identity $\int_{\mathbb{R}} \mathrm{e}^{-x^2/2}\,\mathrm{d}x = \sqrt{2\pi}$. For general $z \in \mathbb{C}$ with $\operatorname{Re} z > 0$, this follows by analytic continuation. The key point to observe that the function

$$z \mapsto \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{-zx^2/2}\,\mathrm{d}x$$

is analytic in $\{z \in \mathbb{C}\,,\quad \operatorname{Re} z > 0\}$. In order to justify the latter claim, we note that by Remark 3.8, we can take the $\frac{\partial}{\partial \bar{z}} = \frac{1}{2}(\frac{\partial}{\partial \operatorname{Re} z} + \mathrm{i}\,\frac{\partial}{\partial \operatorname{Im} z})$ derivative under the integral sign. □

Remark 3.5. *For the purposes of the proof of Proposition 3.3 given below, we only need to use Lemma 3.4 with $z > 0$. The result for general z with $\operatorname{Re} z > 0$ will be applied later; see Section 3.3.1.*

Proof of Proposition 3.3. Given $f \in \mathcal{S}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$, we want to show that

$$f(x) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \widehat{f}(p)\,\mathrm{e}^{\mathrm{i}p\cdot x}\,\mathrm{d}p\,. \tag{3.14}$$

By using Exercise **??** (ii) for the translated function $f(\cdot + x)$, it suffices to show that (3.14) holds for $x = 0$, i.e.

$$f(0) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \widehat{f}(p)\,\mathrm{d}p\,. \tag{3.15}$$

By using Fubini's theorem in (3.5), we get that for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$

$$\int_{\mathbb{R}^d} f(x)\,\widehat{\varphi}(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} \widehat{f}(p)\,\varphi(p)\,\mathrm{d}p\,. \tag{3.16}$$

We now show that (3.15) follows from (3.16) by taking a judicious choice of φ. With notation similar as in (3.8), we let $g(x) = \mathrm{e}^{-|x|^2/2}$. We take $\varphi \equiv \varphi^\lambda := \lambda^{-d} g(\cdot/\lambda)$ for $\lambda > 0$. By Exercise **??** (iii), we then have $\widehat{\varphi^\lambda}(p) = \widehat{g}(\lambda p)$. We substitute this into (3.16) to deduce that

$$\lambda^d \int_{\mathbb{R}^d} f(x)\,\widehat{g}(\lambda x)\,\mathrm{d}x = \int_{\mathbb{R}^d} \widehat{f}(p)\,g(p/\lambda)\,\mathrm{d}p\,. \tag{3.17}$$

By applying the change of variable $y = \lambda x$ on the right-hand side of (3.17), we deduce that

$$\int_{\mathbb{R}^d} f(y/\lambda)\,\widehat{g}(y)\,\mathrm{d}y = \int_{\mathbb{R}^d} \widehat{f}(p)\,g(p/\lambda)\,\mathrm{d}p\,. \tag{3.18}$$

We let $\lambda \to \infty$ and we use the dominated convergence theorem in (3.18) to deduce that

$$f(0) \int_{\mathbb{R}^d} \widehat{g}(y)\,\mathrm{d}y = g(0) \int_{\mathbb{R}^d} \widehat{f}(p)\,\mathrm{d}p\,. \tag{3.19}$$

We now deduce (3.15) from (3.19) because $g(0) = 1$ and $\int_{\mathbb{R}^d} \widehat{g}(y)\,\mathrm{d}y = (2\pi)^{d/2}$ by Lemma 3.4. $\square$

We can extend the Fourier transform from $\mathcal{S}(\mathbb{R}^d)$ to $L^2(\mathbb{R}^d)$. This is possible by the following result.

Proposition 3.6 (Plancherel's theorem)**.** *For all $f \in \mathcal{S}(\mathbb{R}^d)$, we have $\|\widehat{f}\|_{L^2} = \|f\|_{L^2}$. In particular, $\mathcal{F}$ extends to a unitary operator $\mathcal{F} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$.*

Proof. It suffices to prove the first claim. This follows by taking $\varphi = \overline{\widehat{f}} \in \mathcal{S}(\mathbb{R}^d)$ in (3.16); see Exercise **??**. $\square$

For a generalisation of Proposition 3.6, see Exercise **??** below.

Remark 3.7. *Note that* (3.5) *makes sense for $f \in L^1(\mathbb{R}^d)$. Thus, we can extend the Fourier transform from $\mathcal{S}(\mathbb{R}^d)$ to $L^1(\mathbb{R}^d)$ as well. So, we have found two possible extensions! When considering $\widehat{f}$ for f in $L^2(\mathbb{R}^d)$, we cannot in general use* (3.5)*. Instead one should interpret the Fourier transform as being*

$$\widehat{f}(p) = \lim_{N\to\infty} \frac{1}{(2\pi)^{d/2}} \int_{|x|\leq N} f(x)\,\mathrm{e}^{-\mathrm{i}p\cdot x}\,\mathrm{d}x\,, \tag{3.20}$$

which we interpret as an L^2-limit. Note that, in (3.20)*, we are writing $f \in L^2$ as the L^2-limit of the functions $f_N := f \cdot \chi_{|x|\leq N}$, all of which are in $L^1(\mathbb{R}^d)$, so we can compute their Fourier transform using* (3.5)*. When doing these two extensions, one should check that they coincide on $L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$. For a proof of this fact, see* [**?**, Lemma 7.6].

Remark 3.8 (Remark on differentiating under the integral sign)**.** *Throughout, we use the following general fact about differentiating under the integral sign, which follows from the dominated convergence theorem. For details, see e.g.* [**?**, Theorem 2.27] *and* [**?**, Problem A.20]. *Let $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{C}$ be a measurable function such that the following properties hold.*

(i) *$y \mapsto f(x, y)$ is integrable for all $x \in \mathbb{R}$.*
(ii) *$x \mapsto f(x, y)$ is differentiable for almost every $y \in \mathbb{R}^d$.*
(iii) *There exists $g \in L^1(\mathbb{R}^d)$ such that*

$$\left|\frac{\partial f}{\partial x}(x, y)\right| \leq g(y)\,.$$

Then the function $F : \mathbb{R} \to \mathbb{C}$ given by

$$F(x) = \int f(x, y)\,\mathrm{d}y$$

is differentiable and satisfies

$$F'(x) = \int \frac{\partial f}{\partial x}(x, y)\,\mathrm{d}y\,.$$

In other words, we can 'differentiate under the integral sign'.

3.2. **The position and momentum operators.** In what follows, we describe the state of a quantum particle by a function $\psi \in L^2(\mathbb{R}^d)$ with $\|\psi\|_{L^2} = 1$. In this interpretation, the *probability density* of finding this particle at a point $x \in \mathbb{R}^d$ is equal to $|\psi(x)|^2\,\mathrm{d}x$. We sometimes refer to ψ as being a ***wave function***. More precisely, given a Lebesgue-measurable set $\Lambda \subset \mathbb{R}^d$, the probability of finding the particle in the set Λ is equal to $\int_\Lambda |\psi(x)|^2\,\mathrm{d}x$ (note that this makes sense because, by assumption $\int_{\mathbb{R}^d} |\psi(x)|^2\,\mathrm{d}x = 1$). By using Plancherel's theorem (Proposition 3.6), we have that $\|\widehat{\psi}\|_{L^2} = 1$. In particular, as above, we can interpret $|\widehat{\psi}(p)|^2\,\mathrm{d}p$ to be the probability density for the particle to have momentum $p \in \mathbb{R}^d$.

We now study the position and momentum operators in more detail. Part of the discussion of the momentum operator will be heuristic and will be made completely precise in the following section. Throughout, we consider $d = 1$ for simplicity; the analysis can easily be extended to higher dimensions (noting that then position and momentum are d-dimensional vectors).

The ***position operator*** X is defined as multiplication by x, i.e.

$$(X\psi)(x) := x\psi(x)\,. \tag{3.21}$$

The average position is given by

$$\langle \psi, X\psi \rangle = \int_{\mathbb{R}} x\,|\psi(x)|^2\,\mathrm{d}x\,, \tag{3.22}$$

which corresponds to the expectation of x with respect to the probability density $|\psi(x)|^2\,\mathrm{d}x$. In (3.21), we note a difficulty; it is in general not true that $X\psi \in L^2$.

This is resolved by specifying a **domain** $\mathcal{D}(X) \subset L^2(\mathbb{R})$ for X, which is a dense linear subspace of $L^2(\mathbb{R})$. In this case, we can take

$$\mathcal{D}(X) := \left\{ \psi \in L^2(\mathbb{R}) \,, \int_{\mathbb{R}} x^2 \, |\psi(x)|^2 \, \mathrm{d}x < \infty \right\}. \tag{3.23}$$

This is a normed vector space if we take $\|\psi\| := \|x\psi\|_{L^2}$. We then have that $X : \mathcal{D}(X) \to L^2(\mathbb{R})$ is a bounded linear map (if we use the earlier norm on $\mathcal{D}(X)$). Furthermore, by the Cauchy-Schwarz inequality, it follows that for $\psi \in \mathcal{D}(X)$, the average position (3.22) is finite. Note that $C_c^\infty(\mathbb{R}) \subset \mathcal{D}(X) \subset L^2(\mathbb{R})$, hence $\mathcal{D}(X)$ is indeed dense in $L^2(\mathbb{R})$.

The ***momentum operator*** is defined as follows.

$$(P\psi)\hat{}\,(p) := p\,\widehat{\psi}(p)\,. \tag{3.24}$$

By (a suitable extension of) Lemma 3.2 (i), we can rewrite [10] (3.24) in physical (i.e. x) space as

$$(P\psi)(x) = -\mathrm{i}\nabla\psi(x)\,, \tag{3.25}$$

where in $1D$, we take $\nabla = \frac{\mathrm{d}}{\mathrm{d}x}$. We now take

$$\mathcal{D}(P) := \left\{ \psi \in L^2(\mathbb{R}) \,, \frac{\mathrm{d}}{\mathrm{d}x}\,\psi \in L^2(\mathbb{R}) \right\}. \tag{3.26}$$

The expectation of the momentum in the state ψ is equal to

$$\int_{\mathbb{R}} p\,|\widehat{\psi}(p)|^2 \,\mathrm{d}p = \left\langle \widehat{\psi}, \widehat{\left(-\mathrm{i}\,\frac{\mathrm{d}}{\mathrm{d}x}\,\psi\right)} \right\rangle = \left\langle \psi, -\mathrm{i}\,\frac{\mathrm{d}}{\mathrm{d}x}\,\psi \right\rangle. \tag{3.27}$$

In the second equality, we used Parseval's theorem (Exercise **??**).

In addition to considering the position and momentum operators given above, we also consider the **multiplication operator** V. This occurs when the quantum particle moves in an external potential[11] $V : \mathbb{R} \to \mathbb{R}$. The corresponding multiplication operator is given by the map

$$V : \mathcal{D}(V) \subset L^2(\mathbb{R}) \to L^2(\mathbb{R})\,, \qquad (V\psi)(x) := V(x)\,\psi(x)\,. \tag{3.28}$$

In (3.28), we take

$$\mathcal{D}(V) := \left\{ \psi \in L^2(\mathbb{R})\,, V\psi \in L^2(\mathbb{R}) \right\}. \tag{3.29}$$

Remark 3.9. *We make a (heuristic) remark about the momentum operator, which we will rigorously justify in the next section. Given $a \in \mathbb{R}$, we consider the translation*

[10] In Lemma 3.2(i), we proved that $(f')\hat{}\,(p) = (\mathrm{i}p)\,\hat{f}(p)$ for $f \in \mathcal{S}(\mathbb{R})$. If one considers the Fourier transform of tempered distributions, then this identity extends to the general case. In particular, it holds for L^2 functions. We will not go into further details concerning this in our module.

[11] More generally, we can consider $V : \mathbb{R}^d \to \mathbb{R}$; we always work with real-valued potentials.

operator T_a given by $(T_a f)(x) := f(x-a)$. Note that T_a is unitary since $T_a^* = T_{-a} = (T_a)^{-1}$. One can then show that there exists a self-adjoint operator P such that

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

We now show formally that $P = -\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}x}$. We interpret this as saying that conservation of momentum is associated with translation invariance (in the context of Noether's theorem). If f is an analytic function, then we can write it as the following Taylor series.

$$f(x-a) = \sum_{n \geq 0} \sum_{n \geq 0} \frac{(-a)^n}{n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n} f(x) \simeq \mathrm{e}^{-a\,\frac{\mathrm{d}}{\mathrm{d}x}} f(x)\,,$$

where the $\simeq$ should be interpreted as being formal.

From now on, we consider general d, with the understanding that we generalise all of the above constructions from $d = 1$!

We recall the definition of the classical Hamiltonian $H^{\mathrm{cl}}(x,p) = \frac{1}{2m}|p|^2 + V(x)$ in (1.15). Here, we interpret $m > 0$ as being the *mass* of the classical particle. The ***Quantum Hamiltonian*** is defined to be

$$H \equiv H^{\mathrm{Q}} := -\frac{\hbar^2}{2m}\Delta + V\,. \tag{3.30}$$

In (3.30), $\hbar > 0$ is *Planck's constant* and $\Delta = \nabla^2 = -P^2$ is the Laplacian. We can also write

$$\Delta = \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2}\,. \tag{3.31}$$

In the sequel, by using a suitable rescaling argument in space, we set Planck's constant $\hbar = 1$. Furthermore, we assume that the quantum particle has mass $m = 1/2$. This simplifies the notation a bit and (3.30) becomes

$$H \equiv H^{\mathrm{Q}} := -\Delta + V\,. \tag{3.32}$$

We work with the quantum Hamiltonian (3.32) throughout the sequel.

3.3. **The Schrödinger equation.** In this section, we study the *Schrödinger equation for a quantum particle in an external potential V*. We fix an initial wave function $\psi_0 \in L^2(\mathbb{R}^d)$ with $\|\psi_0\|_{L^2} = 1$ describing the quantum particle at time $t = 0$. Recalling (3.32), we consider the following initial value problem.

$$\begin{cases} \mathrm{i}\,\frac{\partial}{\partial t}\psi(x,t) = H\psi(x,t) = -\Delta\psi(x,t) + V(x)\,\psi(x,t) \\ \psi(x,0) = \psi_0(x)\,. \end{cases} \tag{3.33}$$

In (3.33), the spatial variable $x \in \mathbb{R}^d$. By convention, we usually consider $t \geq 0$ (although in principle we could also consider[12] $t < 0$).

In order to rigorously interpret (3.33), we need to add some assumptions on ψ (and consequently on ψ_0).

Definition 3.10 (Classical solutions of the Schrödinger equation (3.33))**.** *We say that ψ solving* (3.33) *is a* ***classical solution*** *if the following properties hold.*

(i) *For every $x \in \mathbb{R}^d$, we have $\psi(x,\cdot) \in C^1([0,\infty))$.*

(ii) *For $s > 0$, there exist $\tau > 0$, $h \in L^2(\mathbb{R}^d)$ (all of which depend on s) such that for all $x \in \mathbb{R}^d$ and $t \in [s-\tau, s+\tau]$ we have*

$$\left|\frac{\partial}{\partial t}\psi(x,t)\right| + |\psi(x,t)| \leq h(x)\,. \tag{3.34}$$

Note that the set of all functions ψ satisfying (3.34) *forms a vector space. Furthermore, it implies that*

$$\left|\frac{\partial}{\partial t}|\psi(x,t)|^2\right| = 2\left|\frac{\partial}{\partial t}\psi(x,t)\right||\psi(x,t)| \leq g(x)\,, \tag{3.35}$$

where $g = 2h^2 \in L^1(\mathbb{R}^d)$.

(iii) *For every $t \geq 0$, we have $\psi(\cdot,t) \in C^2(\mathbb{R}^d)$ and $\nabla\psi(\cdot,t) \in L^2(\mathbb{R}^d)$.*

(iv) *For each fixed $t \geq 0$, we have $V(\cdot)\psi(\cdot,t) \in L^2(\mathbb{R}^d)$.*

For now, we will assume that (3.33) has a classical solution for an appropriate choice of initial condition ψ_0 and potential V. We later show that this is indeed true.

Remark 3.11 (Remark on Definition 3.10)**.** *Before proceeding, let us comment on Definition 3.10.*

- *Condition (i) tells us that we can differentiate in t and consider the left-hand side of* (3.33)*.*
- *Condition (ii) tells us that we can differentiate under the integral sign and compute*

$$\frac{\mathrm{d}}{\mathrm{dt}}\int_{\mathbb{R}^d}|\psi(x,t)|^2\,\mathrm{d}x = \int_{\mathbb{R}^d}\frac{\partial}{\partial t}\,|\psi(x,t)|^2\,\mathrm{d}x\,. \tag{3.36}$$

  Here, we recalled Remark 3.8.
- *Condition (iii) tells us that the first term on the right-hand side of* (3.33) *is well-defined and belongs to $L^2(\mathbb{R}^d)$.*
- *Condition (iv) tells us that the second term on the right-hand side of* (3.33) *is well-defined and belongs to $L^2(\mathbb{R}^d)$. Recalling* (3.29) *(adapted to d dimensions), we can interpret this as saying that, for all fixed $t \geq 0$, we have $\psi(\cdot,t) \in \mathcal{D}(V)$.*

We now prove two important properties of classical solutions.

[12]The situation would be drastically different if one were to delete the factor of i and consider the heat equation! We will not go into this in our module, but it is a good general principle to keep in mind.

(i) The L^2 norm is conserved in time. Therefore, since $\|\psi_0\|_{L^2} = 1$ by assumption, $|\psi(x,t)|^2\,\mathrm{d}x$ is a probability density for all $t \geq 0$. See Proposition 3.12.
(ii) The evolution is *deterministic*, meaning that given a suitable initial condition, the solution is unique. See Proposition 3.13.

Proposition 3.12 (Conservation of L^2 norm for classical solutions)**.** *Let ψ be a classical solution to* (3.33)*. Then the quantity $\int_{\mathbb{R}^d} |\psi(x,t)|^2\,\mathrm{d}x$ is conserved in time.*

Proof. We recall (3.36) and use the product rule to write

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} |\psi(x,t)|^2\,\mathrm{d}x = \int_{\mathbb{R}^d}\biggl[\frac{\partial}{\partial t}\,\overline{\psi(x,t)}\,\psi(x,t) + \overline{\psi(x,t)}\,\frac{\partial}{\partial t}\,\psi(x,t)\biggr]\,\mathrm{d}x\,. \tag{3.37}$$

By using (3.33), we rewrite the right-hand side of (3.37) as

$$= \int_{\mathbb{R}^d}\biggl[-\mathrm{i}\Delta\overline{\psi(x,t)}\,\psi(x,t) + \mathrm{i}V(x)\,|\psi(x,t)|^2 + \mathrm{i}\overline{\psi(x,t)}\,\Delta\psi(x,t) - \mathrm{i}V(x)\,|\psi(x,t)|^2\biggr]\,\mathrm{d}x$$

$$= \int_{\mathbb{R}^d}\biggl[-\mathrm{i}\Delta\overline{\psi(x,t)}\,\psi(x,t) + \mathrm{i}\overline{\psi(x,t)}\,\Delta\psi(x,t)\biggr]\,\mathrm{d}x\,. \tag{3.38}$$

Note that here, we crucially used the assumption that V *is real-valued*. We now integrate by parts and write (3.38) as

$$= \mathrm{i}\int_{\mathbb{R}^d}\biggl[\overline{\nabla\psi(x,t)}\cdot\nabla\psi(x,t) - \overline{\nabla\psi(x,t)}\cdot\nabla\psi(x,t)\biggr]\,\mathrm{d}x = 0\,. \tag{3.39}$$

In (3.39), $\cdot$ denotes the inner product on $\mathbb{R}^d$. In order to obtain (3.39) from (3.38), we integrated by parts[13]. ☐

Proposition 3.13 (Uniqueness of classical solutions)**.** *If φ, ψ are two classical solutions of* (3.33) *with the same initial condition, i.e. $\varphi(x,0) = \psi(x,0)$ for all $x \in \mathbb{R}^d$, then we have $\varphi(x,t) = \psi(x,t)$ for all $x \in \mathbb{R}^d$ and $t \geq 0$.*

Proof. Let φ, ψ be as in the assumption. Since (3.33) is linear, it follows that $\varphi - \psi$ solves the Schrödinger equation with initial data zero. Note that $\varphi - \psi$ is a classical solution[14]

.By using Proposition 3.12, we get that for all $t \geq 0$

$$\int_{\mathbb{R}^d} |\varphi(x,t) - \psi(x,t)|^2\,\mathrm{d}x = \int_{\mathbb{R}^d} |\varphi(x,0) - \psi(x,0)|^2\,\mathrm{d}x = 0\,.$$

The claim now follows once we recall that $\varphi(\cdot,t) - \psi(\cdot,t) \in C^2(\mathbb{R}^d)$ by Definition 3.10 (iii). ☐

[13]It is simple to see this identity for Schwartz functions since the boundary terms vanish; the general case follows by a density argument. We will discuss this in more detail later.

[14]Conditions (i), (iii), and (iv) of Definition 3.10 are immediately verified for $\varphi - \psi$. Bound (3.34) from condition (ii) is satisfied for $\varphi - \psi$, so we deduce that (3.35) is also satisfied. Note that we can in general not deduce a bound of the form of the form(3.35) for $\varphi - \psi$ solely from such a bound for φ and ψ. That is why we need to add the slightly stronger assumption (3.34).

3.3.1. ***The time evolution in the free case.*** We conclude this section by analysing the time evolution of the free case, i.e. when we set the interaction $V = 0$. For simplicity, we assume that the initial data $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$. In this case, the Schrödinger equation (3.33) simplifies and we have

$$\mathrm{i}\,\frac{\partial}{\partial t}\,\psi(x,t) = -\Delta\psi(x,t)\,. \tag{3.40}$$

Taking Fourier transforms in the x variable in (3.40) and using Lemma 3.2 (i), we deduce that

$$\mathrm{i}\,\frac{\partial}{\partial t}\,\widehat{\psi}(p,t) = |p|^2\,\widehat{\psi}(p,t)\,, \tag{3.41}$$

where

$$\widehat{\psi}(p,t) := \frac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d}\psi(x,t)\,\mathrm{e}^{-\mathrm{i}p\cdot x}\,\mathrm{d}x\,. \tag{3.42}$$

Strictly speaking, in order to deduce (3.41) from (3.40), we need to be able to differentiate under the integral sign in (3.42), i.e. we have

$$\frac{\partial}{\partial t}\,\widehat{\psi}(p,t) = \frac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d}\frac{\partial}{\partial t}\,\psi(x,t)\,\mathrm{e}^{-\mathrm{i}p\cdot x}\,\mathrm{d}x\,. \tag{3.43}$$

For fixed $p \in \mathbb{R}^d$, (3.41) is an ordinary differential equation, whose solution is given by

$$\widehat{\psi}(p,t) = \mathrm{e}^{-\mathrm{i}t|p|^2}\,\widehat{\psi_0}(p)\,. \tag{3.44}$$

(We recall here that $\psi(\cdot,0) = \psi_0$ by construction).

Since $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$, (3.44) implies that we can indeed rigorously justify (3.43); see Exercise **??** below. In the same exercise, it is shown that $\psi(\cdot,t) \in \mathcal{S}(\mathbb{R}^d)$. We can hence apply the Fourier inversion formula (Proposition 3.3) in (3.44) to deduce that

$$\psi(x,t) = \frac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d}\mathrm{e}^{\mathrm{i}p\cdot x-\mathrm{i}t|p|^2}\,\widehat{\psi_0}(p)\,\mathrm{d}p\,. \tag{3.45}$$

From (3.44), we can immediately deduce conservation of mass for the solution (3.45) of the free Schrödinger equation; see Exercise **??** below.

We would now like to rewrite (3.45) in terms of just the x-variable, i.e. just in physical space. We prove the following lemma.

Lemma 3.14. *Let $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$. We then have for all $t > 0$ and for almost every $x \in \mathbb{R}^d$*

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

Proof. Let us fix $t > 0$. Given $\varepsilon > 0$, we define

$$\psi^{\varepsilon}(x,t) = \frac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d}\mathrm{e}^{\mathrm{i}p\cdot x-(\mathrm{i}t+\varepsilon)|p|^2}\,\widehat{\psi_0}(p)\,\mathrm{d}p\,. \tag{3.46}$$

We have that for $p \in \mathbb{R}^d$

$$\widehat{\psi^{\varepsilon}}(p,t) = \mathrm{e}^{-(\mathrm{i}t+\varepsilon)|p|^2}\,\widehat{\psi_0}(p)\,. \tag{3.47}$$

Here, we took the Fourier transform in the x-variable as in (3.42) above.

Combining (3.44), (3.47), and using Plancherel's theorem combined with the dominated convergence theorem, we deduce that

$$\lim_{\varepsilon \to 0} \|\psi^\varepsilon(\cdot,t) - \psi(\cdot,t)\|_{L^2} = 0\,. \tag{3.48}$$

We now use the convolution property of the Fourier transform (see Exercise **??** below) and Lemma 3.4 to rewrite (3.46) as

$$\psi^\varepsilon(x,t) = \frac{1}{(4\pi(\mathrm{i}t+\epsilon))^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\frac{|x-y|^2}{4(\mathrm{i}t+\varepsilon)}}\, \psi_0(y)\,\mathrm{d}y\,. \tag{3.49}$$

More precisely, we took

$$\varphi(x) = \frac{1}{(2\pi)^{d/2}} \frac{1}{(2(\mathrm{i}t+\varepsilon))^{d/2}}\, \mathrm{e}^{-\frac{|x|^2}{4(\mathrm{i}t+\varepsilon)}} \in \mathcal{S}(\mathbb{R}^d)$$

and noted that

$$(\varphi * \psi_0)\widehat{\ }(p) = \mathrm{e}^{-(\mathrm{i}t+\varepsilon)|p|^2}\, \widehat{\psi_0}(p)\,.$$

Starting from (3.49) and applying the dominated convergence theorem, we get that pointwise for all $x \in \mathbb{R}^d$

$$\lim_{\varepsilon \to 0} \psi^\varepsilon(x,t) = \frac{1}{(4\pi\mathrm{i}t)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\frac{|x-y|^2}{4t}}\, \psi_0(y)\,\mathrm{d}y\,. \tag{3.50}$$

The claim follows from (3.48) and (3.50). $\square$

Definition 3.15. *A **densely-defined operator** on $L^2(\mathbb{R}^d)$ is a pair $(A, \mathcal{D}(A))$, where the **domain** $\mathcal{D}(A)$ is a dense linear subspace of $L^2(\mathbb{R}^d)$ and $A : \mathcal{D}(A) \to L^2(\mathbb{R}^d)$ is a linear map.*

Definition 3.16. *We say that a densely-defined (linear) operator A on $L^2(\mathbb{R}^d)$ is **symmetric** if for all $f, g \in \mathcal{D}(A)$, we have*

$$\langle Af, g\rangle = \langle f, Ag\rangle\,.$$

Examples of such operators are $X, P, V, \Delta, H, \ldots$. In what follows, we refer to ***observables*** as symmetric operators on $L^2(\mathbb{R}^d)$.

3.4. **The Ehrenfest equations (not examinable).** The trajectory of a free classical particle is ballistic, i.e. it has constant velocity. Recall (1.16) with $V = 0$. We now consider the *average* motion of a quantum particle in an external potential V. Below, we show that this average follows the classical equations of motion. This yields the ***Ehrenfest equations***.

Let ψ be a classical solution of (3.33). We recall (3.2) and given an observable A, we define

$$\langle A\rangle(t) := \langle \psi(\cdot,t), A\psi(\cdot,t)\rangle \tag{3.51}$$

to be the average of the observable A at time t. Here, we implicitly assume that $\psi(\cdot,t) \in \mathcal{D}(A)$.

Let us note that when ψ is a classical solution of (3.33), we have that $|\psi(x,t)|^2\,\mathrm{d}x$ is a probability density and $\psi(\cdot,t)\in D(P)$ for all $t\geq 0$. We assume furthermore that $\psi(\cdot,t)\in\mathcal{D}(X)$ for all $t\geq 0$ (which corresponds to condition (iv) with $V=x$). In particular, we can take $A=X,P$ in (3.51).

We now show the following result.

Theorem 3.17 (The Ehrenfest equations)**.** *Let ψ be a classical solution of* (3.33) *such that there exists $g\in L^1(\mathbb{R}^d)$ such that for all $t\geq 0$*

$$|x|\left|\frac{\partial}{\partial t}|\psi(x,t)|^2\right|+\left|\frac{\partial}{\partial t}\left(\overline{\psi(x,t)}\,\nabla\psi(x,t)\right)\right|+\left|\frac{\partial}{\partial t}\left(\overline{\psi(x,t)}\,(-\Delta+V(x))\psi(x,t)\right)\right|\leq g(x)\,. \tag{3.52}$$

Furthermore, we assume that for all $t\geq 0$

$$\frac{\partial}{\partial t}\psi(x,t)\in\mathcal{D}(H)\,. \tag{3.53}$$

Then, the following identities hold.

(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) $\frac{\mathrm{d}}{\mathrm{d}t}\langle H\rangle(t)=0$.

Remark 3.18. *Before proceeding with the proof, we make a few comments on Theorem 3.17.*

- *The equations (i) and (ii) in Theorem 3.17] can be viewed as analogues of* (1.16)*, where we recall that $m=\frac{1}{2}$, thus giving a factor of* 2 *on the right-hand side in equation (i).*
- *We note that the right-hand side in (ii) is $-\langle\nabla V\rangle(t)$ and not $-V(\langle X\rangle)(t)$.*
- *Conditions* (3.52)–(3.53) *are satisfied e.g. if $\psi_0\in\mathcal{S}(\mathbb{R}^d)$ and $V=0$; see* (3.44) *and Exercise* **??** *(i). We will not go into further detail about this assumption at the moment. It is needed to justify differentiation under the integral sign as in Remark 3.8 above.*

Proof of Theorem 3.17. We prove the claim for $d=1$. The higher-dimensional version is obtained by analogous arguments.

We first prove (i).

By (3.52) and Remark 3.8, we compute

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle X\rangle(t)=\int_{\mathbb{R}}x\left[\frac{\partial}{\partial t}\overline{\psi(x,t)}\,\psi(x,t)+\overline{\psi(x,t)}\,\frac{\partial}{\partial t}\,\psi(x,t)\right]\mathrm{d}x\,,$$

which by using the Schrödinger equation and integrating by parts, analogously as in (3.37)–(3.39) in the proof of Proposition 3.12, is

$$=\mathrm{i}\int_{\mathbb{R}}\left[\overline{\nabla\psi(x,t)}\,\nabla\big(x\psi(x,t)\big)-\nabla\big(x\overline{\psi(x,t)}\big)\,\nabla\psi(x,t)\right]\mathrm{d}x\,,$$

which by the product rule is

$$= \mathrm{i} \int_{\mathbb{R}} \Big[\overline{\nabla \psi(x,t)}\, \psi(x,t) - \overline{\nabla \psi(x,t)}\, \psi(x,t) \Big] \,\mathrm{d}x \,. \tag{3.54}$$

We integrate by parts in the first term in (3.54) to deduce that

$$(3.54) = -2\mathrm{i} \int_{\mathbb{R}} \overline{\psi(x,t)}\, \nabla \psi(x,t) \,\mathrm{d}x = 2\langle P \rangle(t) \,.$$

Claim (i) follows.

We now prove claim (ii). Similarly as for (i), using (3.52) and Remark 3.8, we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \langle P \rangle(t) &= -\mathrm{i} \frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}} \overline{\psi(x,t)}\, \nabla \psi(x,t) \,\mathrm{d}x \\ &= -\mathrm{i} \int_{\mathbb{R}} \Big[\frac{\partial}{\partial t} \overline{\psi(x,t)}\, \nabla \psi(x,t) + \overline{\psi(x,t)}\, \nabla \frac{\partial}{\partial t} \psi(x,t) \Big] \,\mathrm{d}x \,. \end{aligned} \tag{3.55}$$

We use $\frac{\partial \psi}{\partial t} = \mathrm{i}\Delta \psi - \mathrm{i} V \psi$ and $\frac{\partial \overline{\psi}}{\partial t} = -\mathrm{i}\Delta \overline{\psi} + \mathrm{i} V \overline{\psi}$ in (3.55) to write

$$\begin{aligned} (3.55) = \int_{\mathbb{R}} \Big[\Big(-\Delta \overline{\psi(x,t)} + V(x) \overline{\psi(x,t)} \Big) \nabla \psi(x,t) & \\ + \overline{\psi(x,t)}\, \nabla \Big(\Delta \psi(x,t) - V(x) \psi(x,t) \Big) \Big] \,\mathrm{d}x \,. & \end{aligned} \tag{3.56}$$

Using integration by parts, we have

$$\int_{\mathbb{R}} \Big[-\Delta \overline{\psi(x,t)}\, \nabla \psi(x,t) + \overline{\psi(x,t)}\, \nabla \Delta \psi(x,t) \Big] \,\mathrm{d}x = 0 \,. \tag{3.57}$$

By substituting (3.57) into (3.56) and using the product rule, we conclude that

$$(3.56) = -\int_{\mathbb{R}} \nabla V(x) \, |\psi(x,t)|^2 \,\mathrm{d}x = -\langle \nabla V \rangle(t) \,.$$

Claim (ii) now follows.

We now prove claim (iii). We compute, using (3.52) and Remark 3.8

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle H \rangle(t) = \Big\langle \frac{\partial}{\partial t} \psi(\cdot,t), H\psi(\cdot,t) \Big\rangle + \Big\langle \psi(\cdot,t), H \frac{\partial}{\partial t} \psi(\cdot,t) \Big\rangle \,. \tag{3.58}$$

Note that all of the above calculations are well-defined by (3.53). In particular, we can use the symmetry of H (see Exercise **??** below) and use the Schrödinger equation to obtain that

$$(3.57) = 2\mathrm{Re} \Big\langle \frac{\partial}{\partial t} \psi(\cdot,t), H\psi(\cdot,t) \Big\rangle = -2\mathrm{Re}\, \mathrm{i}\, \underbrace{\langle H\psi(\cdot,t), H\psi(\cdot,t) \rangle}_{\in \mathbb{R}} = 0 \,.$$

Claim (iii) now follows. $\square$

3.5. **The Heisenberg uncertainty principle in infinite dimensions.** In this section, we study the *Heisenberg uncertainty principle* in infinite dimensions. We recall that the finite-dimensional case was considered in Proposition 2.11 above. Before proceeding to the general statement, we make an observation. For $\lambda > 0$, we take

$$f(x) = \mathrm{e}^{-\frac{\lambda |x|^2}{2}} \in \mathcal{S}(\mathbb{R}^d)\,. \tag{3.59}$$

Then, by Lemma 3.4 we have that

$$\widehat{f}(p) = \frac{1}{\lambda^{d/2}}\, \mathrm{e}^{-\frac{|p|^2}{2\lambda}}\,. \tag{3.60}$$

If we take $\lambda > 0$ very small, then f in (3.59) is going to be spread out and $\widehat{f}$ in (3.60) is going to be localised near the origin. The opposite will hold if we take $\lambda > 0$ to be large. In this example, we see that f and $\widehat{f}$ are not simultaneously localised. More generally, we could have considered the rescaling of an arbitrary $f \in \mathcal{S}(\mathbb{R}^d)$ and used Exercise **??** (iii).

We now give a qualitative formulation of the above principle.

Theorem 3.19 (The Heisenberg uncertainty principle in infinite dimensions-Version 1)**.** *Suppose that $f \in \mathcal{S}(\mathbb{R}^d)$. Then for all $1 \leq j \leq d$ and $x^0, p^0 \in \mathbb{R}$, we have*

$$\|(x_j - x^0)\, f\|_{L^2}\, \|(p_j - p^0)\, \widehat{f}\|_{L^2} \geq \frac{\|f\|_{L^2}^2}{2}\,. \tag{3.61}$$

Proof. Let us first note that, if we replace f by

$$\mathrm{e}^{\mathrm{i} x_j p^0}\, f(x + x^0\, e_j) \in \mathcal{S}(\mathbb{R}^d)\,, \tag{3.62}$$

and apply the results of Exercise **??** (i)–(ii), we can assume without loss of generality that $x^0 = p^0 = 0$. In (3.62), e_j denotes the coordinate unit vector in $\mathbb{R}^d$ that is equal to 1 in the j-th slot.

We now prove (3.61) when $x^0 = p^0 = 0$. We integrate by parts to deduce that

$$\|f\|_{L^2}^2 = \int_{\mathbb{R}^d} |f(x)|^2\, \mathrm{d}x = -\int_{\mathbb{R}^d} x_j\, \partial_j |f(x)|^2\, \mathrm{d}x = -2\mathrm{Re} \int_{\mathbb{R}^d} x_j \overline{f(x)}\, \partial_j f(x)\, \mathrm{d}x\,. \tag{3.63}$$

We could indeed integrate by parts since $f \in \mathcal{S}(\mathbb{R}^d)$. We now apply Cauchy-Schwarz in (3.63) to deduce that

$$\|f\|_{L^2}^2 \leq 2 \|x_j f\|_{L^2}\, \|\partial_j f\|_{L^2}\,. \tag{3.64}$$

Using Plancherel's theorem and Lemma 3.2 (i) for the second factor in (3.64), we conclude that

$$\|f\|_{L^2}^2 \leq 2 \|x_j f\|_{L^2}\, \|p_j \widehat{f}\|_{L^2}\,,$$

and the claim follows. □

Remark 3.20. *We recall that we interpret $|f(x)|^2\, \mathrm{d}x$ as the probability density for the position of a particle and $|\widehat{f}(p)|^2\, \mathrm{d}p$ as the probability density for its momentum. Theorem 3.19 tells us that the variance of both distributions cannot simultaneously be* *small. In particular, one cannot simultaneously measure the position and momentum of a particle with arbitrary precision.*

We now prove a generalisation of 3.19. Before stating the result, we introduce some notation. Similarly as in (3.51), given an observable A and $\psi \in \mathcal{D}(A)$ with $\|\psi\|_{L^2} = 1$, we define

$$\mathbb{E}_\psi(A) := \langle \psi, A\psi \rangle = \langle A\psi, \psi \rangle \,. \tag{3.65}$$

Similarly to (2.26), with notation as in (3.65), we consider the standard deviation

$$\Delta_\psi(A) := \sqrt{\mathbb{E}_\psi(A^2) - (\mathbb{E}_\psi(A))^2} = \|(A - \mathbb{E}_\psi(A))\psi\|_{L^2} \,. \tag{3.66}$$

From (3.66), it immediately follows that $\Delta_\psi(A) = 0$ if and only if ψ is an eigenstate of A corresponding to the eigenvalue $\mathbb{E}_\psi(A)$.

Theorem 3.21 (The Heisenberg uncertainty principle in infinite dimensions-General version)**.** *Let A, B be two symmetric operators. For all $\psi \in \mathcal{D}(AB) \cap \mathcal{D}(BA)$ with $\|\psi\|_{L^2} = 1$, we have*

$$\Delta_\psi(A)\, \Delta_\psi(B) \geq \frac{1}{2}\, |\mathbb{E}_\psi([A,B])| \,. \tag{3.67}$$

When $d = 1$, we note that Theorem 3.21 is indeed a generalisation of Theorem 3.19 if we take $A = X$, $B = P$, and $\psi \in \mathcal{D}(XP) \cap \mathcal{D}(PX) \supseteq \mathcal{S}(\mathbb{R})$, which satisfies $\mathbb{E}_\psi(X) = x^0$ and $\mathbb{E}_\psi(P) = p^0$. For the right-hand side, we recall Exercise **??** (iii). For general d, we take $A = X_j$ and $B = P_j$ (corresponding to the j-th component in position and momentum respectively) and argue analogously. The only difference is that, instead of Exercise **??** (iii), we use $[X_j, P_j] = \mathrm{i}\mathbf{1}$, which follows by using the same argument.

Proof of Theorem 3.21. Let us fix $\psi \in \mathcal{D}(AB) \cap \mathcal{D}(BA)$. We write

$$\widehat{A} := A - \mathbb{E}_\psi(A) \,, \qquad \widehat{B} := B - \mathbb{E}_\psi(B) \,. \tag{3.68}$$

Here, we take $\mathcal{D}(\widehat{A})) := \mathcal{D}(A)$ and $\mathcal{D}(\widehat{B}) := \mathcal{D}(B)$. Note that $\widehat{A}$ and $\widehat{B}$ are symmetric. Furthermore, we note that

$$\mathcal{D}(AB) \cap \mathcal{D}(BA) \subset \mathcal{D}(\widehat{A}\widehat{B}) \cap \mathcal{D}(\widehat{B}\widehat{A}) \,. \tag{3.69}$$

In order to see the last claim, it suffices to show that

$$\mathcal{D}(AB) \cap \mathcal{D}(BA) \subset \mathcal{D}(\widehat{A}\widehat{B}) \,. \tag{3.70}$$

We write

$$\widehat{A}\widehat{B} = AB - \mathbb{E}_\psi(A)B - \mathbb{E}_\psi(B)A + \mathbb{E}_\psi(A)\, \mathbb{E}_\psi(B) \,. \tag{3.71}$$

If $\psi \in \mathcal{D}(AB)$, then $AB\psi \in L^2$. Moreover $B\psi \in \mathcal{D}(A) \subset L^2$ and $A\psi \in \mathcal{D}(B) \subset L^2$. We hence obtain (3.71), which implies (3.70) and hence (3.69) by symmetry.

By (3.68) and the Cauchy-Schwarz inequality, it follows that

$$\left|\langle \widehat{A}\psi, \widehat{B}\psi \rangle\right| \leq \Delta_\psi(A)\, \Delta_\psi(B) \,. \tag{3.72}$$

Here, we used that (3.66) and (3.68) imply

$$\Delta_\psi(A) = \|\widehat{A}\psi\|_{L^2}\,, \qquad \Delta_\psi(B) = \|\widehat{B}\psi\|_{L^2}\,. \tag{3.73}$$

From (3.68), we compute

$$\widehat{A}\widehat{B} - \widehat{B}\widehat{A} = \big[A - \mathbb{E}_\psi(A)\big]\,\big[B - \mathbb{E}_\psi(B)\big] - \big[B - \mathbb{E}_\psi(B)\big]\,\big[A - \mathbb{E}_\psi(A)\big] = [A,B]\,. \tag{3.74}$$

From (3.74), we deduce that

$$\widehat{A}\widehat{B} = \frac{1}{2}\big(\widehat{A}\widehat{B} + \widehat{B}\widehat{A}\big) + \frac{1}{2}[A,B]\,. \tag{3.75}$$

By Exercise **??**, the operators $\widehat{A}\widehat{B} + \widehat{B}\widehat{A}$ and $\mathrm{i}[A,B]$ are symmetric. Therefore, we have

$$\langle \psi, (\widehat{A}\widehat{B} + \widehat{B}\widehat{A})\psi\rangle \in \mathbb{R}\,, \quad \langle \psi, \mathrm{i}[A,B]\psi\rangle \in \mathbb{R}\,. \tag{3.76}$$

Using (3.75)–(3.76) and the symmetry of $\widehat{A}$, we have that

$$\begin{aligned} |\langle \widehat{A}\psi, \widehat{B}\psi\rangle|^2 = |\langle \psi, \widehat{A}\widehat{B}\psi\rangle|^2 &= \frac{1}{4}\big|\langle \psi, (\widehat{A}\widehat{B} + \widehat{B}\widehat{A})\psi\rangle\big|^2 + \frac{1}{4}\big|\langle \psi, [A,B]\psi\rangle\big|^2 \\ &\geq \frac{1}{4}\big|\langle \psi, [A,B]\psi\rangle\big|^2 = \frac{1}{4}\big|\mathbb{E}_\psi([A,B])\big|^2\,. \end{aligned} \tag{3.77}$$

The claim follows from (3.72) and (3.77). □

We now note an alternate uncertainty principle, which holds only in dimension $d = 1$.

Theorem 3.22. *Let $f \in L^2(\mathbb{R})$ with $\|f\|_{L^2} = 1$ be given. We then have for almost all $x \in \mathbb{R}^d$*

$$|f(x)|^2 \leq \left(\int_{\mathbb{R}} p^2\,|\widehat{f}(p)|^2\,\mathrm{d}p\right)^{1/2}\,. \tag{3.78}$$

Proof. It suffices to consider the case when the right-hand side of (3.78) is finite. By a density argument, it suffices to show that (3.78) holds for all $x \in \mathbb{R}$ with $f \in \mathcal{S}(\mathbb{R})$ and $\|f\| = 1$. We consider $x \in \mathbb{R}$ and $a > 0$. We multiply and divide in (3.7) to deduce that

$$f(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \sqrt{p^2 + a^2}\,\mathrm{e}^{\mathrm{i}px}\,\widehat{f}(p)\,\frac{\mathrm{d}p}{\sqrt{p^2 + a^2}}\,. \tag{3.79}$$

We apply the Cauchy-Schwarz inequality in (3.79) to deduce that

$$|f(x)|^2 \leq \frac{1}{2\pi}\left(\int_{\mathbb{R}} (p^2 + a^2)\,|\widehat{f}(p)|^2\,\mathrm{d}p\right)\left(\int_{\mathbb{R}} \frac{\mathrm{d}p}{p^2 + a^2}\right). \tag{3.80}$$

Using the identity $\int_{\mathbb{R}} \frac{\mathrm{d}p}{p^2 + a^2} = \frac{\pi}{a}$, we can rewrite the right-hand side of (3.80) as

$$\frac{1}{2a}\int_{\mathbb{R}} p^2\,|\widehat{f}(p)|^2\,\mathrm{d}p + \frac{a}{2}\,. \tag{3.81}$$

We note that both terms in (3.81) are equal if we set

$$a = \left(\int_{\mathbb{R}} p^2 \, |\widehat{f}(p)|^2 \, \mathrm{d}p \right)^{1/2} > 0 \,. \tag{3.82}$$

Choosing a as in (3.82), we get that

$$(3.81) = \left(\int_{\mathbb{R}} p^2 \, |\widehat{f}(p)|^2 \, \mathrm{d}p \right)^{1/2} > 0 \,,$$

and the claim follows. $\square$

For a further analysis of Theorem 3.22 and for a justification that an analogous result cannot hold in higher dimensions, see Exercise **??** below. For an application of Theorem 3.22, see Exercise **??**.

3.5.1. ***Further remarks on the uncertainty principle.*** We conclude this section with further remarks on results related to the Heisenberg uncertainty principle.

(i) One can consider the appropriate dual space of $\mathcal{S}(\mathbb{R}^d)$. These are the *tempered distributions* $\mathcal{S}'(\mathbb{R}^d)$. They are studied in detail in MA4J0: Advanced Real Analysis (we will not study them further in our module). On $\mathcal{S}'(\mathbb{R}^d)$, there is a notion of Fourier transform that is consistent with (3.5). It can be shown that the Fourier transform of a *compactly supported*[15] $u \in \mathcal{S}'(\mathbb{R}^d)$ is analytic. Therefore, by analytic continuation it is not possible for both $u \in \mathcal{S}'(\mathbb{R}^d)$ and $\widehat{u} \in \mathcal{S}'(\mathbb{R}^d)$ to be compactly supported unless $u = 0$. This can be viewed as an instance of the uncertainty principle. For a detailed discussion, see [**?**, Chapter 10].

(ii) A more elementary argument can be used to show that it is not possible for $f \in L^2(\mathbb{R}^d)$ and $\widehat{f} \in L^2(\mathbb{R}^d)$ to have compact essential support unless $f = 0$, see [**?**, Theorem 7.12]. We recall that the essential support of a measurable function $f : \mathbb{R}^d \to \mathbb{C}$ is defined to be

$$\operatorname{ess\,supp} f := \mathbb{R}^d \setminus \bigcup \big\{ \Omega \subset \mathbb{R}^d \,, \Omega \text{ is open and } f = 0 \text{ almost everywhere in } \Omega \big\} \,.$$

(iii) We recall Lemma 3.4 above. The following result holds.

Theorem 3.23 (Hardy's uncertainty principle). *Let $f \in \mathcal{S}(\mathbb{R})$ be given. Suppose that there exist $C, C', a > 0$ such that*

$$|f(x)| \leq C \mathrm{e}^{-\frac{ax^2}{2}} \,, \quad |\widehat{f}(p)| \leq C' \mathrm{e}^{-\frac{p^2}{2a}} \,, \tag{3.83}$$

for all $x, p \in \mathbb{R}$. Then there exists $c \in \mathbb{C}$ such that $f(x) = c\, \mathrm{e}^{-\frac{ax^2}{2}}$ for all $x \in \mathbb{R}$.

[15] This notion is properly defined in distribution theory

For a self-contained proof of Theorem 3.23, we refer the reader to [**?**] [16].

One can interpret Theorem 3.23 as an uncertainty principle as follows. Let $f \in \mathcal{S}(\mathbb{R})$ be given. Suppose that there exist $C, C' > 0$ and $a, b > 0$ with $ab > 1$ such that

$$|f(x)| \leq C\mathrm{e}^{-\frac{ax^2}{2}}\,, \quad |\widehat{f}(p)| \leq C'\mathrm{e}^{-\frac{bp^2}{2}}\,, \tag{3.84}$$

for all $x, p \in \mathbb{R}$. Then, Theorem 3.23 implies that $f = 0$. More precisely, (3.83) is satisfied with a replaced by $a' := 1/b < a$. Therefore, we can apply Theorem 3.23 to deduce that there exists $c \in \mathbb{C}$ such that $f(x) = c\,\mathrm{e}^{-\frac{a'x^2}{2}}$ for all $x \in \mathbb{R}$. We substitute this into the first bound in (**??**) and deduce that $c = 0$.

3.6. **Exercises for Section 3.**

Exercise 3.1 (Separability of $L^2(\mathbb{R}^d)$)**.** *In this exercise, we show that $L^2(\mathbb{R}^d)$ is separable, i.e. that it has a countable dense subset. In particular, it has a countable (Hamel) basis. There are several possible approaches here. Let us first consider $d = 1$.*

1. *One possibility is to consider step functions $\chi_{[a,b]}$ for $a, b \in \mathbb{Q}$ and their linear combinations with rational coefficients.*
2. *The other is to start from the observation that $L^2([-a, a])$ is separable for all $a > 0$ by using Fourier series.*
3. *If people find other approaches, that would be great!*

The constructions all easily generalise to $L^2(\mathbb{R}^d)$.

Exercise 3.2. *We show that the Schwartz class* (3.4) *is contained $L^p(\mathbb{R}^d)$ for all $p \in [1, \infty]$.*

(i) *Explain why this follows from the definition if $p = \infty$.*

(ii) *Show the claim for $p \in [1, \infty)$.*
Here, it is helpful to note that for $f \in \mathcal{S}(\mathbb{R}^d)$, and $m \in \mathbb{N}$, there exists $C_{m,f} >$ such that for all $x \in \mathbb{R}^d$ we have $|f(x)| \leq C_{m,f}\,(1 + x_1^2 + \cdots + x_d^2)^{-m}$.

Exercise 3.3 (Further properties of the Fourier transform)**.** *Let $f \in \mathcal{S}(\mathbb{R}^d)$. Then the following identities hold.*

(i) *(Translation) $\big(f(x + a)\big)\widehat{\ }(p) = \mathrm{e}^{\mathrm{i}a\cdot p}\widehat{f}(p)$ for all $a \in \mathbb{R}^d$.*

(ii) *(Modulation) $\big(\mathrm{e}^{\mathrm{i}x\cdot a}f(x)\big)\widehat{\ }(p) = \widehat{f}(p - a)$ for all $a \in \mathbb{R}^d$.*

(iii) *(Dilation) For $\lambda > 0$, let $f_\lambda(x) := \frac{1}{\lambda^d}f(x/\lambda)$. Then, we have*

$$\widehat{f_\lambda}(p) = \widehat{f}(\lambda p)\,.$$

Exercise 3.4. *Show that for $z \in \mathbb{C}$ with $\operatorname{Re} z > 0$, we have $\mathrm{e}^{-z|x|^2/2} \in \mathcal{S}(\mathbb{R}^d)$.*

[16]One should note that, in [**?**], one uses the (harmonic analysis) convention for the Fourier transform

$$\hat{f}(p) \equiv \int_{\mathbb{R}} f(x)\,\mathrm{e}^{-2\pi\mathrm{i}px}\,\mathrm{d}x\,,$$

which is different from the (mathematical physics) convention (3.5) that we have adopted throughout the notes.

Exercise 3.5. *Verify that by taking* $\varphi = \overline{\widehat{f}}$ *in* (3.16) *indeed implies that* $\|f\|_{L^2} = \|\widehat{f}\|_{L^2}$ *for all* $f \in \mathcal{S}(\mathbb{R}^d)$.

Exercise 3.6 (Parseval's theorem)**.** *In the exercise, we prove* ***Parseval's theorem****, which can be viewed as a generalisation of Plancherel's theorem (Proposition 3.6).*

(i) *Recalling* (3.2)*, show that for all* $f, g \in \mathcal{S}(\mathbb{R}^d)$*, one has*

$$\langle \widehat{f}, \widehat{g} \rangle = \langle f, g \rangle \,. \tag{3.85}$$

(ii) *Explain how to obtain* (**??**) *for general* $f, g \in L^2(\mathbb{R}^d)$.

HINT: Use Plancherel's theorem and the polarisation identity.

Exercise 3.7 (Properties of the position, momentum, and Hamiltonian operator)**.** *Recall the position, momentum, and Hamiltonian operators* X, P, *and* H *given in* (3.21), (3.24), *and* (3.32) *above. In this exercise, we examine some properties of these operators. Throughout, we set* $d = 1$.

(i) X, P, H *are unbounded, i.e. their operator norms are infinite. (Here, let us assume* $V \geq 0$ *pointwise for simplicity).*

(ii) X, P, H *are symmetric (recall Definition 3.16).*

(iii) *Show that* $[X, P] = \mathrm{i}\mathbf{1}$. *This should be interpreted as saying that*

$$(XP - PX)\,\psi = \mathrm{i}\psi$$

for ψ *belonging to a set of functions that is dense in* $L^2(\mathbb{R})$. *Give an example of such a set.*

Exercise 3.8. *We assume that* $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$ *and we recall* (3.44).

(i) *Show that for all* $t \geq 0$*, we have* $\psi(\cdot, t) \in \mathcal{S}(\mathbb{R}^d)$.

(ii) *Show that* (3.43) *holds by computing the left-hand side of* (3.43) *from* (3.44) *and by computing the right-hand side of* (3.43) *from* (3.45). *For the latter, one should justify the differentiation under the integral sign from the assumption that* $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$.

Exercise 3.9. *We assume that* $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$.

(i) *Show that for* $\psi(x,t)$ *given by* (3.45)*, we have that the mass*

$$\int_{\mathbb{R}^d} |\psi(x,t)|^2 \,\mathrm{d}x$$

is conserved in time.

HINT: Use (3.44) *and Plancherel's theorem.*

(ii) *State and prove the corresponding conservation of energy result.*

Exercise 3.10 (Convolution and the Fourier transform)**.** *Let* $\varphi, \psi \in \mathcal{S}(\mathbb{R}^d)$ *be given. We define* $\varphi * \psi : \mathbb{R}^d \to \mathbb{C}$ *by*

$$\varphi * \psi(x) := \int_{\mathbb{R}^d} \varphi(x-y)\,\psi(y)\,\mathrm{d}y \,.$$

(i) *Show that* $\varphi * \psi \in \mathcal{S}(\mathbb{R}^d)$.
(ii) *Show that for all* $p \in \mathbb{R}^d$ *we have*

$$(\varphi * \psi)\widehat{\ }(p) = (2\pi)^{d/2}\, \widehat{\varphi}(p)\, \widehat{\psi}(p)\,.$$

Exercise 3.11 (Dispersive estimate)**.** *Suppose that* $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$ *and that* ψ *is a classical solution of* (3.33) *with* $V = 0$. *Show that for all* $t > 0$ *we have*

$$\|\psi(\cdot,t)\|_{L^\infty} \leq \frac{1}{(4\pi t)^{d/2}}\, \|\psi_0\|_{L^1}\,.$$

This is called a ***dispersive estimate****.*

Exercise 3.12. *Suppose that* $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$ *and that* ψ *is a classical solution of* (3.33) *with* $V = 0$.

(i) *Show that for* $t > 0$, *we have*

$$\psi(x,t) = \left(\frac{1}{2\mathrm{i}t}\right)^{d/2} \mathrm{e}^{\mathrm{i}\frac{|x|^2}{4t}} \left(\mathrm{e}^{\mathrm{i}\frac{|y|^2}{4t}}\, \psi_0(y)\right)\widehat{\ } \left(\frac{x}{2t}\right).$$

(ii) *Show that we have*

$$\lim_{t \to \infty} \left\| \psi(\cdot,t) - \left(\frac{1}{2\mathrm{i}t}\right)^{d/2} \mathrm{e}^{\mathrm{i}\frac{|x|^2}{4t}}\, \widehat{\psi_0}\left(\frac{x}{2t}\right) \right\|_{L^2} = 0\,.$$

Exercise 3.13 (Equality in Theorem 3.19)**.** *Let* $d = 1$ *and let* $x^0 = p^0 = 0$. *Show that equality holds in* (3.61) *when* $f(x) = \mathrm{e}^{-x^2}$, *either by a direct calculation or by using the proof of Theorem 3.19 that was given above.*

Exercise 3.14. *Let* A, B *be symmetric operators. Show that the following operators are symmetric.*

(i) $X = AB + BA$.
(ii) $Y = \mathrm{i}[A,B]$.

Exercise 3.15 (Alternative proof of Theorem 3.19)**.** *In this exercise, we outline an alternative proof of Theorem 3.19 to the one that was given above. For simplicity, we set* $d = 1$ *and* $x^0 = p^0 = 0$. *In other words, for* $f \in \mathcal{S}(\mathbb{R})$, *we want to show that*

$$\|Xf\|_{L^2}\, \|Pf\|_{L^2} \geq \frac{1}{2}\, \|f\|_{L^2}^2\,. \tag{3.86}$$

(i) *Recall the result of Exercise* **??** *(iii) and use this to write*

$$\|f\|_{L^2}^2 = -\mathrm{i}\langle f, XPf\rangle + \mathrm{i}\langle f, PXf\rangle\,. \tag{3.87}$$

Rewrite (**??**) *using the symmetry of* X *and* P *and use the Cauchy-Schwarz inequality to deduce the claim.*

(ii) *Explain why it is not possible to apply this proof in the finite-dimensional setting, i.e. why there do not exist matrices* A, B *such that* $[A,B] = c\mathbf{1}$ *for some* $c \neq 0$. *HINT: Assume that it were possible to find such* A, B *and take traces of both sides.*

Exercise 3.16. *We study in more detail Theorem 3.22.*

(i) *Formulate Theorem 3.22 for general $f \in L^2(\mathbb{R})$, i.e. without the assumption that $\|f\|_{L^2} = 1$.*

(ii) *Find a non-zero function $f \in L^2(\mathbb{R})$ such that equality holds in (3.78). You can assume for simplicity that $\|f\|_{L^2} = 1$. Furthermore, you do not need to find a closed form for this function.*

(iii) *We now show that Theorem 3.22 holds only when $d = 1$. More precisely, suppose that there exists $C > 0$ such that for all $f \in L^2(\mathbb{R}^d)$ with $\|f\|_{L^2} = 1$, we have*

$$\sup_{x\in\mathbb{R}^d} |f(x)|^2 \leq C\left(\int_{\mathbb{R}^d} |p|^2 \,|\widehat{f}(p)|^2 \,\mathrm{d}k\right)^{1/2} . \tag{3.88}$$

Then, necessarily $d = 1$.

1. *Given $\lambda > 0$, we let $g_\lambda : \mathbb{R}^d \to \mathbb{C}$ be given by $g_\lambda(x) := \frac{1}{\lambda^{d/2}} f(x/\lambda)$. Show that for all $\lambda > 0$, we have*

$$\|g_\lambda\|_{L^2} = \|f\|_{L^2}\,, \qquad \widehat{g_\lambda}(p) = \lambda^{d/2}\,\widehat{f}(\lambda p)\,.$$

2. *Let $A(f)$ and $B(f)$ denote the left and right-hand side of (**??**) respectively. Show that for all $\lambda > 0$, we have*

$$A(g_\lambda) = \frac{1}{\lambda^d}\,A(f)\,, \qquad B(g_\lambda) = \frac{1}{\lambda}B(f)\,.$$

3. *Use part 2. to deduce that (**??**) can only hold with $d = 1$.*

Exercise 3.17 (An application of Theorem 3.22)**.** *In this exercise, we use Theorem 3.22 to give an example of a potential $V : \mathbb{R} \to \mathbb{R}$ satisfying the following properties.*

(1) *V is continuous on $\mathbb{R} \setminus \{0\}$.*

(2) $\lim_{x\to 0} V(x) = -\infty$.

(3) *$-\Delta + V$ is bounded below, in the sense that there exists $C \in \mathbb{R}$ such that for all $\psi \in L^2(\mathbb{R})$ with $\|\psi\|_{L^2} = 1$, we have*

$$\langle \psi, (-\Delta + V)\psi\rangle \geq C\,.$$

We do this in several steps. Suppose that $V : \mathbb{R} \to \mathbb{R}$ satisfies conditions (1)-(2) above. We want to choose V in such a way that it satisfies condition (3) as well.

(i) *Use Theorem 3.22 and properties of the Fourier transform to obtain that for all $\psi \in L^2(\mathbb{R})$ with $\|\psi\|_{L^2} = 1$, we have*

$$\langle \psi, (-\Delta + V)\psi\rangle \geq \|\Psi\|_{L^\infty}^4 - \|V\|_{L^1}\,\|\psi\|_{L^\infty}^2\,.$$

(ii) *Use the result from part (i) to obtain that*

$$\langle \psi, (-\Delta + V)\psi\rangle \geq -\frac{1}{4}\,\|V\|_{L^1}^2\,.$$

(iii) *Use the result from part (ii) to choose an appropriate V.*

4. **The evolution operator**

In this section, we construct the quantum evolution operator U_t. We first motivate this construction from a general setup in Section **??**. In order to set everything up rigorously, we give a self-contained discussion of densely-defined operators and their spectrum in Section **??**. The construction of the evolution operator is given in Section **??**.

4.1. **Summary of the setup: Axioms of quantum mechanics.**

We recall the definition of a *densely-defined operator* A (given in Definition 3.15), by which this is a pair $(A, \mathcal{D}(A))$, where $\mathcal{D}(A)$ is a dense subset of $L^2(\mathbb{R}^d)$ and $A : \mathcal{D}(A) \to L^2(\mathbb{R}^d)$ is a linear map. We note that position (3.21) and momentum (3.24) are examples of such operators; see Exercise **??** (it is helpful to view them as prototypical examples of what we are considering in this section).

We now want to analyse the time evolution of a general quantum system (recall (3.33) above). Given an initial state of the quantum system $\psi_0 \in L^2(\mathbb{R}^d)$ with $\|\psi_0\|_{L^2} = 1$, we would expect that there exists a unique $\psi_t \equiv \psi(\cdot, t) \in L^2(\mathbb{R}^d)$ with $\|\psi_t\|_{L^2} = 1$, describing the system at time t. We formally write

$$\psi_t = U_t \, \psi_0 \,. \tag{4.1}$$

One would expect for U_t in (**??**) to be linear

$$U_t(\lambda_1 \psi_0^{(1)} + \lambda_2 \psi_0^{(2)}) = \lambda_1 \psi_t^{(1)} + \lambda_2 \psi_t^{(2)} \,,$$

where $\psi_t^{(j)}$ denotes the evolution of the system with state $\psi_0^{(j)}$ at time zero. Physically, linearity corresponds to ***superposition of states***. By construction, we would expect U_t to be a ***one-parameter unitary group***, i.e. that it satisfies the following conditions.

(i) $\|U_t \psi\|_{L^2} = \|\psi\|_{L^2}$ for all $t \in \mathbb{R}$ and $\psi \in L^2(\mathbb{R}^d)$.
(ii) $U_0 = \mathbf{1}$ and $U_{t+s} = U_t \, U_s$ for all $s, t \in \mathbb{R}$.

Furthermore, we would expect U_t to be ***strongly continuous***, i.e.

$$\lim_{t \to t_0} U_t \psi = U_{t_0} \psi \tag{4.2}$$

for all $t_0 \in \mathbb{R}$ and $\psi \in L^2(\mathbb{R}^d)$. We interpret (**??**) as the evolution satisfying the appropriate initial data.

Formally, we would like to write

$$U_t = \mathrm{e}^{-\mathrm{i}tH} \tag{4.3}$$

for an appropriate symmetric operator H. Then, $\psi_t = \mathrm{e}^{-\mathrm{i}tH} \, \psi_0$ satisfies the (Schrödinger) equation

$$\mathrm{i} \, \frac{\mathrm{d}}{\mathrm{d}t} \psi_t = H \psi_t \,,$$

for $\psi_0 \in \mathcal{D}(H)$, which one should compare with (3.33). One can (formally) obtain H from (**??**) as its **inifinitesimal generator**

$$H\psi = \lim_{t\to 0} \frac{\mathrm{i}}{t}\left(U_t\psi - \psi\right) \qquad \mathcal{D}(H) = \left\{\psi \in L^2(\mathbb{R}^d)\,, \lim_{t\to 0}\frac{\mathrm{i}}{t}\left(U_t\psi - \psi\right) \text{ exists}\right\}.$$

The operator H is called the **Hamiltonian** and it corresponds to the energy of the system.

Axioms of quantum mechanics.

1. The configuration space of a quantum system is $L^2(\mathbb{R}^d)$ (or more generally a complex separable Hilbert space). The possible states of this system are elements of $L^2(\mathbb{R}^d)$ that have norm one (i.e. the *wave functions*).
2. Each observable corresponds to a symmetric densely-defined operator A (which we henceforth identify with the operator A).
3. The expectation of an observable A in the state $\psi \in \mathcal{D}(A)$ is given by
$$\langle \psi, A\psi\rangle = \langle A\psi, \psi\rangle \in \mathbb{R}\,.$$
4. The time evolution of the system is given by a strongly continuous one-parameter semigroup U_t. The infinitesimal generator H of this group corresponds to the energy of the system.

We have studied Axioms 1–3 in the previous section. In this section, we want to make precise sense of Axiom 4.

4.2. **Densely-defined operators and their spectral theory.** Let us recall that a (linear) operator $A : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ is ***bounded*** if its operator norm[17]

$$\|A\| := \sup_{\varphi \in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\|A\varphi\|}{\|\varphi\|} < \infty\,. \tag{4.4}$$

(One should compare (**??**) with the analogous notion (2.2) in finite dimensions). In particular, $\mathcal{D}(A) = L^2(\mathbb{R}^d)$. Throughout the sequel, we denote by $\mathcal{L}(L^2(\mathbb{R}^d))$ the class of bounded linear operators on $L^2(\mathbb{R}^d)$. As we noted before, studying such operators is too restrictive. Instead, we will more generally with densely-defined operators, whose theory we now study in more detail.

Let A be a densely-defined operator. We can then define its adjoint.

Definition 4.1. *Given a densely-defined operator A, we define an operator A^* by*

$$\mathcal{D}(A^*) := \left\{\psi \in L^2(\mathbb{R}^d)\,, \exists\tilde{\psi} \in L^2(\mathbb{R}^d) \text{ such that } \langle \psi, A\varphi\rangle = \langle\tilde{\psi}, \varphi\rangle\,\forall \varphi \in \mathcal{D}(A)\right\}$$
$$A^*\psi := \tilde{\psi}\,.$$

We note that, since $\mathcal{D}(A)$ is dense, A^* is well-defined and linear. However, A^* is not necessarily densely-defined; see Exercise **??**.

[17]Using Zorn's lemma, one can show that there exists a linear operator $A : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ with infinite operator norm! We will not consider such operators in our module.

Definition 4.2 (Extension of an operator)**.** *Given operators $A, \tilde{A}$, we say that $\tilde{A}$ is an* ***extension*** *of A and we write $A \subset \tilde{A}$ if the following conditions hold.*

(i) $\mathcal{D}(A) \subset \mathcal{D}(\tilde{A})$
(ii) $A\psi = \tilde{A}\psi$ *for all* $\psi \in \mathcal{D}(A)$.

We write $A = \tilde{A}$ if both $A \subset \tilde{A}$ and $\tilde{A} \subset A$ hold. Let us recall Definition 3.16. We note that for a symmetric operator A, we have $A \subset A^*$; see Exercise **??**.

Definition 4.3. *We say that a symmetric operator A is* ***self-adjoint*** *if $A^* = A$.*

Definition 4.4. *Let $A : \mathcal{D}(A) \subset L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ be a linear operator. The* ***graph*** *of A is the set*

$$\Gamma(A) := \big\{(\varphi, A\varphi)\,, \varphi \in \mathcal{D}(A)\big\} \subset L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)\,.$$

(i) *We say that A is* ***closed*** *if $\Gamma(A) \subset L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ is closed.*
(ii) *We say that A is* ***closable*** *if A has a closed extension.*
(iii) *Every closable operator A has a smallest closed extension*[18]*, called its* ***closure****, which we denote by $\overline{A}$. (In particular, if $A \subset B$ and B is closed, then $\overline{A} \subset B$.)*

We note that, given an operator A, $\overline{\Gamma(A)}$ does not need to be the graph of an operator. For more details, we refer the reader to Exercise **??**.

If A is an operator such that $\mathcal{D}(A^*)$ is dense, then it makes sense to define $(A^*)^* \equiv A^{**}$ in light of Definition **??**. We note the following relationship between the notions of adjoint and closure.

Theorem 4.5. *Let A be a densely-defined operator on $L^2(\mathbb{R}^d)$. Then following properties hold.*

(i) *A^* is closed.*
(ii) *Suppose that $\mathcal{D}(A^*)$ is dense. Then A is closable and $\overline{A} = A^{**}$.*

Proof (optional). We first prove (i). Let us define $V : L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ by

$$V(\varphi, \psi) := (\psi, -\varphi)\,.$$

Note that V is a unitary map. We note that

$$\begin{aligned}\Gamma(A^*) &= \Big\{(\varphi, \tilde{\varphi}) \in L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)\,, \langle \varphi, A\psi\rangle_{L^2(\mathbb{R}^d)} = \langle \tilde{\varphi}, \psi\rangle_{L^2(\mathbb{R}^d)}\ \forall \psi \in \mathcal{D}(A)\Big\} \\ &= \Big\{(\varphi, \tilde{\varphi}) \in L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)\,, \langle (\varphi, \tilde{\varphi}), (\tilde{\psi}, -\psi)\rangle_{L^2(\mathbb{R}^d)\times L^2(\mathbb{R}^d)} = 0\ \forall (\psi, \tilde{\psi}) \in \Gamma(A)\Big\} \\ &= (V\Gamma(A))^{\perp}\,. \end{aligned} \tag{4.5}$$

[18] This is obtained by considering the operator whose domain is given by the intersection of the domains of all of the closed extensions and defining it appropriately on this domain. The details of this construction are left as an exercise.

In order to obtain (**??**), we use $(\cdot)^\perp$ to denote the orthogonal complement for the inner product on $L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ given by

$$\langle (\varphi, \tilde{\varphi}), (\psi, \tilde{\psi}) \rangle_{L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)} := \langle \varphi, \psi \rangle_{L^2(\mathbb{R}^d)} + \langle \tilde{\varphi}, \tilde{\psi} \rangle_{L^2(\mathbb{R}^d)} \,.$$

We recall that orthogonal complements are in general closed. More precisely, let $\mathcal{H}$ be a Hilbert space and let $S \subset \mathcal{H}$ be a nonempty subset. Denote by $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ the inner product on $\mathcal{H}$. Let (y_n) be a sequence in $S^\perp$ such that $y_n \to y$. Then, for all $x \in S$, we have

$$\langle x, y \rangle_{\mathcal{H}} = \lim_n \langle x, y_n \rangle_{\mathcal{H}} = 0 \,.$$

Therefore $y \in S^\perp$ and so $S^\perp$ is closed. Using this general fact and (**??**), claim (i) follows.

We now show claim (ii). Suppose that A^* is densely-defined. We note that $\Gamma(A) \subset L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ is a linear subset. Therefore, by the result of Exercise **??** and using $V^2 = -\mathbf{1}_{L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)}$ as well as the unitarity of V (which follow from the construction of V), we have that

$$\overline{\Gamma(A)} = ((\Gamma(A))^\perp)^\perp = ((V^2\Gamma(A))^\perp)^\perp = (V(V\Gamma(A))^\perp)^\perp \tag{4.6}$$

By using (**??**) in (**??**), we get that

$$\overline{\Gamma(A)} = (V\Gamma(A^*))^\perp \,. \tag{4.7}$$

Arguing as in (i), we deduce from (**??**) that A is closable and $A^{**} = \overline{A}$. For the latter observation we used $\Gamma(A^{**}) = \overline{\Gamma(A)}$, which follows from (**??**). □

We obtain the following result from Theorem **??**.

Corollary 4.6. *If a densely-defined operator A is self-adjoint, it is closed.*

Proof. The claim immediately follows from Theorem **??** since $A = A^*$. □

Definition 4.7 (Kernel and range of an operator)**.** *For an operator A with domain $\mathcal{D}(A)$, its **kernel** is given by*

$$\mathrm{Ker}(A) := \big\{ \varphi \in \mathcal{D}(A) \,,\, A(\varphi) = 0 \big\} \,.$$

*Its **range** is given by*

$$\mathrm{Ran}(A) := \big\{ A(\varphi) \,,\, \varphi \in \mathcal{D}(A) \big\} \,.$$

One can easily check that $\mathrm{Ker}(A)$ is a subspace of $\mathcal{D}(A)$ and $\mathrm{Ran}(A)$ is a subspace of $L^2(\mathbb{R}^d)$. We note the following conditions by which a symmetric operator A on $L^2(\mathbb{R}^d)$ is self-adjoint.

Theorem 4.8 (Criterion for self-adjointness)**.** *Let A be a symmetric operator on $L^2(\mathbb{R}^d)$. The following statements are equivalent.*

(i) *A is self-adjoint.*
(ii) *A is closed and* $\mathrm{Ker}(A^* \pm \mathrm{i}\mathbf{1}) = \{0\}$.
(iii) $\mathrm{Ran}(A \pm \mathrm{i}\mathbf{1}) = L^2(\mathbb{R}^d)$.

In (ii), we mean that $\mathrm{Ker}(A+\mathrm{i}\mathbf{1}) = \mathrm{Ker}(A-\mathrm{i}\mathbf{1}) = \{0\}$ *and in (iii), we mean that* $\mathrm{Ran}(A+\mathrm{i}\mathbf{1}) = \mathrm{Ran}(A-\mathrm{i}\mathbf{1}) = L^2(\mathbb{R}^d)$*.*

Proof. **(i) ⇒ (ii):** Suppose that (i) holds, i.e. A is self-adjoint. By Corollary **??**, we know that it is closed. Furthermore, suppose that $\varphi \in \mathcal{D}(A^*) = \mathcal{D}(A)$ is such that $A^*\varphi = \mathrm{i}\varphi$. Then, since $A^* = A$, we have $A\varphi = \mathrm{i}\varphi$. Therefore

$$-\mathrm{i}\langle\varphi,\varphi\rangle = \langle \mathrm{i}\varphi,\varphi\rangle = \langle A^*\varphi,\varphi\rangle = \langle \varphi, A\varphi\rangle = \langle\varphi,\mathrm{i}\varphi\rangle = \mathrm{i}\langle\varphi,\varphi\rangle\,.$$

It follows that $\langle\varphi,\varphi\rangle = 0$ and therefore $\varphi = 0$. Hence, $\mathrm{Ker}(A-\mathrm{i}\mathbf{1}) = \{0\}$. Analogous arguments show that $\mathrm{Ker}(A+\mathrm{i}\mathbf{1}) = \{0\}$. Hence, the first implication follows.

(ii) ⇒ (iii): Suppose that (ii) holds and we prove (iii).

Step 1: We show that $\mathrm{Ran}(A\pm\mathrm{i}\mathbf{1})$ is dense. We argue by contradiction. If $\mathrm{Ran}(A-\mathrm{i}\mathbf{1})$ were not dense, we could find a nonzero vector $\psi \in \mathrm{Ran}(A-\mathrm{i}\mathbf{1})^\perp$. In particular, such ψ would satisfy

$$\langle\psi,(A-\mathrm{i}\mathbf{1})\varphi\rangle = 0 \tag{4.8}$$

for all $\varphi \in \mathcal{D}(A)$. Recalling Definition **??**, (**??**) implies that

$$\psi \in \mathcal{D}((A-\mathrm{i}\mathbf{1})^*) = \mathcal{D}(A^*+\mathrm{i}\mathbf{1})\,.$$

Moreover, $(A-\mathrm{i}\mathbf{1})^*\psi = (A^*+\mathrm{i}\mathbf{1})\psi = 0$. This yields a contradiction with assumption (ii). Hence it follows that $\mathrm{Ran}(A-\mathrm{i}\mathbf{1})$ is indeed dense. By analogous arguments, we also have that $\mathrm{Ran}(A+\mathrm{i}\mathbf{1})$ is dense. In fact, this proof shows the following equivalences.

$$\begin{cases} \mathrm{Ran}(A-\mathrm{i}\mathbf{1}) \text{ is dense } \iff \mathrm{Ker}(A^*+\mathrm{i}\mathbf{1}) = \{0\} \\ \mathrm{Ran}(A+\mathrm{i}\mathbf{1}) \text{ is dense } \iff \mathrm{Ker}(A^*-\mathrm{i}\mathbf{1}) = \{0\}\,. \end{cases} \tag{4.9}$$

Step 2: $\mathrm{Ran}(A\pm\mathrm{i}\mathbf{1})$ is closed.

Since by Step 1, we know that $\mathrm{Ran}(A\pm\mathrm{i}\mathbf{1})$ is dense, the claim of Step 2 will imply that $\mathrm{Ran}(A\pm\mathrm{i}\mathbf{1}) = L^2(\mathbb{R}^d)$. We first show that $\mathrm{Ran}(A+\mathrm{i}\mathbf{1})$ is closed. For $\varphi \in \mathcal{D}(A) = \mathcal{D}(A+\mathrm{i}\mathbf{1})$, we have

$$\begin{aligned} \|(A+\mathrm{i}\mathbf{1})\varphi\|^2 &= \langle A\varphi+\mathrm{i}\varphi, A\varphi+\mathrm{i}\varphi\rangle = \\ &= \langle A\varphi,A\varphi\rangle - \mathrm{i}\langle\varphi,A\varphi\rangle + \mathrm{i}\langle A\varphi,\varphi\rangle + \langle\varphi,\varphi\rangle = \|A\varphi\|^2 + \|\varphi\|^2\,. \end{aligned} \tag{4.10}$$

In order to deduce the last equality in (**??**), we used the assumption that A is symmetric and $\varphi \in \mathcal{D}(A)$.

Let us consider a sequence $(\varphi_n) \in \mathcal{D}(A)$ such that

$$(A+\mathrm{i}\mathbf{1})\varphi_n \overset{L^2(\mathbb{R}^d)}{\longrightarrow} \psi\,. \tag{4.11}$$

Using (**??**), it follows that

$$\varphi_n \overset{L^2(\mathbb{R}^d)}{\longrightarrow} \varphi_\infty \tag{4.12}$$

for some $\varphi_\infty \in L^2(\mathbb{R}^d)$. (More precisely, we have that for all $m,n\in\mathbb{N}$,

$$\|\varphi_n - \varphi_m\| \leq \|(A+\mathrm{i}\mathbf{1})\varphi_m - (A+\mathrm{i}\mathbf{1})\varphi_n\|\,,$$

so (φ_n) is Cauchy and hence convergent.) Similarly, $A\varphi_n$ converges in $L^2(\mathbb{R}^d)$. Since A is closed by assumption, we deduce that

$$\varphi_\infty \in \mathcal{D}(A) = \mathcal{D}(A + \mathrm{i}\mathbf{1})\,, \qquad A\varphi_n \overset{L^2(\mathbb{R}^d)}{\longrightarrow} A\varphi_\infty\,. \tag{4.13}$$

Using (**??**)–(**??**), it follows that $(A + \mathrm{i}\mathbf{1})\varphi_\infty = \psi$. Hence $\mathrm{Ran}(A + \mathrm{i}\mathbf{1})$ is closed. The proof that $\mathrm{Ran}(A - \mathrm{i}\mathbf{1})$ is closed is analogous.

(iii) $\Rightarrow$ (i): Suppose that $\mathrm{Ran}(A \pm \mathrm{i}\mathbf{1}) = L^2(\mathbb{R}^d)$. We prove (i). We recall that we only need to show that $\mathcal{D}(A^*) \subset \mathcal{D}(A)$ (by symmetry of A we know that $\mathcal{D}(A) \subset \mathcal{D}(A^*)$). Let $\varphi \in \mathcal{D}(A^*) = \mathcal{D}(A^* - \mathrm{i}\mathbf{1})$ be given. Since $\mathrm{Ran}(A - \mathrm{i}\mathbf{1}) = L^2(\mathbb{R}^d)$, there exists $\psi \in \mathcal{D}(A - \mathrm{i}\mathbf{1}) = \mathcal{D}(A)$ such that

$$(A - \mathrm{i}\mathbf{1})\psi = (A^* - \mathrm{i}\mathbf{1})\varphi\,. \tag{4.14}$$

We know that $\mathcal{D}(A) \subset \mathcal{D}(A^*)$. Furthermore, by symmetry of A, we know that $A^* = A$ on $\mathcal{D}(A)$. In particular $A\psi = A^*\psi$. Using these observations in (**??**), we deduce that $\varphi - \psi \in \mathcal{D}(A^*) = \mathcal{D}(A^* - \mathrm{i}\mathbf{1})$ and

$$(A^* - \mathrm{i}\mathbf{1})(\varphi - \psi) = 0\,. \tag{4.15}$$

Using the assumption that $\mathrm{Ran}(A + \mathrm{i}\mathbf{1}) = L^2(\mathbb{R}^d)$ and (**??**), it follows that $\mathrm{Ker}(A^* - \mathrm{i}\mathbf{1}) = \{0\}$. Substituting this into (**??**), we deduce that $\varphi = \psi \in \mathcal{D}(A)$. Hence A is self-adjoint. $\square$

Remark 4.9. *We note that the results of Theorem* **??** *hold if we replace $A \pm \mathrm{i}\mathbf{1}$ and $A^* \pm \mathrm{i}\mathbf{1}$ with $A \pm \mathrm{i}\delta\mathbf{1}$ and $A^* \pm \mathrm{i}\delta\mathbf{1}$ respectively, for arbitrary $\delta > 0$. This is seen by using the same proof.*

Definition 4.10. *Let A be a densely-defined closed operator. The* ***resolvent set*** *of A is defined as*

$$\rho(A) := \left\{\lambda \in \mathbb{C}\,, (A - \lambda\mathbf{1})^{-1} \in \mathcal{L}(L^2(\mathbb{R}^d))\right\}.$$

In other words, $\lambda \in \rho(A)$ if and only if $A - \lambda\mathbf{1}$ is a bijection of $\mathcal{D}(A)$ onto $L^2(\mathbb{R}^d)$ with a bounded inverse. Furthermore, the function $R_A : \rho(A) \to \mathcal{L}(L^2(\mathbb{R}^d))$ given by $R_A(\lambda) := (A - \lambda\mathbf{1})^{-1}$ is called the ***resolvent*** *of A.*

Definition 4.11. *Let A be a densely-defined closed operator. The* ***spectrum*** *of A is defined as*

$$\sigma(A) := \mathbb{C} \setminus \rho(A)\,.$$

The ***point spectrum*** *of A is defined to be the set of all $\lambda \in \sigma(A)$ such that $A - \lambda\mathbf{1}$ has a nontrivial kernel, i.e. there exists nonzero $\varphi \in \mathcal{D}(A)$ such that $(A - \lambda\mathbf{1})\,\varphi = 0$. In this case, we call φ an* ***eigenvector*** *corresponding to the* ***eigenvalue*** *λ. The point spectrum of A is denoted by $\sigma_p(A)$.*

We note a useful identity for resolvents.

Lemma 4.12 (First resolvent identity)**.** *Let A be a densely-defined closed operator. Let $\lambda, \lambda' \in \rho(A)$ be given. We have that*

$$R_A(\lambda) - R_A(\lambda') = (\lambda - \lambda')\, R_A(\lambda)\, R_A(\lambda')\,. \tag{4.16}$$

Proof. We write

$$(A - \lambda'\mathbf{1}) - (A - \lambda\mathbf{1}) = (\lambda - \lambda')\mathbf{1}\,. \tag{4.17}$$

We recall that, since $\lambda, \lambda' \in \rho(A)$, $(A - \lambda\mathbf{1})$ and $(A - \lambda'\mathbf{1})$ are invertible. Therefore, applying $(A - \lambda\mathbf{1})^{-1} = R_A(\lambda)$ on the left and $(A - \lambda'\mathbf{1})^{-1} = R_A(\lambda')$ on the right in (**??**), we deduce (**??**). □

Let us note some further properties of the resolvent.

Theorem 4.13 (Properties of the resolvent)**.** *Let A be a densely-defined closed operator on $L^2(\mathbb{R}^d)$.*

(i) *The resolvent set $\rho(A)$ is open and $R_A : \rho(A) \to \mathcal{L}(L^2(\mathbb{R}^d))$ is holomorphic (i.e. it has an absolutely convergent power series expansion around every $z_0 \in \rho(A)$.*

(ii) *For all $z \in \rho(A)$, we have*

$$\|R_A(z)\| \geq \frac{1}{\operatorname{dist}(z, \sigma(A))}\,. \tag{4.18}$$

(iii) *If A is bounded, we have*

$$\Big\{z \in \mathbb{C}\,, |z| > \|A\|\Big\} \subset \rho(A)\,.$$

Proof. We first prove (i). Let $z_0 \in \rho(A)$. We fix $\lambda' = z_0$ and recursively apply Lemma **??** to deduce that for all $z \in \rho(A)$ and $n \in \mathbb{N}$ we have

$$R_A(z) = \sum_{j=0}^{n} (z - z_0)^j\, R_A(z_0)^{j+1} + (z - z_0)^{n+1}\, R_A(z)\, R_A(z_0)^{n+1}\,. \tag{4.19}$$

In light of (**??**), we define for $z \in \mathbb{C}$ (which is not necessarily an element of $\rho(A)$) and $n \in \mathbb{N}$ the quantity

$$R^{(n)}(z) := \sum_{j=0}^{n} (z - z_0)^j\, R_A(z_0)^{j+1}\,. \tag{4.20}$$

We now show that when

$$|z - z_0| < \|R_A(z_0)\|^{-1}\,, \tag{4.21}$$

we have that $z \in \rho(A)$ and (**??**) converges in norm to $R_A(z)$. In order to do this, we first use the result of Exercise **??** (i) to deduce that for all $j \in \mathbb{N}$, we have

$$\|R_A(z_0)^{j+1}\| \leq \|R_A(z_0)\|^{j+1}\,. \tag{4.22}$$

Using (**??**), we get that as $n \to \infty$, $R^{(n)}$ defined in (**??**) converges in operator norm to

$$R^{(\infty)}(z) := \sum_{j=0}^{\infty} (z - z_0)^j\, R_A(z_0)^{j+1}\,.$$

We now need to check that $R^{(\infty)}(z)$ is the inverse of $A-z\mathbf{1}$. Let $n \in \mathbb{N}$ and $\psi \in L^2(\mathbb{R}^d)$ be given[19]. We define

$$\varphi_n := R^{(n)}(z)\psi\,, \qquad \varphi_\infty := R^{(\infty)}(z)\psi\,. \tag{4.23}$$

In particular, we have $\|\varphi_n - \varphi_\infty\| \to 0$.

We use (**??**), followed by (**??**) to write

$$\begin{aligned} AR^{(n)}(z)\psi = (A - z_0\mathbf{1})R^{(n)}(z)\psi + z_0\varphi_n &= \left(\psi + \sum_{j=1}^{n}(z-z_0)^j\, R_A(z_0)^j\right) + z_0\varphi_n \\ &= \psi + (z-z_0)\varphi_{n-1} + z_0\varphi_n\,. \end{aligned} \tag{4.24}$$

From (**??**), we deduce that

$$(\varphi_n, A\varphi_n) \to (\varphi_\infty, \psi + z\varphi_\infty)\,. \tag{4.25}$$

Since A is closed[20], (**??**) implies that $\varphi_\infty \in \mathcal{D}(A)$ and $A\varphi_\infty = \psi + z\varphi_\infty$. Moreover, by recalling (**??**) we obtain that

$$(A - z\mathbf{1})R^{(\infty)}(z)\psi = \psi\,. \tag{4.26}$$

Analogously as (**??**), we start from

$$R^{(n)}(z)A\psi = \psi + (z-z_0)\varphi_{n-1} + z_0\varphi_n \tag{4.27}$$

and apply (**??**) to deduce that for all $\psi \in L^2(\mathbb{R}^d)$ we have

$$R^{(\infty)}(z)(A - z\mathbf{1})\psi = \psi\,. \tag{4.28}$$

We use (**??**) and (**??**) to deduce that $z \in \rho(A)$ and $R_A(z) = R^{(\infty)}(z)$. All of the claims from (i) now follow.

In order to deduce claim (ii), we use the arguments from the proof of (i). More precisely, we recall that for $z_0 \in \rho(A)$ and for z as in (**??**), we have $z \in \rho(A)$. (Alternatively, one can view this argument as showing that $\mathrm{dist}(z_0, \sigma(A)) \geq \frac{1}{\|R_A(z_0)\|}$). We hence deduce (**??**).

For (iii), we note that if A is bounded, the **Neumann series**

$$-\sum_{j=0}^{\infty} \frac{A^j}{z^{j+1}} \tag{4.29}$$

for $|z| > \|A\|$ gives an inverse for $(A - z\mathbf{1})$. Therefore $z \in \rho(A)$. For the full details of this proof, see Exercise **??**. $\square$

We note a useful criterion.

Proposition 4.14 (Weyl criterion)**.** *Let A be a densely-defined closed operator.*

[19] Since $R^{(n)}(z), R^{(\infty)}(z)$ are bounded, we automatically have $\psi \in \mathcal{D}(R^{(n)}) \cap \mathcal{D}(R^{(\infty)})$.

[20] Note that at this step, we are crucially using the assumption that A is closed.

(i) *If there exists a sequence* $\psi_n \in \mathcal{D}(A)$ *such that* $\|\psi_n\| = 1$ *and* $\|(A - z\mathbf{1})\psi_n\| \to 0$, *then* $z \in \sigma(A)$.
We call (ψ_n) *as above a* ***Weyl sequence associated with*** z.

(ii) *If* $z \in \partial\sigma(A) \subset \sigma(A)$ *(i.e. if* z *is in the topological boundary of* $\sigma(A)$*), the converse claim holds. In other words, there exists a Weyl sequence* (ψ_n) *associated with* z.

Proof. We first prove claim (i). We argue by contradiction. Namely, suppose that z and (ψ_n) are as in the assumptions and $z \in \rho(A)$. Then, we know that $A - z\mathbf{1}$ is invertible and $R_A(z) = (A - z\mathbf{1})^{-1}$ is bounded. We have

$$1 = \|\psi_n\| = \|R_A(z)\,(A - z\mathbf{1})\psi_n\| \leq \|R_A(z)\|\,\|(A - z\mathbf{1})\psi_n\| \to 0\,, \tag{4.30}$$

as $n \to \infty$, which is a contradiction. We note that in (**??**), we used Exercise **??** (i) for the inequality.

In order to prove claim (ii). Let $z \in \partial\sigma(A) = \partial\rho(A)$ be given. We can find a sequence (z_n) of points in $\rho(A)$ converging to z. Furthermore, using (**??**), it follows that $\|R_A(z_n)\| \to \infty$. In particular, there exists a sequence $\varphi_n \in L^2(\mathbb{R}^d) \setminus \{0\}$ satisfying the following properties.

(a) $\|R_A(z_n)\varphi_n\|/\|\varphi_n\| \to \infty$.
(b) $\|R_A(z_n)\varphi_n\| = 1$.

We note that, given any φ_n as in (a), we can achieve (b) by replacing φ_n with $c\varphi_n$ for suitable c. For such a sequence (φ_n) we define $\psi_n := R_A(z_n)\varphi_n$. By construction, we have

(a') $\|\varphi_n\| \to 0$.
(b') $\|\psi_n\| = 1$.

Consequently,

$$\|(A - z\mathbf{1})\psi_n\| = \|(A - z_n\mathbf{1})\psi_n + (z_n - z)\psi_n\| = \|\varphi_n + (z_n - z)\psi_n\| \leq \|\varphi_n\| + |z - z_n| \to 0\,,$$

as $n \to \infty$. In particular (ψ_n) is a Weyl sequence associated with z. $\square$

Remark 4.15. *In the sequel, we only refer to Proposition* **??** *(i). Claim (ii) was added just for completeness.*

4.2.1. ***Some examples***. We now compute some examples of spectra and resolvents for specific operators. Throughout the sequel, given a set $S \subset \mathbb{R}^d$, we denote by $\chi_S : \mathbb{R}^d \to \mathbb{R}$ the associated characteristic function, i.e.

$$\chi_S(x) := \begin{cases} 1 \text{ if } x \in S \\ 0 \text{ if } x \notin S\,. \end{cases}$$

Furthermore, μ denotes Lebesgue measure on $\mathbb{R}^d$ and $C(X)$ denotes the class of continuous functions on X.

Example 1 (Spectrum of the multiplication operator).

Let $f : \mathbb{R}^d \to \mathbb{C}$ be a measurable function. We consider the *multiplication operator*

$$(A\varphi)(x) := f(x)\,\varphi(x)\,, \quad \mathcal{D}(A) := \left\{\varphi \in L^2(\mathbb{R}^d)\,, f\varphi \in L^2(\mathbb{R}^d)\right\}. \tag{4.31}$$

In particular, when $d = 1$ and $f(x) = x$, (**??**) is the position operator (3.21).

Claim 1: $\mathcal{D}(A)$ is dense in $L^2(\mathbb{R}^d)$, i.e. A is a densely-defined operator.

Proof of Claim 1. Given $n \in \mathbb{N}$, let

$$\Omega_n := \left\{x \in \mathbb{R}^d\,, |f(x)| \leq n\right\}. \tag{4.32}$$

Note that $\Omega_n \subset \mathbb{R}^d$ is Lebesgue measurable. For every $\varphi \in L^2(\mathbb{R}^d)$ and $n \in \mathbb{N}$, we have

$$\chi_{\Omega_n}\varphi \in \mathcal{D}(A)\,. \tag{4.33}$$

Namely, by construction, we have that $f_n := \chi_{\Omega_n} f \in L^\infty(\mathbb{R}^d)$. Moreover, by the dominated convergence theorem, we have

$$\chi_{\Omega_n}\varphi \overset{L^2}{\to} \varphi\,.$$

It follows that $\mathcal{D}(A)$ is dense in $L^2(\mathbb{R}^d)$. □

Claim 2: A is bounded (as an operator) if and only if $f \in L^\infty(\mathbb{R}^d)$, in which case $\|A\| = \|f\|_{L^\infty}$.

Proof of Claim 2. Suppose that $f \in L^\infty(\mathbb{R}^d)$. We note that for all $\varphi \in L^2(\mathbb{R}^d)$, we have

$$\|f\varphi\|_{L^2} \leq \|f\|_{L^\infty}\|\varphi\|_{L^2} < \infty\,. \tag{4.34}$$

From (**??**), we deduce that $\mathcal{D}(A) = L^2(\mathbb{R}^d)$ and

$$\|A\| \leq \|f\|_{L^\infty}\,. \tag{4.35}$$

In order to show the converse inequality, let $\varepsilon > 0$ be given and let $\Omega_\varepsilon \subset \mathbb{R}^d$ be given by

$$\Omega^\varepsilon := \left\{x \in \mathbb{R}^d\,, |f(x)| \geq \|f\|_{L^\infty} - \varepsilon\right\}.$$

We know that $\mu(\Omega^\varepsilon) > 0$. We can then choose $\varphi = \chi_S$ for some $S \subset \Omega^\varepsilon$ with $0 < \mu(S) < \infty$. By construction, we have $\varphi \in L^2(\mathbb{R}^d) \setminus \{0\}$. Moreover,

$$\|A\varphi\|_{L^2}^2 = \int_S |f(x)|^2\,\mathrm{d}x \geq (\|f\|_{L^\infty} - \varepsilon)^2\,\mu(S) = (\|f\|_{L^\infty} - \varepsilon)^2\,\|\varphi\|_{L^2}^2\,.$$

It follows that $\|A\| \geq \|f\|_{L^\infty} - \varepsilon$ for all $\varepsilon > 0$. Therefore, letting $\varepsilon \to 0$, we get

$$\|A\| \geq \|f\|_{L^\infty}\,. \tag{4.36}$$

We deduce $\|A\| = \|f\|_{L^\infty}$ from (**??**) and (**??**). When $f \notin L^\infty(\mathbb{R}^d)$, we show in Exercise **??** that A is unbounded. □

Claim 3: A^* is given by multiplication with $\bar{f}$ and $\mathcal{D}(A^*) = \mathcal{D}(A)$. In particular, A is self-adjoint if f is real-valued.

Proof of Claim 3. We compute the adjoint of A according to Definition **??**. Let $\psi \in \mathcal{D}(A^*)$ be given. Then there exists $\tilde{\psi} \equiv A^*\psi \in L^2(\mathbb{R}^d)$ such that for all $\varphi \in \mathcal{D}(A)$, we have

$$\int_{\mathbb{R}^d} \overline{\psi(x)}\, f(x)\, \varphi(x)\, \mathrm{d}x = \int_{\mathbb{R}^d} \overline{\tilde{\psi}(x)}\, \varphi(x)\, \mathrm{d}x\,,$$

which we rewrite as

$$\int_{\mathbb{R}^d} \Big(\overline{\psi(x)}\, f(x) - \overline{\tilde{\psi}(x)}\Big)\, \varphi(x)\, \mathrm{d}x = 0\,. \tag{4.37}$$

Recalling (**??**)–(**??**), we deduce from (**??**) that for all $n \in \mathbb{N}$ and $\varphi \in L^2(\mathbb{R}^d)$

$$\int_{\mathbb{R}^d} \chi_{\Omega_n}(x)\, \Big(\overline{\psi(x)}\, f(x) - \overline{\tilde{\psi}(x)}\Big)\, \varphi(x)\, \mathrm{d}x = 0\,. \tag{4.38}$$

From (**??**), it follows that for all $n \in \mathbb{N}$, we have

$$\chi_{\Omega_n}(x)\, \Big(\overline{\psi(x)}\, f(x) - \overline{\tilde{\psi}(x)}\Big) = 0 \tag{4.39}$$

for almost every $x \in \mathbb{R}^d$. Taking complex conjugates in (**??**), we have

$$\chi_{\Omega_n}(x)\, \Big(\psi(x)\, \overline{f(x)} - \tilde{\psi}(x)\Big) = 0 \tag{4.40}$$

for almost every $x \in \mathbb{R}^d$. Letting $n \to \infty$ in (**??**), we deduce that

$$A^*\psi = \tilde{\psi} = \bar{f}\, \psi\,. \tag{4.41}$$

In particular, from (**??**), it follows that A^* is given by multiplication with $\bar{f}$. Moreover, recalling (**??**), we have

$$\mathcal{D}(A^*) = \big\{\varphi \in L^2(\mathbb{R}^d)\,, \bar{f}\varphi \in L^2(\mathbb{R}^d)\big\} = \big\{\varphi \in L^2(\mathbb{R}^d)\,, f\varphi \in L^2(\mathbb{R}^d)\big\} = \mathcal{D}(A)\,.$$

□

Example 2 (Spectrum of the momentum operator)**.**

We set $d = 1$ and consider the momentum operator (3.24) $-\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}x}$ on an appropriate domain.

Claim 1: If we take $P_0 = -\mathrm{i}\,\frac{\mathrm{d}}{\mathrm{d}x}$ with

$$\mathcal{D}(P_0) = \big\{\varphi \in C^1([0,2\pi])\,, \varphi(0) = \varphi(2\pi) = 0\big\}\,, \tag{4.42}$$

then P_0 is symmetric, but not self-adjoint.

Claim 2: If we take $P = -\mathrm{i}\,\frac{\mathrm{d}}{\mathrm{d}x}\psi$ with

$$\mathcal{D}(P) = \big\{\varphi \in C^1([0,2\pi])\,, \varphi(0) = \varphi(2\pi)\big\}\,,$$

then $\overline{P}$ is self-adjoint.

In particular, we see that the choice of domain can play a crucial role in whether a symmetric operator is self-adjoint or not. Note that P and P_0 are densely-defined[21] operators on $L^2([0,2\pi])$.

[21]Here, we slightly deviate from our convention that the big Hilbert space we are considering is $L^2(\mathbb{R}^d)$).

Proof of Claim 1. We first show that P_0 is symmetric. Let $\varphi, \psi \in \mathcal{D}(P_0)$ be given. We integrate by parts and use $\varphi(0) = \varphi(2\pi) = 0$ to deduce that

$$\langle \varphi, P_0 \psi \rangle = -\mathrm{i} \int_0^{2\pi} \overline{\varphi(x)}\, \frac{\mathrm{d}\psi}{\mathrm{d}x}(x)\, \mathrm{d}x = \mathrm{i} \int_0^{2\pi} \frac{\mathrm{d}\overline{\varphi}}{\mathrm{d}x}(x)\, \psi(x)\, \mathrm{d}x = \langle P_0 \varphi, \psi \rangle\,.$$

Hence, P_0 is symmetric as claimed.

Let $\psi \in \mathcal{D}(P_0^*)$ be given. Let $\tilde{\psi} = P_0^* \psi \in L^2([0, 2\pi])$. We then have that for all $\varphi \in \mathcal{D}(P_0)$

$$\int_0^{2\pi} \overline{\psi(x)}\, \big(-\mathrm{i}\varphi'(x)\big)\, \mathrm{d}x = \int_0^{2\pi} \overline{\tilde{\psi}(x)}\, \varphi(x)\, \mathrm{d}x\,. \tag{4.43}$$

We integrate by parts in (??) to deduce that for all $\varphi \in \mathcal{D}(P_0)$

$$\int_0^{2\pi} \varphi'(x)\, \overline{\left(\psi(x) - \mathrm{i}\int_0^x \tilde{\psi}(t)\, \mathrm{d}t\right)}\, \mathrm{d}x = 0\,. \tag{4.44}$$

We give a full justification of passing from (??) to (??) in Exercise ??.

We can rewrite (??) as

$$\psi(x) - \mathrm{i}\int_0^x \tilde{\psi}(t)\, \mathrm{d}t \in \big\{\varphi'\,, \varphi \in \mathcal{D}(P_0)\big\}^{\perp}\,. \tag{4.45}$$

On the other hand, we note that (see Exercise ??)

$$\big\{\varphi'\,, \varphi \in \mathcal{D}(P_0)\big\} = \left\{h \in C([0, 2\pi])\,, \int_0^{2\pi} h(t)\, \mathrm{d}t = 0\right\}. \tag{4.46}$$

We now show that (??) implies that the function $x \mapsto \psi(x) - \mathrm{i}\int_0^x \tilde{\psi}(t)\, \mathrm{d}t$ is constant, and therefore

$$\psi(x) = \psi(0) + \mathrm{i}\int_0^x \tilde{\psi}(t)\, \mathrm{d}t\,. \tag{4.47}$$

In order to see this, we note that the L^2 closure of (??) is equal to

$$\big\{h \in L^2([0, 2\pi])\,, \langle 1, h \rangle = 0\big\} = \{1\}^{\perp}\,, \tag{4.48}$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $L^2([0, 2\pi])$. Substituting (??) into (??)–(??), we obtain (??).

We now rewrite (??) in terms of *absolutely continuous functions* (see [?, Section 2.7]). We recall that

$$AC([0, 2\pi]) := \left\{f \in C([0, 2\pi])\,, f(x) = f(0) + \int_0^x g(t)\, \mathrm{d}t\,, g \in L^1([0, 2\pi])\right\}. \tag{4.49}$$

With notation as in (??), we write $g = f'$. We hence rewrite (??) as[22]

$$\psi \in \mathcal{H}^1([0, 2\pi]) := \big\{f \in AC([0, 2\pi])\,, f' \in L^2([0, 2\pi])\big\}\,. \tag{4.50}$$

[22] $\mathcal{H}^1([0, 2\pi])$ can be shown to be a Hilbert space when equipped with the norm given by $\|f\|_{\mathcal{H}^1}^2 := \int_0^{2\pi} (|f(t)|^2 + |f'(t)|^2)\, \mathrm{d}t$. Furthermore, it can be shown that $C^1([0, 2\pi])$ is dense in $\mathcal{H}^1([0, 2\pi])$ with respect to this norm. We accept these facts without proof.

Conversely, if we reverse the above arguments we get that for ψ as in (**??**), equation (**??**) holds with $\tilde{\psi} = -\mathrm{i}\psi' \in L^2([0,2\pi])$. In conclusion, we get that

$$P_0^* = -\mathrm{i}\,\frac{\mathrm{d}}{\mathrm{d}x}\,, \qquad \mathcal{D}(P_0^*) = \mathcal{H}^1([0,2\pi])\,.$$

It follows that P_0 is *not self-adjoint.* □

We can also compute $\overline{P_0}$. We know by Theorem **??** (ii) that $\overline{P_0} = P_0^{**} \subset P_0^*$. Here, we also used the symmetry of P_0 and Exercise **??**. We note that for all $\varphi \in \mathcal{D}(P_0) \subset \mathcal{H}^1([0,2\pi])$ and $\psi \in \mathcal{H}^1([0,2\pi])$ we have, using integration by parts

$$0 = \langle \psi, \overline{P_0}\varphi\rangle - \langle P_0^*\psi, \varphi\rangle = \mathrm{i}\big(\varphi(0)\overline{\psi(0)} - \varphi(2\pi)\overline{\psi(2\pi)}\big)\,. \tag{4.51}$$

Since $\psi(0)$ and $\psi(2\pi)$ can be arbitrarily prescribed, (**??**) implies that $\varphi(0) = \varphi(2\pi) = 0$. In particular, we obtain

$$\overline{P_0}\varphi = -\mathrm{i}\,\frac{\mathrm{d}}{\mathrm{d}x}\varphi\,, \qquad \mathcal{D}(\overline{P_0}) = \big\{\varphi \in \mathcal{H}^1([0,2\pi])\,, \varphi(0) = \varphi(2\pi) = 0\big\}\,. \tag{4.52}$$

Proof of Claim 2. Recalling Definition **??**, we have that $P_0 \subset P$. By Exercise **??**, we have $P^* \subset P_0^*$. Hence $\mathcal{D}(P^*) \subset \mathcal{D}(P_0^*) = \mathcal{H}^1([0,2\pi])$.

For all $\psi \in \mathcal{D}(P^*) \subset \mathcal{H}^1([0,2\pi])$ and $\varphi \in \mathcal{D}(P)$, we have

$$0 = \langle \psi, P\varphi\rangle - \langle P^*\psi, \varphi\rangle = \mathrm{i}\varphi(0)\,\overline{(\psi(0) - \psi(2\pi))}\,. \tag{4.53}$$

In order to deduce (**??**), we used an integration by parts step; see Exercise **??**. From (**??**), we deduce that $\psi(0) = \psi(2\pi)$. Hence, we can conclude

$$P^*\varphi = -\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}x}\varphi\,, \qquad \mathcal{D}(P^*) = \big\{\varphi \in \mathcal{H}^1([0,2\pi])\,, \varphi(0) = \varphi(2\pi)\big\}\,. \tag{4.54}$$

From (**??**) and arguing analogously as for (**??**), we can deduce that $P^* = \overline{P}$, from where we deduce that $(\overline{P})^* = P^{**} = \overline{P}$. In particular, $\overline{P}$ is self-adjoint. □

For an analysis of the eigenvectors of P and P_0, see Exercise **??**.

Example 3 (Resolvent of the multiplication operator)**.**

Let $f : \mathbb{R}^d \to \mathbb{C}$ be a measurable function. We recall the multiplication operator defined in (**??**) above.

$$(A\varphi)(x) = f(x)\,\varphi(x)\,, \quad \mathcal{D}(A) = \big\{\varphi \in L^2(\mathbb{R}^d)\,, f\varphi \in L^2(\mathbb{R}^d)\big\}\,.$$

For $z \in \mathbb{C}$, we note that $(A - z\mathbf{1})^{-1}$ is also given by the multiplication operator

$$(A - z\mathbf{1})^{-1} = \frac{1}{f(x) - z}\,\varphi(x)\,, \quad \mathcal{D}((A - z\mathbf{1})^{-1}) = \left\{\varphi \in L^2(\mathbb{R}^d)\,, \frac{1}{f - z}\,\varphi \in L^2(\mathbb{R}^d)\right\}.$$

Recalling Claim 2 from Example 1, it follows that $(A - z\mathbf{1})^{-1}$ is bounded if and only if

$$\frac{1}{f - z} \in L^\infty(\mathbb{R}^d)\,. \tag{4.55}$$

Furthermore, we note that for $\varepsilon > 0$,

$$\|(A - z\mathbf{1})^{-1}\| = \left\|\frac{1}{f - z}\right\|_{L^\infty} \leq \frac{1}{\epsilon} \tag{4.56}$$

is equivalent to

$$\mu\Big(\Big\{x \in \mathbb{R}^d \,, |f(x) - z| < \varepsilon\Big\}\Big) = 0\,. \tag{4.57}$$

Using (**??**)–(**??**), it follows that

$$\rho(A) = \bigg\{z \in \mathbb{C}\,, \exists \varepsilon > 0 : \mu\Big(\Big\{x \in \mathbb{R}^d \,, |f(x) - z| < \varepsilon\Big\}\Big) = 0\bigg\}\,. \tag{4.58}$$

Taking complements in (**??**), we get

$$\sigma(A) = \bigg\{z \in \mathbb{C}\,, \forall \varepsilon > 0 : \mu\Big(\Big\{x \in \mathbb{R}^d \,, |f(x) - z| < \varepsilon\Big\}\Big) > 0\bigg\}\,. \tag{4.59}$$

The set in (**??**) is called the ***essential range*** of f. Furthermore, if $\mu(f^{-1}(z)) > 0$, we have that z is an eigenvalue of A. A corresponding eigenvector is $\chi_B \in L^2(\mathbb{R}^d)$, where $B \subset f^{-1}(\{z\})$ is such that $0 < \mu(B) < \infty$.

Example 4 (Spectrum of the momentum operator)**.**

We set $d = 1$ and consider the momentum operator $P = -\mathrm{i}\,\frac{\mathrm{d}}{\mathrm{d}x}$ with

$$\mathcal{D}(P) = \big\{\varphi \in AC([0, 2\pi])\,, \, \varphi' \in L^2([0, 2\pi])\,, \, \varphi(0) = \varphi(2\pi)\big\}\,.$$

Here, we recall (**??**)–(**??**). The eigenvalues of P are the integers and the corresponding normalised eigenfunctions

$$u_n(x) = \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{inx}$$

form an orthonormal basis; see Exercise **??** (iii).

In order to compute the resolvent, given $\psi \in L^2([0, 2\pi])$, we want to solve the following inhomogeneous ODE

$$-\,\mathrm{i}\varphi'(x) - z\varphi(x) = \psi(x) \tag{4.60}$$

We verify that the solution to (**??**) is given by

$$\varphi(x) = \varphi(0)\mathrm{e}^{\mathrm{i}zx} + \mathrm{i}\int_0^x \mathrm{e}^{\mathrm{i}z(x-t)}\,\psi(t)\,\mathrm{d}t\,. \tag{4.61}$$

Since we want $\varphi \in \mathcal{D}(P)$, we want $\varphi(0) = \varphi(2\pi)$. Substituting this into (**??**), we get that

$$\varphi(0) = \frac{\mathrm{i}}{\mathrm{e}^{-2\pi\mathrm{i}z} - 1}\int_0^{2\pi} \mathrm{e}^{-\mathrm{i}zt}\,\psi(t)\,\mathrm{d}t\,, \quad z \in \mathbb{C} \setminus \mathbb{Z}\,. \tag{4.62}$$

In (**??**), we note that it is necessary to consider $z \in \mathbb{C} \setminus \mathbb{Z}$. We now substitute (**??**) into (**??**) and consider the domain of integration $t \leq x$ and $t > x$ separately to deduce that

$$(P - z\mathbf{1})^{-1}\psi(x) = \int_0^{2\pi} G(z, x, t)\,\psi(t)\,\mathrm{d}t\,, \tag{4.63}$$

where in (**??**), we take for $z \in \mathbb{C} \setminus \mathbb{Z}$

$$G(z,x,t) := \mathrm{e}^{\mathrm{i}z(x-t)} \begin{cases} \frac{\mathrm{i}}{1-\mathrm{e}^{2\pi \mathrm{i} z}}\,, & t \leq x \\ \frac{\mathrm{i}}{\mathrm{e}^{-2\pi \mathrm{i} z}-1}\,, & t > x\,. \end{cases}$$

It follows that $\rho(P) = \mathbb{C} \setminus \mathbb{Z}, \sigma(P) = \mathbb{Z}$.

4.3. **Construction of the evolution operator.** We want to construct the operator $U_t = \mathrm{e}^{-\mathrm{i}tA}$ for A ($\equiv H$) a self-adjoint operator on $L^2(\mathbb{R}^d)$, which is possibly unbounded. We first consider the simpler case when A is bounded.

Proposition 4.16. *Let A be a bounded operator on $L^2(\mathbb{R}^d)$. Then the following properties hold.*

(i) *Let $t \in \mathbb{C}$. The sequence of operators*

$$\left(\sum_{k=0}^{n} \frac{(-\mathrm{i}t)^k}{k!} A^k \right)_n \tag{4.64}$$

is Cauchy in operator norm. In particular, the following operator norm limit of (**??**) *is well-defined.*

$$U_t := \sum_{k=0}^{\infty} \frac{(-\mathrm{i}t)^k}{k!} A^k\,. \tag{4.65}$$

(ii) *With objects constructed as (i), we have that $U_s U_t = U_{s+t}$ for all $s,t \in \mathbb{C}$.*

(iii) *For all $t \in \mathbb{C}$, we have*

$$\frac{\mathrm{d}}{\mathrm{d}t} U_t = -\mathrm{i}AU_t\,. \tag{4.66}$$

(iv) *If A* ***is self-adjoint****, we have that for $t \in \mathbb{R}$, U_t is unitary. In particular, we have $U_t^* = U_{-t}$.*

Proof. Many of the proofs are analogous to the study of the matrix exponential in the finite-dimensional case; recall Exercise 2.13.

We first prove (i). Let $m < n$. We compute

$$\sum_{k=0}^{n} \frac{(-\mathrm{i}t)^k}{k!} A^k - \sum_{k=0}^{m} \frac{(-\mathrm{i}t)^k}{k!} A^k = \sum_{k=m+1}^{n} \frac{(-\mathrm{i}t)^k}{k!} A^k \tag{4.67}$$

Taking operator norms in (**??**) and using Exercise **??** (i), we get that the operator norm of the above expression is

$$\leq \sum_{k=m+1}^{\infty} \frac{|t|^k}{k!} \|A\|^k\,,$$

which can be made arbitrarily small if we choose m large enough. By recalling that the space of bounded operators on $L^2(\mathbb{R}^d)$ is complete with respect to the operator norm, we deduce that the limit (**??**) is well-defined.

We now show (ii). Note that, from the proof of (i), we have that the series (**??**) converges absolutely with respect to the operator norm. In particular, we can compute

$$U_s U_t = \sum_{k,\ell=0}^{\infty} \frac{(-\mathrm{i}s)^k}{k!} \frac{(-\mathrm{i}t)^\ell}{\ell!} A^k A^\ell = \sum_{n=0}^{\infty} \frac{1}{n!} A^n \sum_{m=0}^{n} \frac{n!}{(n-m)!\, m!} (-\mathrm{i}s)^{n-m} (-\mathrm{i}t)^m$$
$$= \sum_{n=0}^{\infty} \frac{(-\mathrm{i}(s+t))^n}{n!} A^n = U_{s+t}\,. \quad (4.68)$$

For the full details of (**??**), see Exercise **??** below.

We now show (iii). We compute for $\varepsilon > 0$

$$\frac{1}{\varepsilon}\big(U_{t+\varepsilon} - U_t\big) = \sum_{n=0}^{\infty} \frac{\big(-\mathrm{i}(t+\varepsilon)\big)^n - \big(-\mathrm{i}t\big)^n}{n!\varepsilon} A^n = \left[-\mathrm{i}\sum_{n=1}^{\infty} \frac{\big(-\mathrm{i}t\big)^{n-1}}{(n-1)!} A^n\right] + B_\varepsilon\,, \quad (4.69)$$

where B_ε is a bounded operator satisfying

$$\|B_\varepsilon\| \leq C(t, \|A\|)\, \varepsilon\,. \quad (4.70)$$

Here $C(t, \|A\|) > 0$ denotes a quantity depending on $t, \|A\|$. In order to obtain (**??**)–(**??**), we used Exercise **??** (i) combined with Taylor's theorem, by which we have

$$\Big|(t+\varepsilon)^n - t^n - \varepsilon n t^{n-1}\Big| \leq C(t)\varepsilon^2 n^2\,,$$

where $C(t) > 0$ denotes a quantity depending on t. Letting $\varepsilon \to 0$, we hence deduce that

$$\frac{\mathrm{d}}{\mathrm{d}t} U_t = -\mathrm{i}\sum_{n=1}^{\infty} \frac{\big(-\mathrm{i}t\big)^{n-1}}{(n-1)!} A^n = -\mathrm{i}AU_t\,. \quad (4.71)$$

In order to obtain the last step in fully rigorous fashion, see Exercise **??** below.

We now conclude with the proof of (iv). The identity $U_t^* = U_{-t}$ follows by taking adjoints in the partial sums of order n in (**??**) and then taking $n \to \infty$. By using (ii), we hence deduce that

$$U_t^*\, U_t = U_{-t}\, U_t = U_0 = \mathbf{1}\,.$$

The claim now follows. □

We note the following lemma.

Lemma 4.17. *Let A be a self-adjoint operator on $L^2(\mathbb{R}^d)$. For $\delta \in \mathbb{R}\backslash\{0\}$, the following estimates hold.*

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

Proof. We first prove (i). Let us first note that

$$\mathrm{i}\delta \in \rho(A)\,. \quad (4.72)$$

By Remark **??**, we get that $\mathrm{Ker}(A + \mathrm{i}\delta\mathbf{1}) = \{0\}$, i.e. $A + \mathrm{i}\delta\mathbf{1}$ is injective and $\mathrm{Ran}(A + \mathrm{i}\delta\mathbf{1}) = L^2(\mathbb{R}^d)$. Furthermore, we have that $A + \mathrm{i}\delta\mathbf{1}$ is closed. We now recall the

closed graph theorem (see e.g. [?, Theorem III.12]), which tells us that a linear map $T : X \to Y$ between two Banach spaces is bounded if and only if its graph $\Gamma(T) = \{(x, T(x))\, x \in X\}$ is a closed subset of $X \times Y$. Applying this result, it follows that $(A + \mathrm{i}\delta\mathbf{1})^{-1} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ is bounded. Therefore, (**??**) follows.

We hence compute

$$\|(A+\mathrm{i}\delta\mathbf{1})^{-1}\| = \sup_{\varphi \in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\|(A+\mathrm{i}\delta\mathbf{1})^{-1}\varphi\|}{\|\varphi\|} = \sup_{\psi \in \mathcal{D}(A)\setminus\{0\}} \frac{\|\psi\|}{\|(A+\mathrm{i}\delta\mathbf{1})\psi\|}\,. \tag{4.73}$$

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

$$\begin{aligned}\|(A+\mathrm{i}\delta\mathbf{1})\psi\|^2 &= \langle (A+\mathrm{i}\delta\mathbf{1})\psi, (A+\mathrm{i}\delta\mathbf{1})\psi\rangle \\ &= \|A\psi\|^2 - \mathrm{i}\delta\langle\psi, A\psi\rangle + \mathrm{i}\delta\langle A\psi,\psi\rangle + \delta^2\|\psi\|^2 = \|A\psi\|^2 + \delta^2\|\psi\|^2\,. \end{aligned} \tag{4.74}$$

Claim (i) follows by substituting (**??**) into (**??**).

We now show claim (ii) by arguing as for (**??**) to note that

$$\|A(A+\mathrm{i}\delta\mathbf{1})^{-1}\| = \sup_{\varphi \in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\|A(A+\mathrm{i}\delta\mathbf{1})^{-1}\varphi\|}{\|\varphi\|} = \sup_{\psi \in \mathcal{D}(A)} \frac{\|A\psi\|}{\|(A+\mathrm{i}\delta\mathbf{1})\psi\|}\,. \tag{4.75}$$

Claim (i) follows by substituting (**??**) into (**??**). □

Let H be a self-adjoint operator on $L^2(\mathbb{R}^d)$ and let $\lambda \in \mathbb{R}\setminus\{0\}$ be given. By (**??**) we know that $\mathrm{i}\lambda \in \rho(H)$. We can define

$$R_\lambda := (H - \mathrm{i}\lambda\mathbf{1})^{-1} = R_H(\mathrm{i}\lambda)\,. \tag{4.76}$$

R_λ commutes with H; see Exercise **??**.

Definition 4.18 (Yosida approximation of H)**.** *With notation as above, we define the* ***Yosida approximation*** *of H by*

$$H_\lambda := -\mathrm{i}\lambda H R_\lambda\,. \tag{4.77}$$

We observe that by Lemma **??** (ii), H_λ is bounded and $\|H_\lambda\| \leq |\lambda|$.

Let us prove the following useful approximation result.

Lemma 4.19. *For all $\varphi \in \mathcal{D}(H)$, we have*

$$\lim_{\lambda\to\pm\infty} H_\lambda\varphi = H\varphi\,. \tag{4.78}$$

Remark 4.20. *We can formally see that Lemma* **??** *holds from*

$$\text{“}H_\lambda = \frac{-\mathrm{i}\lambda H}{H - \mathrm{i}\lambda\mathbf{1}} = \frac{H}{-\frac{H}{\mathrm{i}\lambda} + \mathbf{1}} \overset{\lambda\to\pm\infty}{\longrightarrow} H\,.\text{”}$$

Proof of Lemma **??**. We first show that

$$\lim_{\lambda\to\pm\infty} H R_\lambda \varphi = 0\,, \tag{4.79}$$

for all $\varphi \in L^2(\mathbb{R}^d)$. Since, by Lemma **??** (ii), we have $\|HR_\lambda\| \leq 1$, it suffices to show that (**??**) holds for $\varphi \in \mathcal{D}(H)$. Indeed, using Exercise **??** followed by Lemma **??** (i), we get that for all $\varphi \in \mathcal{D}(H)$ and $\lambda \in \mathbb{R} \setminus \{0\}$

$$\|HR_\lambda \varphi\| = \|R_\lambda H\varphi\| \leq \frac{\|H\varphi\|}{|\lambda|}\,. \tag{4.80}$$

We hence deduce (**??**) from (**??**).

Starting from $\mathbf{1} = (H - \mathrm{i}\lambda\mathbf{1})R_\lambda$, we have that

$$HR_\lambda = \mathbf{1} + \mathrm{i}\lambda R_\lambda\,. \tag{4.81}$$

Recalling (**??**) and using (**??**) followed by (**??**) (with φ replaced by $H\varphi$), we get that for $\varphi \in \mathcal{D}(H)$

$$\|H\varphi - H_\lambda\varphi\| = \|(\mathbf{1} + \mathrm{i}\lambda R_\lambda)H\varphi\| = \|HR_\lambda H\varphi\| \to 0 \tag{4.82}$$

as $\lambda \to \pm\infty$. The claim follows from (**??**). □

Since H_λ given by (**??**) is bounded, we can use Proposition **??** to define for $\lambda \in \mathbb{R}\setminus\{0\}$ the operator

$$U_t^{(\lambda)} := \mathrm{e}^{-\mathrm{i}tH_\lambda}\,. \tag{4.83}$$

We note the following result (which is not obvious since in general H_λ is not symmetric).

Lemma 4.21. *For all $\lambda, t \in \mathbb{R}$ such that $t\lambda \geq 0$, with notation as in* (**??**)*, we have $\|U_t^{(\lambda)}\| \leq 1$.*

Proof. We note that, by using (**??**) in (**??**), we have

$$U_t^{(\lambda)} = \mathrm{e}^{-t\lambda - \mathrm{i}t\lambda^2 R_\lambda} = \mathrm{e}^{-t\lambda}\,\mathrm{e}^{-\mathrm{i}t\lambda^2 R_\lambda}\,. \tag{4.84}$$

Recalling the construction from Proposition **??** and applying the triangle inequality for the operator norm, and Exercise **??** (i) in (**??**), we get that

$$\|U_t^{(\lambda)}\| \leq \mathrm{e}^{-t\lambda} \sum_{n=0}^{\infty} \frac{(|t|\lambda^2)^n}{n!}\,\|R_\lambda\|^n\,. \tag{4.85}$$

By using Lemma **??** (i), we deduce that

$$(\mathbf{??}) \leq \mathrm{e}^{-t\lambda} \sum_{n=0}^{\infty} \frac{(|t|\lambda^2)^n}{n!}\,\frac{1}{|\lambda|^n} = 1\,.$$

The claim now follows. □

We now come to the construction of the evolution operator U_t.

Theorem 4.22 (The evolution operator U_t)**.** *We recall that the operator H is assumed to be a self-adjoint on $L^2(\mathbb{R}^d)$. The following claims hold.*

(i) *Let $t \geq 0$ be given. As $\lambda \to \infty$, the operator $U_t^{(\lambda)}$* ***converges strongly*** *to an operator U_t that satisfies*

$$\frac{\mathrm{d}}{\mathrm{d}t} U_t = -\mathrm{i} H U_t = -\mathrm{i} U_t H\,. \tag{4.86}$$

We interpret (??) as an identity holding on $\mathcal{D}(H)$. By the above strong convergence, we mean that for all $\varphi \in L^2(\mathbb{R}^d)$

$$\lim_{\lambda \to \infty} \|(U_t^{(\lambda)} - U_t)\varphi\| = 0\,.$$

(ii) *An analogous statement to (i) holds for $t \leq 0$ with $\lambda \to -\infty$.*

(iii) *We have $U_t^* = U_{-t}$. Furthermore, $U_0 = \mathbf{1}$ and $U_s U_t = U_{s+t}$ for all $s, t \in \mathbb{R}$. In particular, U_t is unitary.*

Proof. We prove (i). The proof of (ii) is analogous. The proof of (iii) is left as Exercise **??**. We first note that for $\lambda, \mu > 0$ and $\varphi \in L^2(\mathbb{R}^d)$

$$U_t^{(\lambda)}\varphi - U_t^{(\mu)}\varphi = -\int_0^t \frac{\mathrm{d}}{\mathrm{d}s}\, U_{t-s}^{(\lambda)}\, U_s^{(\mu)}\, \varphi \, \mathrm{d}s\,. \tag{4.87}$$

We note that

$$[H_\lambda, U_t^{(\mu)}] = 0\,. \tag{4.88}$$

For a proof of (**??**), see Exercise **??**. Differentiating in s, and using (**??**) followed by Lemma **??**, it follows that

$$\left\|\frac{\mathrm{d}}{\mathrm{d}s}\, U_{t-s}^{(\lambda)}\, U_s^{(\mu)}\, \varphi\right\| = \|U_{t-s}^{(\lambda)}\,(H_\mu - H_\lambda)\,U_s^{(\mu)}\,\varphi\| = \|U_{t-s}^{(\lambda)}\,U_s^{(\mu)}\,(H_\mu - H_\lambda)\,\varphi\| \leq \|(H_\mu - H_\lambda)\,\varphi\|\,. \tag{4.89}$$

Substituting (**??**) into (**??**) and using Lemma **??**, we deduce that for $\varphi \in \mathcal{D}(H)$

$$\|U_t^{(\lambda)}\varphi - U_t^{(\mu)}\varphi\| \leq t\,\|(H_\mu - H_\lambda)\,\varphi\| \to 0 \tag{4.90}$$

as $\lambda, \mu \to \infty$. By a density argument, we deduce that (**??**) holds for all $\varphi \in L^2(\mathbb{R}^d)$. In particular, for all $\varphi \in L^2(\mathbb{R}^d)$, $(U_t^{(\lambda)}\varphi)_{\lambda > 0}$ is Cauchy and therefore converges as $\lambda \to \infty$ to a limit which we denote by $U_t \varphi$. We note that $\varphi \mapsto U_t \varphi$ is linear by construction.

In order to see that U_t is bounded, we use continuity of the norm[23] and Lemma **??** to deduce that

$$\|U_t\| = \sup_{\|\varphi\|=1} \|U_t \varphi\| = \sup_{\|\varphi\|=1} \left\|\lim_{\lambda \to \infty} U_t^{(\lambda)}\varphi\right\| = \sup_{\|\varphi\|=1} \lim_{\lambda \to \infty} \|U_t^{(\lambda)}\varphi\| \leq 1\,. \tag{4.91}$$

Finally, for $h \in \mathbb{R}$ and $\varphi \in L^2(\mathbb{R}^d)$, we have

$$\frac{1}{h}\Big(U_{t+h}^{(\lambda)}\varphi - U_t^{(\lambda)}\varphi\Big) = -\frac{\mathrm{i}}{h} \int_t^{t+h} U_s^{(\lambda)}\, H_\lambda \varphi \, \mathrm{d}s\,. \tag{4.92}$$

[23] in the third equality in (**??**).

Letting $\lambda \to \infty$ and then letting $h \to 0$ in (**??**), we obtain that for all $\varphi \in \mathcal{D}(H)$

$$\frac{\mathrm{d}}{\mathrm{d}t} U_t \varphi = -\mathrm{i} H U_t \varphi = -\mathrm{i} U_t H \varphi \,. \tag{4.93}$$

The details of the last step are left as Exercise **??**. □

The converse of Theorem **??** is **Stone's theorem**, whose proof can be found in [**?**, Theorem 5.3].

Theorem 4.23 (Stone's theorem)**.** *Let U_t be a strongly continuous one-parameter semigroup of unitary operators on $L^2(\mathbb{R}^d)$, i.e. it satisfies the following properties.*

(i) $U_0 = \mathbf{1}$ *and* $U_s U_t = U_{s+t}$ *for all* $s, t \in \mathbb{R}$.
(ii) $\lim_{\varepsilon \to 0} \| U_{t+\varepsilon} \varphi - U_t \varphi \| = 0$ *for all* $\varphi \in L^2(\mathbb{R}^d)$.

Then there exists a self-adjoint operator A such that for all $t \in \mathbb{R}$ we have $U_t = \mathrm{e}^{-\mathrm{i}tA}$.

4.4. **Exercises for Section ??.**

Exercise 4.1 (Position and momentum as unbounded densely-defined operators)**.** *Throughout the exercise, we set $d = 1$. We recall the position operator X and momentum operator P given by* (3.21) *and* (3.24) *above.*

(i) *Show carefully that X and P are densely-defined.*
(ii) *Find an explicit $\psi \in L^2(\mathbb{R})$ such that $X(\psi) \notin L^2(\mathbb{R})$.*
(iii) *Find an explicit $\psi \in L^2(\mathbb{R})$ such that $P(\psi) \notin L^2(\mathbb{R})$.*
HINT: After solving (ii), one can obtain (iii) by using the Fourier transform.

Exercise 4.2 (The adjoint of a densely-defined operator)**.** *Let A be a densely-defined operator.*

(i) *Check carefully that A^* given by Definition* **??** *is well-defined and linear on $\mathcal{D}(A^*)$.*
(ii) *In this part of the exercise, we give an example of a densely-defined operator such that A^* is not densely-defined. Fix $d = 1$, $f \in L^\infty(\mathbb{R})$ such that $f \notin L^2(\mathbb{R})$ and a nonzero $\psi_0 \in L^2(\mathbb{R})$. We define A by*

$$\mathcal{D}(A) := \{ \psi \in L^2(\mathbb{R}) \,, |\langle f, \psi \rangle| < \infty \}$$
$$(A\psi) := \langle f, \psi \rangle \, \psi_0 \,.$$

(1) *Check that A is densely-defined. (HINT: It suffices to find a dense subset of ψ in $L^2(\mathbb{R})$ for which $\langle f, \psi \rangle$ is finite).*
(2) *Show that for all $\varphi \in \mathcal{D}(A^*)$ and $\psi \in \mathcal{D}(A)$, we have*

$$\langle \psi, A^* \varphi \rangle = \langle \psi, \langle \psi_0, \varphi \rangle f \rangle \,.$$

(3) *Use part (2) to deduce that for $\varphi \in \mathcal{D}(A^*)$ we have $A^* \varphi = \langle \psi_0, \varphi \rangle f$.*
(4) *When is $A^* \varphi$ from part (3) an element of $L^2(\mathbb{R})$. Is the set of such φ dense in $L^2(\mathbb{R})$?*
(5) *How does A^* act on $\mathcal{D}(A^*)$?*

Exercise 4.3 (Extensions of operators)**.** *We study the properties of extensions of operators.*

(i) *Check carefully that for a symmetric operator A, we have $A \subset A^*$.*
(ii) *Let A, B be densely-defined operators. Show that*
$$A \subset B \Rightarrow B^* \subset A^* \,.$$

Exercise 4.4 ($\overline{\Gamma(A)}$ does not need to be a graph of an operator)**.** *Let (φ_n) be an orthonormal basis of the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^d)$. Let $e_\infty \in \mathcal{H}$ be a vector that is not a finite linear combination of the basis vectors (φ_n) and e_∞. Given $n \in \mathbb{N}$ and $b, c_1, \ldots, c_n$, we define*
$$A(be_\infty + \sum_{i=1}^{n} c_i \varphi_i) = be_\infty \,. \tag{4.94}$$

(i) *Show that* (**??**) *is a densely-defined linear operator on $\mathcal{H}$.*
(ii) *Show that $(e_\infty, e_\infty) \in \overline{\Gamma(A)}$ and $(e_\infty, 0) \in \overline{\Gamma(A)}$.*
(iii) *Deduce that there does not exist a linear operator B on $\mathcal{H}$ such that $\overline{\Gamma(A)} = \Gamma(B)$.*

It can be shown that if A is closable, then $\Gamma(\overline{A}) = \overline{\Gamma(A)}$; see [**?**, VIII.1]. *In particular, A constructed above is an example of an operator which is not closable.*

Exercise 4.5. *Let $\mathcal{H}$ be a Hilbert space and let $S \subset \mathcal{H}$ be a subspace. In this exercise, we show that*
$$(S^\perp)^\perp = \overline{S} \,.$$

(i) *Show first that $\overline{S} \subset (S^\perp)^\perp$.*
In order to do this, let $y \in \overline{S}$ be given. Write $y = \lim_n y_n$ with $y_n \in S$ and use continuity of the inner product.
(ii) *We now show $(S^\perp)^\perp \subset \overline{S}$. Let $y \in (S^\perp)^\perp$. Since $\overline{S} \subset \mathcal{H}$ is closed, we can write $\mathcal{H} = \overline{S} \oplus \overline{S}^\perp$.*
Therefore, we can write
$$y = y_1 + y_2 \,, \tag{4.95}$$
where $y_1 \in \overline{S}$ and $y_2 \in \overline{S}^\perp$ are uniquely determined. We want to argue that $y_2 = 0$.
HINT: Take inner products of both sides of (**??**) *with y_2.*

Exercise 4.6 (Properties of the operator norm)**.** *Let A, B be bounded operators. Show that the following claims hold for the operator norm; recall* (**??**).

(i) $\|AB\| \leq \|A\| \|B\|$.
(ii) $\|A^*\| = \|A\|$.
(iii) *(Optional) Let A be a bounded operator. We prove that $\lim_n \|A^n\|^{1/n}$ exists and that it equals $\inf_n \|A^n\|^{1/n}$.*
    (1) *Let $a_n := \log \|A^n\|$. Prove that for all $m, n \in \mathbb{N}$, we have $a_{m+n} \leq a_m + a_n$. (The latter property is referred to as **subadditivity**).*

(2) *For a fixed* $m \in \mathbb{N}_+$, *set* $n = mq + r$ *were* $q, r \in \mathbb{N}_+$ *and* $0 \le r \le m - 1$. *Use (1) to show that*

$$\limsup_n \frac{a_n}{n} \le \frac{a_m}{m}\,.$$

(3) *Prove that*

$$\lim_n \frac{a_n}{n} = \inf_n \frac{a_n}{n}\,.$$

The quantity $\lim_n \|A^n\|^{1/n}$ *can be shown to equal* $\sup_{\lambda\in\sigma(A)} |\lambda|$, *which is referred to as the* ***spectral radius*** *of* A. *For a proof of this fact, see* [**?**, Theorem VI.6].

Exercise 4.7 (The Neumann series and applications to resolvents of bounded operators)**.** *In this exercise, we give the full details of the proof of Theorem* **??** *(iii).*

(i) *Let* T *be a bounded operator with* $\|T\| < 1$. *Show that the sequence of operators*

$$S_n := \sum_{j=0}^{n} T^j \tag{4.96}$$

converges in norm to an operator S_∞.
HINT: Recall (**??**) *(i).*

(ii) *Show that* $S_\infty(\mathbf{1} - T) = (\mathbf{1} - T)S_\infty = \mathbf{1}$.
We refer to

$$S_\infty = \sum_{j=0}^{\infty} T^j$$

as a ***Neumann series***.

(iii) *Suppose that* A, B *are bounded operators such that* A *is invertible and* $\|A^{-1}B\| < 1$. *Show that* $A + B$ *is invertible.*

(iv) *Fill in the details of the proof of* **??** *(iii). In particular, prove that* (**??**) *gives the resolvent of* A *for* $|z| > \|A\|$.

Exercise 4.8 (Multiplication by an unbounded function)**.** *Suppose that* $f : \mathbb{R}^d \to \mathbb{C}$ *is a measurable function* $f \notin L^\infty(\mathbb{R}^d)$. *We recall the definition of the multiplication operator* (**??**). *In this exercise, we show that* A *is not bounded.*

(i) *Given* $M > 0$, *show that* $\|A\| \ge M$.
  (1) *Explain why there exists a set* S_M *with* $0 < \mu(S_M) < \infty$ *such that* $|f(x)| \ge M$ *for all* $x \in S_M$.
  (2) *Find a nonzero* $\varphi \in L^2(\mathbb{R}^d)$ *which vanishes almost everywhere outside of* S_M *such that* $\|A\varphi\|_{L^2} \ge M\|\varphi\|_{L^2}$.
  (3) *Deduce that* $\|A\| \ge M$.

(ii) *Conclude that* $\|A\| = \infty$.

Exercise 4.9 (Justification of passing from (**??**) to (**??**))**.** *In this exercise, we give a rigorous justification of why we could integrate by parts in* (**??**) *to deduce* (**??**).

(i) *Recall* (**??**). *Let* $f \in L^2([0,2\pi])$ *and* $g \in \mathcal{D}(P_0)$ *be given. Show that*

$$\int_0^{2\pi} f(x)\, g(x)\, \mathrm{d}x = -\int_0^{2\pi} \left(\int_0^x f(t)\, \mathrm{d}t \right) g'(x)\, \mathrm{d}x\,. \tag{4.97}$$

(1) *Explain why* (**??**) *holds for* f *which is continuous.*
(2) *Argue by density to deduce* (**??**) *for general* $f \in L^2([0,2\pi])$.

(ii) *Use part (i) to justify the integration by parts used to obtain* (**??**).

Exercise 4.10. *Show that* (**??**) *holds.*

Exercise 4.11. *Show that for all* $f \in C^1([0,2\pi])$ *with* $f(0) = f(2\pi)$ *and for all* $g \in \mathcal{H}^1([0,2\pi])$, *we have*

$$\int_0^{2\pi} f'(x)\, g(x)\, \mathrm{d}x = -\int_0^{2\pi} f(x)\, g'(x)\, \mathrm{d}x + f(0)(g(2\pi) - g(0))\,.$$

Exercise 4.12. *Let* P_0, P *be given as in Example 2 above.*

(i) *Suppose that* $\lambda \in \mathbb{C}$ *is given. Show that there does not exist a nonzero* $\varphi \in \mathcal{D}(P_0)$ *such that* $P_0\varphi = \lambda\varphi$.
HINT: Solve the ODE $-\mathrm{i}\varphi'(x) = \lambda\varphi(x)$ *and consider the boundary conditions.*

(ii) *Show that there exists nonzero* $\varphi \in \mathcal{D}(P)$ *such that* $P\varphi = \lambda\varphi$ *if and only if* $\lambda \in \mathbb{Z}$.

(iii) *Verify the claim about the eigenvalues and eigenvectors given in Example 4.*

Exercise 4.13 (Justification of (**??**))**.** *In this exercise, we justify* (**??**) *in more detail. Given* $n \in \mathbb{N}$, *define*

$$U_t^{[n]} := \sum_{k=0}^{n} \frac{(-\mathrm{i}t)^k}{k!}\, A^k$$

(i) *Show that as* $n \to \infty$

$$\left\| U_s^{[n]}\, U_t^{[n]} - U_s\, U_t \right\| \to 0\,.$$

(ii) *Write* $U_s^{[n]}\, U_t^{[n]}$ *by expanding the sum as in* (**??**) *(noting now that the sum is finite and that all of the manipulations are fully rigorous). Use this to show that as* $n \to \infty$

$$\left\| U_s^{[n]}\, U_t^{[n]} - U_{s+t} \right\| \to 0\,.$$

(iii) *Conclude* (**??**) *from (i) and (ii).*

Exercise 4.14 (Justification of the last identity in (**??**))**.** *Verify the last identity in* (**??**) *rigorously.*
HINT: Truncate the sum and argue as in Exercise **??**.

Exercise 4.15. *Let* H *be a self-adjoint operator on* $L^2(\mathbb{R}^d)$ *and let* R_λ *be given as in* (**??**). *Show that* H *and* R_λ *commute on a dense subset of* $L^2(\mathbb{R}^d)$.
HINT: It suffices to show that $H(H - \mathrm{i}\lambda\mathbf{1})^{-1}\psi = (H - \mathrm{i}\lambda\mathbf{1})^{-1}H\psi$ *for all* $\psi \in \mathcal{D}(H)$. *In order to deduce the latter identity, it is helpful to write* $H = (H - \mathrm{i}\lambda\mathbf{1}) + \mathrm{i}\lambda\mathbf{1}$.

Exercise 4.16 (Proof of (**??**)). *We now prove* (**??**).

(i) *Show that* (**??**) *follows if we prove that the bounded operators HR_μ and HR_λ commute.*

(ii) *Show that the claim in (i) follows if we show that*

$$[R_\lambda, R_\mu] = 0\,. \tag{4.98}$$

HINT: Use that H commutes with R_λ and R_μ on a dense subset of $L^2(\mathbb{R}^d)$ by Exercise **??**. *Namely,* (**??**) *and Exercise* **??** *would imply that for $\psi \in \mathcal{D}(H)$*

$$HR_\mu HR_\lambda\psi = HR_\mu R_\lambda H\psi = HR_\lambda R_\mu H\psi = HR_\lambda HR_\mu\psi\,.$$

(ii) *Show that* (**??**) *follows from Lemma* **??**.

Exercise 4.17 (Optional: Proof of Theorem **??** (iii)). *Prove Theorem* **??** *(iii). (Some parts of this exercise are quite involved, so it is optional).*

(i) *We first want to show that $U_t^* = U_t$ for all $t \in \mathbb{R}$. This can be obtained by an approximation argument using the analogous claim for finite λ.*

(ii) *Second, we want to show that $U_tU_{-t} = \mathbf{1}$ for all $t \in \mathbb{R}$. This is the most difficult claim to prove. We sketch its proof. Throughout, we can assume $t > 0$.*

  (1) *Let us note that*

$$U_t\,U_{-t} = \lim_{\lambda\to\infty} \mathrm{e}^{-\mathrm{i}t(H_\lambda - H_{-\lambda})} \tag{4.99}$$

  in the strong sense.
  HINT: In order to prove (**??**), *argue similarly as in the discussion preceding* (**??**).
  We would now like to see that the right-hand side of (**??**) *is indeed the identity operator.*

  (2) *Using Lemma* **??** *that fo all $\varphi \in L^2(\mathbb{R}^d)$, we have*

$$\lim_{\lambda\to\infty}(H_\lambda - H_{-\lambda})\varphi = 0\,. \tag{4.100}$$

  Using (**??**) *and the Uniform boundedness (Banach-Steinhaus) theorem*[24], *it follows that there exists $C > 0$ such that*

$$\|H_\lambda - H_{-\lambda}\| \leq C \tag{4.101}$$

  for all $\lambda \in \mathbb{R}$.

  (3) *We now write for $\varphi \in L^2(\mathbb{R}^d)$*

$$\mathrm{e}^{-\mathrm{i}t(H_\lambda - H_{-\lambda})}\,\varphi = \varphi + \sum_{n=1}^{\infty}\frac{t^n}{n!}(H_\lambda - H_{-\lambda})^n\,\varphi\,. \tag{4.102}$$

---

[24] If $A_n : \mathcal{H} \to \mathcal{H}$ are bounded operators such that for all $f \in \mathcal{H}$ we have $\sup_n \|A_nf\| < \infty$, then $\sup_n \|A_n\| < \infty$; see [**?**, Theorem 0.44] for a proof.

We iteratively use Exercise **??** *(i) and* (**??**) *to estimate*

$$\left\| \sum_{n=1}^{\infty} \frac{t^n}{n!} (H_\lambda - H_{-\lambda})^n \varphi \right\| \leq \sum_{n=1}^{\infty} \frac{t^n}{n!} C^{n-1} \|(H_\lambda - H_{-\lambda}) \varphi\| , \tag{4.103}$$

which converges to zero as $\lambda \to \infty$ *by* (**??**).

(4) *We hence deduce the claim from* (**??**), (**??**), *and* (**??**).

(iii) *In order to show that* $U_s U_t = U_{s+t}$ *for all* $s, t \in \mathbb{R}$, *we first note that the identity holds for* $s, t > 0$ *and* $s, t < 0$ *since it holds for* $U_s^{(\lambda)}$ *and* $U_t^{(\lambda)}$. *For* $s < 0 < t$, *with* $|s| < t$, *use* $U_t = U_{-s} U_{s+t}$ *and* $U_s U_{-s} = \mathbf{1}$ *to obtain the claim.*

Exercise 4.18 (Optional: Justification of (**??**))**.** *Explain how* (**??**) *implies* (**??**).
HINTS: It is helpful to use Lemma **??** *and Lemma* **??**.
As a first step, it is good to note that with $\varphi \in \mathcal{D}(H)$, *the right-hand side of* (**??**) *can be written as*

$$-\frac{\mathrm{i}}{h} \int_t^{t+h} U_s^{(\lambda)} H\varphi \, \mathrm{d}s + \varepsilon_\lambda , \tag{4.104}$$

where $\varepsilon_\lambda \to 0$ *as* $\lambda \to \infty$. *For the second step, one wants to show that*

$$-\frac{\mathrm{i}}{h} \int_t^{t+h} (U_s - U_s^{(\lambda)}) H\varphi \, \mathrm{d}s \to 0 . \tag{4.105}$$

as $\lambda \to \infty$. *We view* (**??**) *as an integral of an* $L^2(\mathbb{R}^d)$ *valued function, and we want to use the dominated convergence theorem in this context (one can use this without proof). In order to do so, we note that*

$$\|(U_s - U_s^{(\lambda)}) H\varphi\|_{L^2} \leq 2 \|H\varphi\|_{L^2} .$$

Letting $\lambda \to \infty$, *we therefore deduce that for* $\varphi \in \mathcal{D}(H)$, *we have*

$$\frac{1}{h}\Big(U_{t+h}\varphi - U_t\varphi\Big) = -\frac{\mathrm{i}}{h} \int_t^{t+h} U_s H\varphi \, \mathrm{d}s . \tag{4.106}$$

We note that (**??**) *implies the claim if we show that the integrand on the right-hand side is continuous (in the* L^2 *norm). In order to do this, it suffices to show that for* $\psi \in L^2(\mathbb{R}^d)$, *we have*

$$\|U_t \psi - U_s \psi\|_{L^2} \to 0 \tag{4.107}$$

as $|t - s| \to 0$. *Note that* (**??**) *follows if we show that for* $\lambda > 0$

$$\|U_t^{(\lambda)} \psi - U_s^{(\lambda)} \psi\|_{L^2} \to 0 \tag{4.108}$$

as $|t - s| \to 0$. *In order to show* (**??**), *we use the fundamental theorem of calculus as in* (**??**) *and write*

$$U_t^{(\lambda)}\psi - U_s^{(\lambda)}\psi = -\mathrm{i} \int_s^t U_\tau^{(\lambda)} H_\lambda \psi \, \mathrm{d}\tau ,$$

from where we deduce that

$$\|U_t^{(\lambda)}\psi - U_s^{(\lambda)}\psi\| \leq |t - s| \, \|H_\lambda \psi\|_{L^2} . \tag{4.109}$$

Since H_λ is bounded, we obtain (**??**) *from* (**??**)*. Finally, the equality of the last two expressions in* (**??**) *follows from* (**??**) *by letting $\lambda = \mu \to \infty$.*

5. Applications and examples

In this section, we give applications of the theory developed in Sections 3–**??**. Furthermore, we give several examples. Section **??** is devoted to understanding when the operator $H = -\Delta + V$ is self-adjoint, thus allowing us to construct the evolution operator U_t as in Theorem **??**.

5.1. Construction of self-adjoint operators.

We recall from Proposition 3.6 that, given $f \in L^2(\mathbb{R}^d)$, we can define its Fourier transform $\widehat{f} \in L^2(\mathbb{R}^d)$.

Definition 5.1 (Sobolev spaces)**.** *Given $s \geq 0$, we define the* ***Sobolev space of order*** *s, which we denote by $H^s \equiv H^s(\mathbb{R}^d)$ as*

$$H^s(\mathbb{R}^d) := \Big\{ f \in L^2(\mathbb{R}^d)\,, \;\; (1+|p|^2)^{s/2}\,\widehat{f}(p) \in L^2(\mathbb{R}^d) \Big\}\,. \tag{5.1}$$

We define the associated ***Sobolev norm*** *by*

$$\|f\|_{H^s(\mathbb{R}^d)} := \|(1+|p|^2)^{s/2}\,\widehat{f}\|_{L^2(\mathbb{R}^d)}\,. \tag{5.2}$$

It can be shown that $H^s(\mathbb{R}^d)$ is a Hilbert space with inner product given by

$$\langle f, g\rangle_{H^s} = \int_{\mathbb{R}^d} (1+|p|^2)^s\,\overline{\widehat{f}(p)}\,\widehat{g}(p)\,\mathrm{d}p\,.$$

In particular, the Hilbert space norm corresponds to (**??**).

Remark 5.2. *It is possible to consider Sobolev spaces with $s < 0$. For this, we need to rigorously define $\widehat{f}$ for objects f which are not in L^2 (and sometimes not even functions!). This is done in full rigour in Advanced Real Analysis (MA4J0). For the purposes of our module, Definition* **??** *will suffice and we will not consider Sobolev spaces of order $s < 0$. For a self-contained account of the general theory of Sobolev spaces, one can consult* [**?**, Section 9.2].

Remark 5.3. *In light of Proposition 3.6 and Lemma 3.2 (i), when $s = k \in \mathbb{N}$, we can formally consider* (**??**) *as the space of L^2 functions whose derivatives of order at most k belong to $L^2(\mathbb{R}^d)$.*

We note the following result, which is a special case of **Sobolev embedding**.

Lemma 5.4. *Let $s > d/2$ be given. We have that $H^s(\mathbb{R}^d) \subset L^\infty(\mathbb{R}^d)$.*

Lemma **??** is proved in Exercise **??**.

We first consider $H = -\Delta$. The self-adjointness of H with domain $H^2(\mathbb{R}^d)$ is given by the following lemma.

Proposition 5.5. *If $\mathcal{D}(\Delta) = H^2(\mathbb{R}^d)$, then Δ is self-adjoint.*

Proof. Recalling Lemma 3.2 (i) and (3.31), we deduce that for $\varphi \in \mathcal{S}(\mathbb{R}^d)$, we have for all $p \in \mathbb{R}^d$

$$(\Delta\varphi)\widehat{\;}(p) = -|p|^2\,\widehat{\varphi}(p)\,. \tag{5.3}$$

By Plancherel's theorem and density of $\mathcal{S}(\mathbb{R}^d)$ in $H^2(\mathbb{R}^d)$, we get that (**??**) holds as an identity of L^2 functions whenever $\varphi \in H^2(\mathbb{R}^d)$. Using (**??**) and Parseval's theorem (recall Exercise **??**), it follows that Δ is symmetric[25]. Therefore, we need to show that $\varphi \in \mathcal{D}(\Delta^*)$ implies that $\varphi \in H^2(\mathbb{R}^d) = \mathcal{D}(\Delta)$.

By Definition **??**, using (**??**) and Parseval's theorem, it follows that given $\varphi \in \mathcal{D}(\Delta^*)$, there exists $\psi \in L^2(\mathbb{R}^d)$ such that

$$-\int_{\mathbb{R}^d} |p|^2\,\overline{\widehat{\varphi}(p)}\,\widehat{\zeta}(p)\,\mathrm{d}p = \int_{\mathbb{R}^d} \overline{\widehat{\psi}(p)}\,\widehat{\zeta}(p)\,\mathrm{d}p \tag{5.4}$$

for all $\zeta \in H^2(\mathbb{R}^d)$. From (**??**), we deduce that for almost every $p \in \mathbb{R}^d$

$$\widehat{\psi}(p) = -|p|^2\,\widehat{\varphi}(p)\,. \tag{5.5}$$

Since $\psi \in L^2(\mathbb{R}^d)$, we deduce from (**??**) that $\varphi \in H^2(\mathbb{R}^d) = \mathcal{D}(\Delta)$. The claim then follows. $\square$

Before stating the next result, we note a general criterion for self-adjointness.

Proposition 5.6. *Let A be a closed symmetric operator. Then A is self-adjoint if and only if there exists $\lambda \in \mathbb{C}$ such that $\lambda \in \rho(A)$ and $\bar{\lambda} \in \rho(A)$.*

Proof. If A is self-adjoint, then we know that by (**??**) that $\mathbb{C} \setminus \mathbb{R} \subset \rho(A)$, we obtain that there exists $\lambda \in \mathbb{C}$ such that $\lambda \in \rho(A)$ and $\bar{\lambda} \in \rho(A)$.

Conversely, suppose that there exists a $\lambda \in \mathbb{C}$ with the above properties. Since A is symmetric, we know that A^* is an extension of A. We want to show that

$$\mathcal{D}(A^*) \subset \mathcal{D}(A)\,. \tag{5.6}$$

Since $\lambda \in \rho(A)$, we note that[26] $\mathrm{Ran}(A - \lambda \mathbf{1}) = L^2(\mathbb{R}^d)$. In particular, given $\varphi \in \mathcal{D}(A^*)$, there exists $\psi \in \mathcal{D}(A)$ such that

$$(A^* - \lambda \mathbf{1})\varphi = (A - \lambda \mathbf{1})\psi\,. \tag{5.7}$$

Recalling that A is symmetric, we can rewrite (**??**) as

$$(A^* - \lambda \mathbf{1})\varphi = (A^* - \lambda \mathbf{1})\psi\,. \tag{5.8}$$

In particular, from (**??**) it follows that

$$(A^* - \lambda \mathbf{1})(\varphi - \psi) = 0\,. \tag{5.9}$$

We obtain (**??**) from (**??**) if we show that

$$\mathrm{Ker}(A^* - \lambda \mathbf{1}) = \{0\}\,. \tag{5.10}$$

In order to show (**??**), we note that for all densely-defined operators B, we have

$$\overline{\mathrm{Ran}\,B} \oplus \mathrm{Ker}\,B^* = L^2(\mathbb{R}^d)\,. \tag{5.11}$$

[25] Alternatively, we can integrate by parts to note that for $\varphi, \psi \in \mathcal{S}(\mathbb{R}^d)$ we have $\int_{\mathbb{R}^d} \Delta\varphi(x)\,\psi(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} \varphi(x)\,\Delta\psi(x)\,\mathrm{d}x$ and then deduce the analogous identity for $\varphi, \psi \in H^2(\mathbb{R}^d)$ by density.

[26] We are taking a slightly different convention than in the previous years' lecture notes [**?**], and hence this conclusion is automatic

Namely, (**??**) follows from (**??**) with $B = A - \bar{\lambda}\mathbf{1}$, since then $B^* = A^* - \lambda\mathbf{1}$ and $\operatorname{Ran} B = L^2(\mathbb{R}^d)$ by the assumption that $\bar{\lambda} \in \rho(A)$.

We now prove (**??**). By Definition **??**, we deduce that $\psi \in \operatorname{Ker}(B^*)$ is equivalent to

$$\langle \psi, B\varphi \rangle = 0 \ \ \forall \varphi \in \mathcal{D}(B)\,,$$

which in turn is equivalent to

$$\psi \in (\operatorname{Ran} B)^{\perp}$$

Hence,

$$\operatorname{Ker} B^* = (\operatorname{Ran} B)^{\perp}\,. \tag{5.12}$$

Taking orthogonal complements in (**??**) and recalling Exercise **??**, we obtain

$$(\operatorname{Ker} B^*)^{\perp} = \overline{\operatorname{Ran} B}\,. \tag{5.13}$$

We deduce (**??**) from (**??**). $\square$

We can now state a useful general criterion that allows us to deduce self-adjointness.

Theorem 5.7. *Let A be self-adjoint and let B be closed and symmetric with $\mathcal{D}(A) \subset \mathcal{D}(B)$. Assume that there exists $\varepsilon \in [0, 1)$ and $c \geq 0$ such that for all $\varphi \in \mathcal{D}(A)$, we have*

$$\|B\varphi\| \leq \varepsilon \|A\varphi\| + c\|\varphi\|\,. \tag{5.14}$$

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

Proof. We know that A is closed by Corollary **??**. Therefore $T = A + B$ with $\mathcal{D}(T) = \mathcal{D}(A)$ is closed. We show that for $\lambda \in \mathbb{R}$ with $|\lambda|$ sufficiently large, we have

$$\mathrm{i}\lambda \in \rho(T)\,. \tag{5.15}$$

The claim then follows from Proposition **??**. We note that for $\lambda \in \mathbb{R} \setminus \{0\}$

$$T - \mathrm{i}\lambda\mathbf{1} = A + B - \mathrm{i}\lambda\mathbf{1} = \big(\mathbf{1} + B(A - \mathrm{i}\lambda\mathbf{1})^{-1}\big)\,\big(A - \mathrm{i}\lambda\mathbf{1}\big)\,, \tag{5.16}$$

since $\mathrm{i}\lambda \in \rho(A)$ (recall (**??**)). Again using $\mathrm{i}\lambda \in \rho(A)$, the claim follows from (**??**) if we show that $\mathbf{1} + B(A - \mathrm{i}\lambda\mathbf{1})^{-1}$ has a bounded inverse. We do this by considering a Neumann series, i.e. recalling Exercise **??** (iii), it suffices to show that

$$\|B(A - \mathrm{i}\lambda\mathbf{1})^{-1}\| < 1\,. \tag{5.17}$$

In order to show (**??**), we first note that since $\mathrm{i}\lambda \in \rho(A)$, we have that

$$(A - \mathrm{i}\lambda\mathbf{1})^{-1} : L^2(\mathbb{R}^d) \to \mathcal{D}(A)$$

is a bijection. We can therefore apply (**??**) followed by Lemma **??** to deduce that for all $\varphi \in L^2(\mathbb{R}^d)$, we have

$$\big\|B(A - \mathrm{i}\lambda\mathbf{1})^{-1}\varphi\big\| \leq \varepsilon\big\|A(A - \mathrm{i}\lambda\mathbf{1})^{-1}\varphi\big\| + c\big\|(A - \mathrm{i}\lambda\mathbf{1})^{-1}\varphi\big\| \leq \varepsilon\|\varphi\| + \frac{c}{|\lambda|}\,\|\varphi\|\,. \tag{5.18}$$

We choose λ such that $\varepsilon + \frac{c}{|\lambda|} < 1$. Then we obtain (**??**) from (**??**). Note that here we used the fact that $(A - \mathrm{i}\lambda\mathbf{1})^{-1} : L^2(\mathbb{R}^d) \to \mathcal{D}(A)$ is a bijection. Therefore, we deduce (**??**). $\square$

From Theorem **??**, we deduce the following corollary.

Corollary 5.8. *Let A be self-adjoint. Let B be bounded and symmetric. Then $A+B$ is self-adjoint.*

Proof. We apply Theorem **??** with $\varepsilon = 0$. $\square$

For further applications of Theorem **??**, see Exercise **??** below.

5.2. **The Harmonic oscillator.** In this section, we consider the ***Harmonic oscillator*** when $d = 1$

$$H = \frac{1}{2}(-\Delta + x^2)\,. \tag{5.19}$$

It can be shown that there exists a domain $\mathcal{D}(H) \subset L^2(\mathbb{R})$ which is dense for which H is self-adjoint. See [**?**, Theorem 2.12] (and the reference given there) for a general statement. Alternatively, see [**?**, Section 8.3]. We will accept this fact without proof here.

Before proceeding with the analysis, let us introduce a few more notions from spectral theory. We recall Definition **??**, where we defined the spectrum $\sigma(A)$ and point spectrum $\sigma_p(A)$ of a densely-defined closed operator A. We now define a few more related objects.

Definition 5.9. *Let A be a densely-defined closed operator.*

- (i) *The **continuous spectrum** of A is denoted by $\sigma_c(A)$. It consists of all $\lambda \in \mathbb{C}$ with the property that $A - \lambda\mathbf{1}$ is injective,* $\mathrm{Ran}(A - \lambda\mathbf{1})$ *is dense and* $(A - \lambda\mathbf{1})^{-1}$ *is not bounded.*
- (ii) *The **residual spectrum** of A is denoted by $\sigma_r(A)$. It consists of all $\lambda \in \mathbb{C}$ with the property that $A - \lambda\mathbf{1}$ is injective and* $\mathrm{Ran}(A - \lambda\mathbf{1})$ *is not dense.*

In the sequel, we will not study the residual spectrum in light of the following result.

Proposition 5.10. *Let A be a self-adjoint operator. We then have $\sigma_r(A) = \emptyset$.*

Proof. The claim follows from (**??**) with $B = A - \lambda\mathbf{1}$, since by assumption $\mathrm{Ker}B = \{0\}$ and $B^* = B$ (recall that we are considering $\lambda \in \sigma(A) \subset \mathbb{R}$). $\square$

We recall Proposition **??** and we consider a slight modification of the assumptions.

Definition 5.11 (Spreading sequence)**.** *Let A be a self-adjoint operator and let $z \in \mathbb{C}$ be given. We say that a sequence (ψ_n) in $\mathcal{D}(A)$ is a **spreading sequence** for (A, z) if the following conditions hold.*

- (i) $\|\psi_n\| = 1$ *for all $n \in \mathbb{N}$.*
- (ii) $\|(A - z\mathbf{1})\psi_n\| \to 0$.
- (iii) *For all $R > 0$, there exists $N \in \mathbb{N}$ such that* $\operatorname{ess\,supp} \psi_n \cap B_R(0) = \emptyset$ *for $n > N$.*

We recall that conditions (i) and (ii) above say that (ψ_n) is a Weyl sequence associated with z. The notion from Definition **??** lets us characterise elements of the continuous spectrum of a self-adjoint operator.

Proposition 5.12. *Let A be a self-adjoint operator. If $z \in \sigma_c(A)$, then there exists a spreading sequence for (A, z).*

Sketch of proof. We prove a slightly weaker statement, in which (iii) in Definition **??** is replaced by a weaker condition.

(iii') For all $g \in L^2(\mathbb{R}^d)$, we have $\langle \psi_n, g \rangle \to 0$, i.e. $\psi_n \rightharpoonup 0$.

It is shown in [**?**, Theorem 6.15] that the sequence (ψ_n) which we construct satisfies (iii). This is quite an involved argument which we will not study in our module. The result in [**?**, Theorem 6.15] is referred to as the *Weyl-Zhislin theorem.*

Let $z \in \sigma_c(A)$. By Definition **??**, we know that $(A - z\mathbf{1})^{-1}$ is unbounded. Therefore, there exists a sequence (φ_n) in $\operatorname{Ran}(A - z\mathbf{1})$ such that

$$\|\varphi_n\| = 1\,, \quad \|(A - z\mathbf{1})^{-1}\varphi_n\| \to \infty\,. \tag{5.20}$$

We define

$$\psi_n := \frac{(A - z\mathbf{1})^{-1}\varphi_n}{\|(A - z\mathbf{1})^{-1}\varphi_n\|}\,. \tag{5.21}$$

By assumption, $(A - z\mathbf{1})^{-1}$ is densely-defined. For all h belonging to the dense set $\operatorname{Ran}(A - z\mathbf{1})$, we compute, using (**??**)–(**??**), the self-adjointness of A, and the Cauchy-Schwarz inequality that

$$|\langle \psi_n, h \rangle| = \frac{1}{\|(A - z\mathbf{1})^{-1}\varphi_n\|}\,|\langle \varphi_n, (A - z\mathbf{1})^{-1}h \rangle| \leq \frac{1}{\|(A - z\mathbf{1})^{-1}\varphi_n\|}\,\|(A - z\mathbf{1})^{-1}h\| \to 0\,. \tag{5.22}$$

By a density argument, we get that (**??**) holds for all $h \in L^2(\mathbb{R}^d)$. $\square$

Proposition 5.13. *Let $V : \mathbb{R}^d \to \mathbb{R}$ be continuous such that $V(x) \to \infty$ as $|x| \to \infty$. Then H is self-adjoint with suitable choice of domain $\mathcal{D}(H) \subset L^2(\mathbb{R}^d)$ (see* [**?**, Theorem 2.12] *and the reference therein for the precise statement) and the spectrum consists of eigenvalues $\lambda_1 < \lambda_2 < \cdots$ such that $\lambda_n \to \infty$ as $n \to \infty$.*

Sketch of proof. We only prove that $\sigma_c(H) = \emptyset$. For the proof of the remaining claims, see [**?**, Theorem 6.18]. Suppose that there exists $z \in \sigma_c(H)$. Recalling (**??**), we know that $z \in \mathbb{R}$. By Proposition **??**, there exists an associated spreading sequence. Therefore, we have by (i) and (ii) of Definition **??** and the Cauchy-Schwarz inequality that

$$\langle \psi_n, (H - z\mathbf{1})\psi_n \rangle \to 0\,. \tag{5.23}$$

We compute

$$\begin{aligned}\langle \psi_n, (H - z\mathbf{1})\psi_n \rangle &= \langle \psi_n, -\Delta\psi_n \rangle + \langle \psi_n, V\psi_n \rangle - z \\ &= \int_{\mathbb{R}^d} |\nabla\psi_n|^2 \,\mathrm{d}x + \int_{\mathbb{R}^d} V|\psi_n|^2 \,\mathrm{d}x - z \geq \int_{\mathbb{R}^d} V|\psi_n|^2 \,\mathrm{d}x - z \to \infty\,. \end{aligned} \tag{5.24}$$

For the last step of (**??**), we used conditions (i) and (iii) of Definition **??** and the assumption that $V(x) \to \infty$ as $|x| \to \infty$. We hence obtain a contradiction with (**??**) and therefore $\sigma_c(H) = \emptyset$. $\square$

We now continue with the analysis of the harmonic oscillator (**??**). For the remainder of the subsection, we set $d = 1$. By Proposition **??**, we know that the spectrum of H consists of isolated eigenvalues. We can rewrite H using the ***creation*** and ***annihilation operators*** which are respectively defined as

$$a^* := \frac{1}{\sqrt{2}}(X - \mathrm{i}P)\,, \qquad a = \frac{1}{\sqrt{2}}(X + \mathrm{i}P)\,. \tag{5.25}$$

These are operators that are defined on $\mathcal{D}(X) \cap \mathcal{D}(P)$, which we take to be their domain. Let us note that a^* is indeed the adjoint of a. Furthermore, we define the operator

$$\mathcal{N} := a^* a\,. \tag{5.26}$$

Recall that by Exercise **??** (iii) we have $[X, P] = \mathrm{i}\mathbf{1}$, so we can rewrite (**??**) as[27]

$$H = \mathcal{N} + \frac{1}{2}\,. \tag{5.27}$$

From (**??**), we see that $\mathcal{N}$ is self-adjoint if we take its domain to be the same as that of H. Since $\mathcal{N} \geq 0$, we obtain from (**??**) that $H \geq \frac{1}{2}$. We note the following general result about the harmonic oscillator.

Theorem 5.14. *The spectrum of the self-adjoint operator* $H = \frac{1}{2}(-\Delta + x^2)$ *is*

$$\sigma(H) = \sigma_p(H) = \left\{\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots\right\}.$$

All eigenvalues of H *have multiplicity* 1.

Proof. By (**??**), it is equivalent to prove that $\sigma(\mathcal{N}) = \sigma_p(\mathcal{N}) = \mathbb{N}_0$ and that each eigenvalue of $\mathcal{N}$ has multiplicity 1. By Proposition **??**, the spectrum of $\mathcal{N}$ consists of isolated eigenvalues. Since $\mathcal{N} \geq 0$, all the eigenvalues are nonnegative.

Let us first show that $\lambda = 0$ is an eigenvalue of $\mathcal{N}$. We note that $af = 0$ implies that $\mathcal{N}f = a^* af = 0$. The converse is also true since

$$\langle f, \mathcal{N}f\rangle = \langle f, a^* af\rangle = \langle af, af\rangle = \|af\|^2\,, \tag{5.28}$$

and since $\mathcal{N}$ is positive. Let

$$f_0(x) := \pi^{-1/4}\, \mathrm{e}^{-\frac{1}{2}x^2}\,. \tag{5.29}$$

Then $\|f_0\| = 1$. Furthermore, we compute

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

We deduce that f_0 is indeed an eigenvector of $\mathcal{N}$. Conversely, suppose that $\mathcal{N}f_0 = 0$ with $\|f_0\| = 1$. Recalling (**??**), it follows that $af_0 = 0$. Now, by uniqueness theory of ODEs, we get that f_0 is the unique solution of the ODE $g' = -xg$ with $g(0) = \pi^{-1/4}$. Hence, the eigenvalue 0 has multiplicity 1.

[27] Throughout the sequel, we write $\frac{1}{2}$ instead of $\frac{1}{2}\mathbf{1}$ etc. for simplicity of notation.

Given $n \in \mathbb{N}$, we let

$$f_n := c_n (a^*)^n f_0 \,, \tag{5.30}$$

for c_n to be determined below. By Exercise **??** (i), we note that $f_n \in \mathcal{S}(\mathbb{R})$. In Exercise **??** (ii), we prove the identity

$$\mathcal{N} a^* = a^* (\mathcal{N} + 1) \,. \tag{5.31}$$

Using (**??**) iteratively in (**??**), we obtain that

$$\mathcal{N} f_n = c_n (a^*)^n (\mathcal{N} + n) f_0 \,. \tag{5.32}$$

Recalling that $\mathcal{N} f_0 = 0$ and substituting this into (**??**), it follows that

$$\mathcal{N} f_n = n c_n (a^*) f_0 = n f_n \,.$$

Here, we also recalled (**??**). We note that the function in (**??**) satisfies $\|f_n\| = 1$ if we take $c_n = \frac{1}{\sqrt{n!}}$; see Exercise **??**. In particular, $f_n \neq 0$. We therefore obtain that $\sigma_p(\mathcal{N})$ contains all the nonnegative integers.

Conversely, suppose that $\lambda \in \mathbb{C}$ with $\lambda \neq 0$ is an eigenvalue of $\mathcal{N}$. Let f be an associated eigenvector. Using the identity $\mathcal{N} a = a(\mathcal{N} - 1)$ (see Exercise **??** (ii)), it follows that

$$\mathcal{N} a f = a\,(\mathcal{N} - 1) f = (\lambda - 1) a f \,. \tag{5.33}$$

Moreover,

$$\|af\|^2 = \langle af, af \rangle = \langle f, a^* a f \rangle = \langle f, \mathcal{N} f \rangle = \lambda \|f\|^2 \neq 0 \,. \tag{5.34}$$

From (**??**)–(**??**), it follows that af is an eigenvector of $\mathcal{N}$ with eigenvalue $\lambda - 1$. If λ is not a nonnegative integer, we can iterate this construction to see that all $\lambda-1, \lambda-2, \lambda-3, \ldots$ are all eigenvalues[28]. This is a contradiction since $\mathcal{N}$ is a positive operator, so its eigenvalues are all nonnegative.

Finally, we show that all eigenvalues have multiplicity 1. In order to do this, we argue by contradiction. Suppose that f_n and g_n are orthogonal eigenvectors with eigenvalue n. Then af_n and ag_n are orthogonal eigenvectors with eigenvalue $n-1$. In order to see this, we use (**??**) to deduce that

$$\mathcal{N} a f_n = (n-1) a f_n \,,$$

and similarly for g_n. Furthermore, similarly as in (**??**)

$$\langle a f_n, a g_n \rangle = \langle f_n, a^* a g_n \rangle = \langle f_n, \mathcal{N} g_n \rangle = n \langle f_n, g_n \rangle = 0 \,.$$

By (**??**), we furthermore have that $af_n, ag_n \neq 0$. Iterating this procedure, we obtain that $a^n f_n, a^n g_n$ are orthogonal eigenvectors with eigenvalue 0. However, from before we know that 0 has multiplicity 1, so we obtain a contradiction. $\square$

Remark 5.15. *It can be shown that one can write* (**??**) *in terms of* ***Hermite polynomials***

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

[28]It is important that all of these numbers are non-zero, so we can repeat the above argument.

One then has $f_n(x) = \mathrm{e}^{-\frac{x^2}{2}} H_n(x)$*. We will not use this further in our module. For details, see* [**?**, Section 8.3].

5.3. **The Schrödinger operator with the Coulomb potential (not examinable).** In classical mechanics, atoms and molecules are unstable. In particular, as electrons orbit the nuclei, they radiate away energy and fall onto the nucleus. One of the greatest achievements of the theory of quantum mechanics was to show that this is not the case. Mathematically, stability is interpreted as saying that Hamiltonian H of the system is bounded from below.

Let us fix $d = 3$. When studying the hydrogen atom, one is led to consider the Schrödinger operator.

$$H = -\Delta - \frac{1}{|x|}\,. \tag{5.35}$$

In (**??**), the interaction potential is $V(x) = -\frac{1}{|x|}$, which is referred to as the ***Coulomb potential***.

Proposition 5.16. *Let H be given by* (**??**)*. Then the following properties hold.*

(i) *If we take $\mathcal{D}(H) = H^2(\mathbb{R}^3)$, then H is self-adjoint.*

(ii) *We have*

$$-\Delta \geq \frac{1}{4|x|^2}\,. \tag{5.36}$$

We interpret (**??**) *as saying that for all φ belonging to a dense set of $L^2(\mathbb{R}^3)$ (e.g. $C_c^\infty(\mathbb{R}^3)$ or $\mathcal{S}(\mathbb{R}^3)$), we have*

$$\langle \varphi, -\Delta\varphi\rangle = \int_{\mathbb{R}^3} |\nabla\varphi(x)|^2\,\mathrm{d}x \geq \frac{1}{4}\int_{\mathbb{R}^3} \frac{|\varphi(x)|^2}{|x|^2}\,\mathrm{d}x\,. \tag{5.37}$$

(By a density argument, we hence obtain (**??**) *for all $\varphi \in H^1(\mathbb{R}^3)$.)*

Remark 5.17. *Before proceeding to the proof of Proposition* **??***, we make several comments.*

(1) *We recall that for an operator A, we write $A \geq 0$ if $\langle\psi, A\psi\rangle \geq 0$ for all $\psi \in \mathcal{D}(A)$. More generally, given operators A, B, we say that $A \geq B$ if $A - B \geq 0$. This is what we mean in (ii).*

(2) *The estimate* (**??**) *is sometimes referred to as* ***Hardy's inequality****. In this context, it is usually written as*

$$\left\|\frac{\varphi}{|x|}\right\|_{L^2(\mathbb{R}^3)} \leq 2\|\nabla\varphi\|_{L^2(\mathbb{R}^3)}\,. \tag{5.38}$$

(3) *The result in* (**??**) *generalises to $d \geq 3$*

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

The proof of Proposition **??** *(i) requires $d \leq 3$, so we state the complete result for $d = 3$.*

Proof of Proposition **??**. We note that claim (i) follows from the result of Exercise **??** (i). Note that here, we need $d \leq 3$ since then $\frac{1}{|x|}$ is square integrable near the origin (recall (**??**)). We now prove (ii). For $1 \leq j \leq 3$, we take $P_j = -\mathrm{i}\frac{\partial}{\partial x_j}$ and compute for $\varphi \in C_c^\infty(\mathbb{R}^3)$,

$$\Big[P_j, \frac{1}{|x|}\Big]\,\varphi = -\mathrm{i}\frac{\partial}{\partial x_j}\Big(\frac{\varphi}{|x|}\Big) + \frac{\mathrm{i}}{|x|}\frac{\partial \varphi}{\partial x_j} = \mathrm{i}\,\frac{x_j}{|x|^3}\,\varphi\,.$$

In other words, we have

$$\Big[P_j, \frac{1}{|x|}\Big] = \mathrm{i}\,\frac{x_j}{|x|^3}\,. \tag{5.39}$$

Using the observation that for $1 \leq j \leq 3$, X_j and $\frac{1}{|x|}$ commute and recalling that by arguing analogously as in Exercise **??** (iii), we have $[X_j, P_j] = \mathrm{i}$, it follows that

$$\frac{\mathrm{i}}{3}\sum_{j=1}^{3}\Big[\frac{1}{|x|}\,P_j\,\frac{1}{|x|}, X_j\Big] = \frac{\mathrm{i}}{3}\sum_{j=1}^{3}\frac{1}{|x|}\,\underbrace{\Big(P_j\,X_j - X_j\,P_j\Big)}_{=-\mathrm{i}}\,\frac{1}{|x|} = \frac{1}{|x|^2}\,. \tag{5.40}$$

Using (**??**) and the symmetry of $X_j, P_j, \frac{1}{|x|}$, it follows that for $\varphi \in C_c^\infty(\mathbb{R}^d)$

$$\begin{aligned}\Big\langle \varphi, \frac{1}{|x|^2}\,\varphi\Big\rangle &= \frac{\mathrm{i}}{3}\sum_{j=1}^{3}\Big\langle \varphi, \Big[\frac{1}{|x|}\,P_j\,\frac{1}{|x|}, X_j\Big]\,\varphi\Big\rangle \\ &= \frac{1}{3}\sum_{j=1}^{3}\Big\{\Big\langle \frac{1}{|x|}\,P_j\,\frac{1}{|x|}\,\varphi, \mathrm{i}X_j\varphi\Big\rangle + \Big\langle \mathrm{i}X_j\varphi, \frac{1}{|x|}\,P_j\,\frac{1}{|x|}\,\varphi\Big\rangle\Big\} = -\frac{2}{3}\sum_{j=1}^{3}\mathrm{Im}\,\Big\langle \frac{1}{|x|}\,P_j\,\frac{1}{|x|}\,\varphi, X_j\varphi\Big\rangle\,.\end{aligned} \tag{5.41}$$

By using (**??**) and the symmetry of $\frac{1}{|x|}$, it follows that for $1 \leq j \leq 3$, we have

$$\Big\langle \frac{1}{|x|}\,P_j\,\frac{1}{|x|}\,\varphi, X_j\varphi\Big\rangle = \Big\langle P_j\varphi, \frac{x_j}{|x|^2}\,\varphi\Big\rangle - \mathrm{i}\Big\langle \varphi, \frac{x_j^2}{|x|^4}\varphi\Big\rangle\,. \tag{5.42}$$

We take imaginary parts in (**??**), recall (**??**), and sum in j to deduce that

$$\Big\langle \varphi, \frac{1}{|x|^2}\varphi\Big\rangle = -2\,\mathrm{Im}\,\sum_{j=1}^{3}\Big\langle P_j\varphi, \frac{x_j}{|x|^2}\,\varphi\Big\rangle\,. \tag{5.43}$$

We use the Cauchy-Schwarz inequality (in $L^2(\mathbb{R}^3)$ in (**??**) to deduce that

$$\int_{\mathbb{R}^3}\frac{|\varphi(x)|^2}{|x|^2}\,\mathrm{d}x = \Big|\Big\langle \varphi, \frac{1}{|x|^2}\varphi\Big\rangle\Big| \leq 2\,\sum_{j=1}^{3}\|P_j\varphi\|_{L^2(\mathbb{R}^3)}\,\Big\|\frac{x_j}{|x|^2}\,\varphi\Big\|_{L^2(\mathbb{R}^3)}\,. \tag{5.44}$$

We now use the Cauchy-Schwarz inequality in j in (**??**). The claim then follows from (**??**) by using the Cauchy-Schwarz inequality in $\mathbb{R}^3$. $\square$

From Proposition **??**, we obtain the following result.

Corollary 5.18. *Let H be defined as in (**??**) above. For $\varphi \in \mathcal{D}(H)$ with $\|\varphi\|_{L^2(\mathbb{R}^3)} = 1$, we have $\langle \varphi, H\varphi \rangle \geq -1$, i.e. $H \geq -1$.*

Proof. Let us note that for all $x \in \mathbb{R}^3 \setminus \{0\}$, we have

$$\frac{1}{4|x|^2} - \frac{1}{|x|} = \left(\frac{1}{2|x|} - 1\right)^2 - 1 \geq -1\,.$$

For φ as above, we therefore have by Proposition **??** that

$$\langle \varphi, H\varphi \rangle \geq \left\langle \varphi, \left(\frac{1}{4|x|^2} - \frac{1}{|x|}\right)\varphi \right\rangle \geq -\|\varphi\|^2_{L^2(\mathbb{R}^3)} = -1\,.$$

$\square$

5.4. **Exercises for Section ??.**

Exercise 5.1. *Let $s > 0$ be given. Show that $\mathcal{S}(\mathbb{R}^d)$ is dense in $H^s(\mathbb{R}^d)$, with respect to the norm given by (**??**) above.*
HINT: Work in the Fourier variables.

Exercise 5.2 (Proof of Lemma **??**)**.** *We prove the result of Lemma **??**.*

(i) *Let $\varphi \in \mathcal{S}(\mathbb{R}^d)$ be given. Use Proposition 3.3 to write*

$$\varphi(x) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \widehat{\varphi}(p)\, \mathrm{e}^{\mathrm{i}p\cdot x}\, \mathrm{d}p\,. \tag{5.45}$$

Use this to show the result for such φ.
*HINT: Multiply and divide the integrand in (**??**) by $(1+|p|^2)^{s/2}$ and use the Cauchy-Schwarz inequality.*

(ii) *Deduce the general result. Here, it is useful to recall the result of Exercise **??**.*

Exercise 5.3. *Let $d \leq 3$ and let $V : \mathbb{R}^d \to \mathbb{R}$ be a measurable function satisfying*

$$\int_{|V(x)|>m} V^2(x)\, \mathrm{d}x < \infty \tag{5.46}$$

*for some value of $m \geq 0$. (In (**??**), we are integrating over $x \in \mathbb{R}^d$ with $|V(x)| > m$).*

(i) *Show that for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$, we have*

$$\|V\varphi\|^2 \leq m^2\|\varphi\|^2 + \|\varphi\|^2_{L^\infty} \int_{|V(x)|>m} V^2(x)\, \mathrm{d}x\,. \tag{5.47}$$

(ii) *Explain why (**??**) holds and gives us a finite bound for all $\varphi \in H^2(\mathbb{R}^d)$.*
*HINT: Use Exercise **??**. This is where the assumption $d \leq 3$ is important.*

(iii) *Use Theorem* **??** *to deduce that the Schrödinger operator* $H = -\Delta - \frac{1}{|x|}$ *with* $\mathcal{D}(H) = H^2(\mathbb{R}^3)$ *is self-adjoint.*
HINT: For part (iii), it is helpful to notice that the operator given with multiplication by $\frac{1}{|x|}$ *is closed when we take its domain to be*

$$\left\{\psi \in L^2(\mathbb{R}^3)\,, \quad \frac{1}{|x|}\,\psi \in L^2(\mathbb{R}^3)\right\}. \tag{5.48}$$

Moreover, by (**??**), *it follows that* $H^2(\mathbb{R}^3)$ *is contained in the set given by* (**??**). *Therefore, one can justify the application of Theorem* **??**.

Exercise 5.4. *Verify that condition (iii) of Definition* **??** *implies condition (iii') from the sketch of the proof of Proposition* **??**.

Exercise 5.5. *We analyse the operator* $\mathcal{N}$ *from* (**??**) *in more detail.*

(i) *Check that the operators* a^*, a *defined in* (**??**) *map* $\mathcal{S}(\mathbb{R})$ *to itself.*
(ii) *Use (i) to deduce that the operator* $\mathcal{N}$ *given by* (**??**) *is densely-defined.*
(iii) *Verify the identity* (**??**).
HINT for (iii): Recall that Exercise **??** *(iii) tells us that* $[X, P] = \mathrm{i}\mathbf{1}$.

Exercise 5.6 (Commutation identities for a^*, a)**.** *We show commutation identities that hold for* a^* *and* a.

(i) *Show that* $[a, a^*] = \mathbf{1}$. *More precisely, for any* f, g *belonging to an appropriate dense subset of* $L^2(\mathbb{R})$ *(e.g.* $\mathcal{S}(\mathbb{R})$*), we have*

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

(ii) *Show that the following identities hold.*

$$\mathcal{N}a = a\,(\mathcal{N} - 1)\,, \qquad \mathcal{N}a^* = a^*(\mathcal{N} + 1)\,.$$

(iii) *Given* $f \in \mathcal{S}(\mathbb{R})$ *and* $n \in \mathbb{N}$ *show that for* f_0 *as in* (**??**), *we have*

$$\|(a^*)^n f_0\|^2 = n\|(a^*)^{n-1} f_0\|^2\,.$$

HINT: Write

$$\|(a^*)^n f_0\|^2 = \langle (a^*)^{n-1} f_0, a(a^*)^n f_0\rangle \tag{5.49}$$

and show that the right-hand side of (**??**) *equals*

$$\langle (a^*)^{n-1} f_0, (\mathcal{N} + 1)(a^*)^{n-1} f_0\rangle\,.$$

(iv) *Check that* f_n *given by* (**??**) *satisfies* $\|f_n\| = 1$ *if* $c_n = \frac{1}{\sqrt{n!}}$.

6. The Feynman-Kac formula

6.1. **The Trotter product formula.** Given a self-adjoint operator H on $L^2(\mathbb{R}^d)$, we saw in Theorem **??** how to construct the operator $\mathrm{e}^{-\mathrm{i}tH}$. An analogous result can be shown for the operator e^{-tH} provided that H is bounded below and that we consider $t \geq 0$. This can be deduced from the general ***Hille-Yosida theorem***, which we will not prove, but which can be found in [**?**, Theorem X.47a].

Theorem 6.1. *Assume that H is a self-adjoint operator on $L^2(\mathbb{R}^d)$ and that*

$$E_0 := \inf_{\substack{f \in \mathcal{D}(H) \\ \|f\|=1}} \langle f, Hf \rangle > -\infty \,. \tag{6.1}$$

For every $t \geq 0$, there exists a bounded operator e^{-tH} that satisfies the following properties.

(i) ***Strong continuity:*** *For all $\varphi \in L^2(\mathbb{R}^d)$ and for all $t \geq 0$, we have*
$$\lim_{s \to t} \big\| \mathrm{e}^{-sH} \varphi - \mathrm{e}^{-tH} \varphi \big\| = 0 \,.$$
(ii) ***Boundedness:*** *For all $t \geq 0$, we have*
$$\big\| \mathrm{e}^{-tH} \big\| \leq \mathrm{e}^{-tE_0} \,,$$
where E_0 is given as in (**??**).

(iii) ***Semigroup property:*** *For all $s, t \geq 0$, we have*
$$\mathrm{e}^{-sH} \, \mathrm{e}^{-tH} = \mathrm{e}^{-(s+t)H} \,.$$
(iv) ***Derivative property:*** *On $\mathcal{D}(H)$, we have*
$$\frac{\mathrm{d}}{\mathrm{d}t} \, \mathrm{e}^{-tH} = -H \, \mathrm{e}^{-tH} = -\mathrm{e}^{-tH} \, H \,.$$

We say that an operator H satisfying (**??**) is ***bounded from below***.

Theorem 6.2 (Trotter product formula)**.** *Let A and B be self-adjoint operators on $L^2(\mathbb{R}^d)$ satisfying the following properties.*

(i) *A and B are bounded from below.*

(ii) *$A + B$ is self adjoint with domain $\mathcal{D}(A) \cap \mathcal{D}(B)$.*

Then for all $\varphi \in L^2(\mathbb{R}^d)$, we have

$$\mathrm{e}^{-(A+B)} \, \varphi = \lim_{n \to \infty} \Big(\mathrm{e}^{-\frac{1}{n}A} \, \mathrm{e}^{-\frac{1}{n}B} \Big)^n \, \varphi \,. \tag{6.2}$$

Proof. We note that by assumption (ii), $\mathcal{D}(A+B) = \mathcal{D}(A) \cap \mathcal{D}(B)$ is dense in $L^2(\mathbb{R}^d)$. Furthermore, by assumption (i), we know that A, B, and $A+B$ are all bounded below. Therefore, using Theorem **??** (ii), it suffices to show (**??**) for $\varphi \in \mathcal{D}(A) \cap \mathcal{D}(B) =: \mathcal{D}$. This is what we show. For $\tau \in [0,1]$, we define the operator

$$K_\tau := \frac{1}{\tau} \big(\mathrm{e}^{-\tau A} \, \mathrm{e}^{-\tau B} - \mathrm{e}^{-\tau(A+B)} \big) \,. \tag{6.3}$$

Let us analyse the operator (**??**). For $\varphi \in \mathcal{D}$, we have by Theorem **??** that the following limits hold as $\tau \to 0$.

$$\frac{1}{\tau}\big(\mathrm{e}^{-\tau A}\,\mathrm{e}^{-\tau B} - \mathbf{1}\big)\,\varphi = \frac{1}{\tau}\big(\mathrm{e}^{-\tau A} - \mathbf{1}\big)\,\varphi + \frac{1}{\tau}\,\mathrm{e}^{-\tau A}\,\big(\mathrm{e}^{-\tau B} - \mathbf{1}\big) \to -A\varphi - B\varphi \tag{6.4}$$

$$\frac{1}{\tau}\big(\mathrm{e}^{-\tau(A+B)} - \mathbf{1}\big)\varphi \to -(A+B)\,\varphi\,. \tag{6.5}$$

From (**??**)–(**??**), it follows that $K_\tau\,\varphi \to 0$ as $\tau \to 0$ for every $\varphi \in \mathcal{D}$. From (**??**) and Theorem **??** (i), we know that $\tau \mapsto \|K_\tau\,\varphi\|$ is a continuous function. Therefore, for each $\varphi \in \mathcal{D}$, $\|K_\tau\varphi\|$ is bounded uniformly in $\tau \in [0,1]$ by some quantity depending on φ. Since $A+B$ is self-adjoint with domain $\mathcal{D}$, we have that $\mathcal{D}$ is a Banach space under the norm given by

$$\|\varphi\|_{\mathcal{D}} := \|(A+B)\varphi\| + \|\varphi\|\,. \tag{6.6}$$

See Exercise **??**. By the earlier discussion and the uniform boundedness principle, it follows that there exists $C > 0$ such that for all $\tau \in [0,1]$ and $\varphi \in \mathcal{D}$, we have

$$\|K_\tau\,\varphi\| \leq C\|\varphi\|_{\mathcal{D}}\,. \tag{6.7}$$

Furthermore, we fix $\varphi \in \mathcal{D}$. Given $\tau \geq 0$ and K_τ as in (**??**), we define $F_\tau : [0,1] \to \mathbb{R}$ by

$$F_\tau(s) := \|K_\tau\,\mathrm{e}^{-s(A+B)}\,\varphi\| = \frac{1}{\tau}\big\|\big(\mathrm{e}^{-\tau A}\,\mathrm{e}^{-\tau B} - \mathrm{e}^{-\tau(A+B)}\big)\,\mathrm{e}^{-s(A+B)}\,\varphi\big\|\,. \tag{6.8}$$

Note that $\mathrm{e}^{-s(A+B)}\,\varphi \in \mathcal{D}$ for $s \in [0,1]$. We know that for all $s \in [0,1]$

$$\lim_{\tau\to 0} F_\tau(s) = 0\,. \tag{6.9}$$

We now show that (**??**) holds *uniformly in* $s \in [0,1]$ (i.e. given $\varepsilon > 0$, we can choose τ sufficiently small independently of s such that $F_\tau(s) < \varepsilon$).

Before proceeding with the proof of uniformity in s in (**??**), we use (**??**), the triangle inequality, and (**??**) to deduce that for $r, s \in [0,1]$ we have

$$\begin{aligned}|F_\tau(s) - F_\tau(r)| &\leq \frac{1}{\tau}\big\|\big(\mathrm{e}^{-\tau A}\,\mathrm{e}^{-\tau B} - \mathrm{e}^{-\tau(A+B)}\big)\,\big(\mathrm{e}^{-s(A+B)} - \mathrm{e}^{-r(A+B)}\big)\,\varphi\big\| \\ &\leq C\big\|\big(\mathrm{e}^{-s(A+B)} - \mathrm{e}^{-r(A+B)}\big)\,\varphi\big\|_{\mathcal{D}}\,.\end{aligned} \tag{6.10}$$

The expression on the right-hand side of (**??**) no longer involves τ! Furthermore, by Theorem **??** (i), we deduce that

$$\big\|\big(\mathrm{e}^{-s(A+B)} - \mathrm{e}^{-r(A+B)}\big)\,\varphi\big\|_{\mathcal{D}} \to 0 \tag{6.11}$$

as $|r-s| \to 0$. Note that, in order to deduce (**??**), we are using Theorem **??** (i) with argument φ and $(A+B)\varphi$.

Let us now explain how this implies that (**??**) holds uniformly in $s \in [0,1]$. Let $\varepsilon > 0$ be given. For all $s \in [0,1]$, there exists by (**??**) a quantity τ_s such that $F_\tau(s) < \frac{\varepsilon}{4}$ for all $\tau \in (0,\tau_s)$. By (**??**)–(**??**), it follows that there exists $\delta > 0$, depending on ε such that $F_\tau(r) < \frac{\varepsilon}{2}$ for all $\tau \in (0,\tau_s)$ and for all $r \in (s-\delta, s+\delta) \cap [0,1]$. By compactness of $[0,1]$, there exists finitely many $s_1, \ldots, s_n$ such that $(s_i - \delta, s_i + \delta), i = 1, \ldots, n$ cover $[0,1]$. Let us take

$$\tilde{\tau} := \min_{1 \leq i \leq n} \tau_{s_i} > 0\,.$$

The above arguments then show that $F_\tau(s) < \varepsilon$ for all $s \in [0,1]$ and for all $\tau \in (0, \tilde{\tau})$. Therefore, we deduce that (**??**) holds uniformly in $s \in [0,1]$.

We use telescoping to write (see Exercise **??**)

$$\Big(\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B}\Big)^n - \mathrm{e}^{-(A+B)} = \sum_{k=0}^{n-1} \Big(\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B}\Big)^k \Big(\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B} - \mathrm{e}^{-\frac{1}{n}(A+B)}\Big)\Big(\mathrm{e}^{-\frac{1}{n}(A+B)}\Big)^{n-1-k}\,. \tag{6.12}$$

We use Exercise **??** (i) in (**??**), combined with Theorem **??** (ii) to deduce there exists $\tilde{C} > 0$ such that for all $0 \leq k \leq n-1$, we have

$$\Big\|\Big(\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B}\Big)^k\Big\| \leq \big\|\mathrm{e}^{-\frac{1}{n}A}\big\|^k\,\big\|\mathrm{e}^{-\frac{1}{n}B}\big\|^k \leq \tilde{C}\,. \tag{6.13}$$

In (**??**), we crucially used that $\frac{k}{n} \leq 1$ and Theorem **??** (ii). Using (**??**)–(**??**), Exercise **??** (i) again, and recalling (**??**), we deduce that

$$\begin{aligned}
&\Big\|\Big[\Big(\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B}\Big)^n - \mathrm{e}^{-(A+B)}\Big]\,\varphi\Big\| \\
&\qquad \leq \tilde{C} \sup_{s \in [0,1]} \Big\|\Big(\frac{1}{n}\Big)^{-1}\Big(\mathrm{e}^{-\frac{1}{n}A}\,\mathrm{e}^{-\frac{1}{n}B} - \mathrm{e}^{-\frac{1}{n}(A+B)}\Big)\,\mathrm{e}^{-s(A+B)}\,\varphi\Big\| \\
&\qquad = \tilde{C} \sup_{s \in [0,1]} F_{1/n}(s)\,.
\end{aligned} \tag{6.14}$$

The claim now follows from (**??**) and the fact that (**??**) holds uniformly in $s \in [0,1]$. $\square$

6.2. **The Feynman-Kac formula.** Given a measurable function $a : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$, we formally define the operator A by

$$(A\varphi)(x) := \int_{\mathbb{R}^d} a(x,y)\,\varphi(y)\,\mathrm{d}y\,. \tag{6.15}$$

We refer to operators of the form (**??**) as ***integral operators***. With this notation, we refer to the function a as the associated ***integral kernel***. We note the following general result, whose proof is given in Exercise **??** below.

Proposition 6.3 (Schur's test)**.** *Suppose that $a : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ is measurable and suppose that it satisfies.*

$$M := \sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} |a(x,y)|\,\mathrm{d}y < \infty\,, \qquad N := \sup_{y \in \mathbb{R}^d} \int_{\mathbb{R}^d} |a(x,y)|\,\mathrm{d}x < \infty\,. \tag{6.16}$$

*Then the operator A given by (**??**) is bounded on $L^2(\mathbb{R}^d)$ and it satisfies*

$$\|A\| \leq \sqrt{MN}\,. \tag{6.17}$$

In this section, we analyse a specific integral operator whose integral kernel is given by $a(x,y) = g_t(x-y)$, where $g_t : \mathbb{R}^d \to \mathbb{R}$ is the Gaussian

$$g_t(x) := \frac{1}{(2\pi t)^{d/2}} \, \mathrm{e}^{-\frac{|x|^2}{2t}} \,, \quad t > 0 \,. \tag{6.18}$$

We refer to (**??**) as the ***heat kernel***. The name is justified by the following result.

Proposition 6.4. *For $t > 0$, $\mathrm{e}^{t\Delta/2}$ (given by Theorem **??**) is an integral operator with integral kernel $a(x,y) = g_t(x-y)$. More precisely, for all $\varphi \in L^2(\mathbb{R}^d)$, we have*

$$\big(\mathrm{e}^{t\Delta/2}\,\varphi\big)(x) = \int_{\mathbb{R}^d} g_t(x-y)\,\varphi(y)\,\mathrm{d}y \,. \tag{6.19}$$

Proof. We note that g_t given by (**??**) is a positive function that satisfies

$$\|g_t\|_{L^1} = 1 \,. \tag{6.20}$$

By Proposition **??**, it follows that the right-hand side of (**??**) indeed defines a bounded operator on $L^2(\mathbb{R}^d)$, whose operator norm is at most 1. Since by Theorem **??**, we know that the left-hand side of (**??**) is a bounded operator on $L^2(\mathbb{R}^d)$, it suffices to verify that (**??**) holds for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$.

Given $\varphi \in \mathcal{S}(\mathbb{R}^d)$, let $u_t := g_t * \varphi$ denote the right-hand side of (**??**). Since $\mathcal{S}(\mathbb{R}^d)$ is closed under convolution (recall Exercise **??** (i)), it follows that $u_t \in \mathcal{S}(\mathbb{R}^d)$. Using Exercise **??** (ii) and Lemma 3.4, it follows that

$$\widehat{u}_t(p) = \mathrm{e}^{-\frac{t|p|^2}{2}}\,\widehat{\varphi}(p) \,. \tag{6.21}$$

In light of (**??**), we take $u_0 := \varphi$. (One can then check that u_s is smooth in $s \in [0,\infty)$). We now take $\frac{\partial}{\partial t}$ in (**??**) to get

$$\frac{\partial}{\partial t}\,\widehat{u}_t(p) = -\frac{|p|^2}{2}\,\widehat{u}_t(p) \,. \tag{6.22}$$

We use Remark 3.8 and Lemma 3.2 (i) to rewrite (**??**) as

$$\left(\frac{\partial}{\partial t}u_t - \frac{\Delta}{2}u_t\right)^{\widehat{}}(p) = 0 \tag{6.23}$$

for all $p \in \mathbb{R}^d$. By Fourier inversion (**??**) implies that u_t solves the heat equation with initial data φ. The same is true for $\mathrm{e}^{t\Delta/2}\,\varphi$ so the claim follows by uniqueness of solutions to the heat equation (see Exercise **??**). □

We now introduce the concept of ***Wiener measure***. Given $x \in \mathbb{R}^d$, let $\mathcal{C}_x$ denote the set of all continuous paths $\omega : [0,\infty) \to \mathbb{R}^d$ with $\omega(0) = x$. Let $\Sigma \equiv \Sigma_x$ be the smallest σ-algebra on $\mathcal{C}_x$ that contains all sets of the form[29]

$$\big\{\omega \in \mathcal{C}_x \,;\, \omega(t_1) \in I_1, \ldots, \omega(t_n) \in I_n\big\} \,, \tag{6.24}$$

where $n \in \mathbb{N}, 0 < t_1 < \cdots < t_n$ and $I_1, \ldots, I_n \subset \mathbb{R}^d$ are open.

[29]Sets of the form (**??**) are sometimes referred to as *cylinder sets*.

Theorem 6.5. *Given $x \in \mathbb{R}^d$, there exists a unique probability measure μ_x on $(\mathcal{C}_x, \Sigma)$ such that for all $n \in \mathbb{N}$, $f : \mathbb{R}^{nd} \to \mathbb{R}$ continuous, and $0 < t_1 < \cdots < t_n$, we have*

$$\int_{\mathcal{C}_x} f(\omega(t_1), \ldots, \omega(t_n))\, \mu_x(\mathrm{d}\omega) = \int_{\mathbb{R}^{nd}} g_{t_1}(x - x_1)\, g_{t_2 - t_1}(x_1 - x_2) \cdots g_{t_n - t_{n-1}}(x_{n-1} - x_n)\, f(x_1, \ldots, x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n\,. \tag{6.25}$$

The rigorous construction of the Wiener measure is based on the Riesz-Markov theorem. We will not go through the proof in our module. For the details, one can refer to [**?**, Section X.11: Pages 277–278]. In the probability literature, the paths ω are referred to as ***Brownian paths***.

One can show that the Wiener measure concentrates on paths of appropriate Hölder regularity. Let us recall the relevant definition.

Definition 6.6 (Hölder continuity)**.** *Let $\alpha \in [0,1]$ be given. For $t > 0$, we say that $\omega : [0,t] \to \mathbb{R}^d$ is **Hölder continuous with parameter** α (or α-**Hölder continuous**) if there exists a constant K such that*

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

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

Let us denote by $H_x^\alpha([0,t])$ the set of all continuous paths $\omega : [0,t] \to \mathbb{R}^d$ with $\omega(0) = x$ that are Hölder continuous with parameter α. We identify $H_x^\alpha([0,t])$ with a subset of $\mathcal{C}_x$ by extending $\omega(t') = \omega(t)$ for $t' \geq t$.

We will not prove the following result.

Theorem 6.7. *Let $x \in \mathbb{R}^d$ and $t > 0$ be fixed. The following claims hold.*

(i) *For $\alpha < \frac{1}{2}$, we have that*

$$\mu_x(H_x^\alpha([0,t])) = 1\,.$$

(ii) *For $\alpha \geq \frac{1}{2}$, we have that*

$$\mu_x(H_x^\alpha([0,t])) = 0\,.$$

For a proof of this result in a more general context, see [**?**, Section 1.2].

We note that for $V : \mathbb{R}^d \to \mathbb{R}$ continuous, and $\omega \in \mathcal{C}_x$, we have

$$\int_0^t V(\omega(s))\, \mathrm{d}s = \lim_{n \to \infty} \frac{t}{n} \sum_{j=1}^n V\Big(\omega\Big(\frac{jt}{n}\Big)\Big)\,. \tag{6.26}$$

We can view (**??**) as the definition of the integral of V along a Brownian path.

We now state the first version of the Feynman-Kac formula.

Theorem 6.8 (The Feynman-Kac formula, Version 1)**.** *Consider $H = -\frac{\Delta}{2} + V$ where V satisfies the following properties.*

(i) *V is continuous and bounded below.*
(ii) *H is self-adjoint with domain*[30] $\mathcal{D}(\Delta) \cap \mathcal{D}(V) = H^2(\mathbb{R}^d) \cap \mathcal{D}(V)$.

Then for all $\varphi \in L^2(\mathbb{R}^d)$, $x \in \mathbb{R}^d$, and $t > 0$, we have

$$(\mathrm{e}^{-tH}\,\varphi)(x) = \int_{\mathcal{C}_x} \exp\biggl\{-\int_0^t V(\omega(s))\,\mathrm{d}s\biggr\}\,\varphi(\omega(t))\,\mu_x(\mathrm{d}\omega)\,. \tag{6.27}$$

Proof. Without loss of generality, we can assume that $V \geq 0$ (otherwise, we just add a constant to V). Let us first note that it suffices to show (**??**) for $\varphi \in \mathcal{S}(\mathbb{R}^d)$ (in doing so, we also rigorously interpret the right-hand side of (**??**)). By Theorem **??**, we know that the left-hand side of (**??**) depends continuously on $\varphi \in L^2$. Suppose that $f \in \mathcal{S}(\mathbb{R}^d)$ is given. By Theorem **??**, Young's inequality, and (**??**), we have that

$$\biggl\|\int_{\mathcal{C}_x} |f(\omega(t))|\,\mu_x(\mathrm{d}\omega)\biggr\|_{L^2_x} = \biggl\|\int_{\mathbb{R}^d} g_t(x-x_1)\,|f(x_1)|\,\mathrm{d}x_1\biggr\|_{L^2_x} \leq \|f\|_{L^2}\,. \tag{6.28}$$

Now, we find a sequence (φ_n) in $\mathcal{S}(\mathbb{R}^d)$ such that $\lim_n \|\varphi_n - \varphi\|_{L^2} = 0$. Note that (**??**) follows if we show that for all $n \in \mathbb{N}$ we have

$$(\mathrm{e}^{-tH}\,\varphi_n)(x) = \int_{\mathcal{C}_x} \exp\biggl\{-\int_0^t V(\omega(s))\,\mathrm{d}s\biggr\}\,\varphi_n(\omega(t))\,\mu_x(\mathrm{d}\omega)\,. \tag{6.29}$$

Namely, we take limits as $n \to \infty$ and note that the right-hand side of (**??**) is Cauchy in $L^2(\mathbb{R}^d)$ by (**??**) (where we take $f = \varphi_m - \varphi_n$).

We henceforth consider $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Given $\omega \in \mathcal{C}_x$, we define for $n \in \mathbb{N}$ the function $F_n : \mathcal{C}_x \to \mathbb{R}$ by

$$F_n(\omega) := \exp\biggl\{-\frac{t}{n}\sum_{j=1}^n V\Bigl(\omega\Bigl(\frac{jt}{n}\Bigr)\Bigr)\biggr\} \leq 1\,. \tag{6.30}$$

With F_n as in (**??**), we use (**??**) followed by the dominated convergence theorem[31] to write

$$\begin{aligned}\int_{\mathcal{C}_x} \mathrm{e}^{-\int_0^t V(\omega(s))\,\mathrm{d}s}\,\varphi(\omega(t))\,\mu_x(\mathrm{d}\omega) &= \int_{\mathcal{C}_x} \lim_{n\to\infty} F_n(\omega)\,\varphi(\omega(t))\,\mu_x(\mathrm{d}\omega)\\ &= \lim_{n\to\infty}\int_{\mathcal{C}_x} F_n(\omega)\,\varphi(\omega(t))\,\mu_x(\mathrm{d}\omega)\,,\end{aligned} \tag{6.31}$$

which is justified provided that the limit on the right-hand side of (**??**) exists. We now show that this is indeed the case.

We now use (**??**) and recall (**??**) to rewrite the right-hand side of (**??**) as

$$\lim_{n\to\infty}\int_{\mathbb{R}^{nd}} g_{\frac{t}{n}}(x-x_1)\,\cdots\,g_{\frac{t}{n}}(x_{n-1}-x_n)\,\mathrm{e}^{-\frac{t}{n}\sum_{j=1}^n V(x_j)}\,\varphi(x_n)\,\mathrm{d}x_1\,\cdots\,\mathrm{d}x_n\,. \tag{6.32}$$

[30]Recall Proposition **??**.
[31]whose application is justified by (**??**).

Note that, in the application of (**??**), we took $t_j = \frac{jt}{n}, j = 1, \ldots, n$ and

$$f(x_1, \ldots, x_n) = \mathrm{e}^{-\frac{t}{n}\sum_{j=1}^{n} V(x_j)}\,\varphi(x_n)\,.$$

By Proposition **??**, we have

$$(\textbf{??}) = \lim_{n\to\infty}\Big(\underbrace{\mathrm{e}^{\frac{t}{2n}\Delta}\,\mathrm{e}^{-\frac{t}{n}V}\,\cdots\,\mathrm{e}^{\frac{t}{2n}\Delta}\,\mathrm{e}^{-\frac{t}{n}V}}_{n \text{ times}}\,\varphi\Big)(x)\,. \tag{6.33}$$

By the Trotter product formula (Theorem **??**), we have that

$$(\textbf{??}) = \big(\mathrm{e}^{-t(-\Delta/2+V)}\,\varphi\big)(x) = \big(\mathrm{e}^{-tH}\,\varphi\big)(x)\,. \tag{6.34}$$

Note that the application of the Trotter product formula was justified in (**??**) by the assumptions of the theorem. □

Remark 6.9. *Let us make several comments on Theorem* **??**

(i) *We note that in the assumptions of Theorem* **??**, *we can take real-valued* $V \in L^\infty(\mathbb{R}^d)$ *which is continuous.*
(ii) *When* $d = 3$, *assumption (ii) of Theorem* **??** *holds for real-valued*
$$V \in L^2(\mathbb{R}^d) + L^\infty(\mathbb{R}^d) \equiv \big\{f_1 + f_2\,;\, f_1 \in L^2(\mathbb{R}^d)\,,\, f_2 \in L^\infty(\mathbb{R}^d)\big\}\,.$$
See [**?**, Theorem X.15] *for details.*
(iii) *In* [**?**, Theorem X.68], *it is shown that, when* $d = 3$, (**??**) *holds when we take real-valued* $V \in L^2(\mathbb{R}^d) + L^\infty(\mathbb{R}^d)$, *as in (ii). In particular, there is no assumption about* V *being bounded below. A possible example for* V *is the Coulomb potential* $V(x) = -\frac{1}{|x|}$.
(iv) *Note that, with* $d = 3$, *by (iii), we can consider* $V \in L^p(\mathbb{R}^d)$ *for* $2 \leq p \leq \infty$.

Using the same ideas as before, we can write the operator $\mathrm{e}^{-t(-\Delta/2+V)} \equiv \mathrm{e}^{-tH}$ as an integral operator with a kernel $(\mathrm{e}^{-tH})(x,y)$ that can be expressed using an appropriate modification of the Wiener measure. In order to do this, we need to consider Brownian paths where *both endpoints are fixed.* More precisely, for $t > 0$ and $x, y \in \mathbb{R}^d$, we let $\mathcal{C}^{(t)}_{x,y}$ denote the set of continuous paths $\omega : [0,t] \to \mathbb{R}^d$ such that $\omega(0) = x, \omega(t) = y$. Such objects are sometimes referred to as ***Brownian bridges***.

Let $\tilde{\Sigma} \equiv \tilde{\Sigma}^{(t)}_{x,y}$ be the smallest σ-algebra on $\mathcal{C}^{(t)}_{x,y}$ that contains all sets of the form

$$\big\{\omega \in \mathcal{C}^{(t)}_{x,y}\,;\, \omega(t_1) \in I_1, \ldots, \omega(t_n) \in I_n\big\}\,,$$

where $n \in \mathbb{N}, 0 < t_1 < \cdots < t_n < t$ and $I_1, \ldots, I_n \subset \mathbb{R}^d$ are open. It is then possible to use the Riesz-Markov theorem and prove a result analogous to Theorem **??**, which we state below.

Theorem 6.10. *Given* $x, y \in \mathbb{R}^d$ *and* $t > 0$, *there exists a unique measure* $\mu^{(t)}_{x,y}$ *on* $(\mathcal{C}^{(t)}_{x,y}, \tilde{\Sigma})$ *such that for all* $n \in \mathbb{N}$, $f : \mathbb{R}^{nd} \to \mathbb{R}$ *continuous, and* $0 < t_1 < \cdots < t_n$, *we* *have*

$$\int_{\mathcal{C}^{(t)}_{x,y}} f(\omega(t_1), \ldots, \omega(t_n))\, \mu^{(t)}_{x,y}(\mathrm{d}\omega) = \int_{\mathbb{R}^{nd}} g_{t_1}(x - x_1)\, g_{t_2 - t_1}(x_1 - x_2) \cdots g_{t - t_n}(x_n - y)\, f(x_1, \ldots, x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n\,. \tag{6.35}$$

Note that $\mu^{(t)}_{x,y}$ is ***not a probability measure***. One can verify that

$$\mu^{(t)}_{x,y}\big(\mathcal{C}^{(t)}_{x,y}\big) = g_t(x - y) \neq 1\,. \tag{6.36}$$

See Exercise **??**.

Theorem 6.11 (The Feynman-Kac formula, Version 2)**.** *Suppose that V satisfies the same assumptions as Theorem **??**. Then for $t > 0$, $\mathrm{e}^{-t(-\Delta/2 + V)} \equiv \mathrm{e}^{-tH}$ is an integral operator with kernel*

$$k(x,y) = \int_{\mathcal{C}^{(t)}_{x,y}} \exp\bigg\{- \int_0^t V(\omega(s))\, \mathrm{d}s\bigg\}\, \mu^{(t)}_{x,y}(\mathrm{d}\omega)\,. \tag{6.37}$$

Proof. It suffices to show that for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$, we have

$$\big(\mathrm{e}^{-tH}\varphi\big)(x) = \int_{\mathbb{R}^d} k(x,y)\, \varphi(y)\, \mathrm{d}y\,. \tag{6.38}$$

Using (**??**), the dominated convergence theorem, and Theorem **??**, it follows that for $\varphi \in \mathcal{S}(\mathbb{R}^d)$ (which in particular is continuous), we have

$$\int_{\mathbb{R}^d} k(x,y)\, \varphi(y)\, \mathrm{d}y = \lim_{n \to \infty} \int_{\mathbb{R}^d} \int_{\mathbb{R}^{(n-1)d}} g_{\frac{t}{n}}(x - x_1) \cdots g_{\frac{t}{n}}(x_{n-1} - y)\, \mathrm{e}^{-\frac{t}{n}\sum_{j=1}^{n-1} V(x_j) - \frac{t}{n} V(y)}\, \varphi(y)\, \mathrm{d}x_1 \cdots \mathrm{d}x_{n-1}\, \mathrm{d}y\,. \tag{6.39}$$

By the arguments in (**??**)–(**??**) (here we replace x_n by y), we deduce that

$$(\mathbf{??}) = \big(\mathrm{e}^{-tH}\varphi\big)(x)\,,$$

which implies the claim. ∎

From Theorem **??**, we deduce the following pointwise nonnegativity result.

Corollary 6.12. *Suppose that V satisfies the same assumptions as Theorem **??**. Then for $t > 0$, the integral kernel $\mathrm{e}^{-t(-\Delta/2 + V)} \equiv \mathrm{e}^{-tH}$ is pointwise nonnegative.*

Proof. Recalling (**??**), and arguing analogously as for (**??**), we have that

$$k(x,y) = \lim_{n} \int_{\mathbb{R}^{(n-1)d}} g_{\frac{t}{n}}(x - x_1) \cdots g_{\frac{t}{n}}(x_{n-1} - y)\, \mathrm{e}^{-\frac{t}{n}\sum_{j=1}^{n-1} V(x_j) - \frac{t}{n} V(y)}\, \mathrm{d}x_1 \cdots \mathrm{d}x_{n-1}\,. \tag{6.40}$$

In particular, the claim then follows from (**??**) and the positivity of the heat kernel (**??**). □

For an application of Corollary **??**, see Exercise **??** below.

6.3. **Exercises for Section 6.**

Exercise 6.1. *In the proof of Theorem* **??***, show that* $\mathcal{D}$ *is a Banach space under the norm* (**??**).
HINT: Recall that $A+B$ *is closed. It is helpful to review the argument in the proof of the (ii)* $\Rightarrow$ *(iii) claim in Theorem* **??**.

Exercise 6.2. *Verify* (**??**).
HINT: Write

$$\mathrm{e}^{-(A+B)} = \left(\mathrm{e}^{-\frac{1}{n}(A+B)}\right)^n$$

and write $X^n - Y^n = X^n - X^{n-1}Y + X^{n-1}Y - \cdots - Y^n$.

Exercise 6.3 (The Lie product formula)**.** *Let* A, B *be complex* $d \times d$ *matrices for some finite* d. *In this exercise, we show directly that*

$$\mathrm{e}^{A+B} = \lim_{n\to\infty} \left[\mathrm{e}^{\frac{1}{n}A}\,\mathrm{e}^{\frac{1}{n}B}\right]^n . \tag{6.41}$$

We note that the convergence in (**??**) *is in a finite-dimensional space of matrices so we do not have to specify the norm. Furthermore, there are no additional assumptions on* A *and* B*, as opposed to the case of unbounded operators.*

(i) *Given* $n \in \mathbb{N}$, *let* $S_n := \mathrm{e}^{\frac{1}{n}(A+B)}$ *and* $T_n := \mathrm{e}^{\frac{1}{n}A}\,\mathrm{e}^{\frac{1}{n}B}$. *Show that*

$$S_n^n - T_n^n = \sum_{m=0}^{n-1} S_n^m \,(S_n - T_n)\, T_n^{n-1-m} .$$

(ii) *Given a complex* $d \times d$ *matrix* C, *show that*

$$\|\mathrm{e}^C\| \leq \mathrm{e}^{\|C\|} .$$

HINT: Write the exponential as in (2.54) *and use Exercise 2.1 (iii).*

(iii) *Using (i) and (ii), show that*

$$\|S_n^n - T_n^n\| \leq n\|S_n - T_n\|\,\mathrm{e}^{\|A\|+\|B\|} .$$

(iv) *Show that*

$$\|S_n - T_n\| \leq \frac{C}{n^2} ,$$

for some constant $C > 0$ *depending on* $\|A\|$ *and* $\|B\|$.
HINT: Expand S_n *and* T_n *according to* (2.54) *and notice that in* $S_n - T_n$, *some terms cancel.*

(v) *Deduce* (**??**).

Exercise 6.4. *In this exercise, we note an analogue of Theorem* **??** *when* $\mathrm{e}^{-(A+B)}, \mathrm{e}^{-A}, \mathrm{e}^{-B}$ *are replaced by* $\mathrm{e}^{\mathrm{i}(A+B)}, \mathrm{e}^{\mathrm{i}A}, \mathrm{e}^{\mathrm{i}B}$ *respectively. In other words, instead of the time-evolution given by Theorem* **??**, *we are using the time-evolution given by Theorem* **??**. *We now give the precise statement.*

Let A and B be self-adjoint operators on $L^2(\mathbb{R}^d)$ such that $A+B$ is self adjoint with domain $\mathcal{D}(A+B)$. Then for all $\varphi \in L^2(\mathbb{R}^d)$, we have

$$\mathrm{e}^{\mathrm{i}(A+B)}\,\varphi = \lim_{n\to\infty}\Big(\mathrm{e}^{\frac{\mathrm{i}}{n}A}\,\mathrm{e}^{\frac{\mathrm{i}}{n}B}\Big)^n\,\varphi\,. \tag{6.42}$$

In some textbooks, (**??**) *is referred to as a part of the Trotter product formula.*
HINT: Argue analogously as in the proof of Theorem **??**. *Note that in this case, we de not need to assume that A and B are bounded from below.*

Exercise 6.5 (Proof of Proposition **??**). *In this exercise, we outline a proof of Schur's test given in Proposition* **??**.

(i) *With notation as in Proposition* **??**, *explain why it suffices to show that for all $\varphi, \psi \in L^2(\mathbb{R}^d)$, we have*

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |a(x,y)|\,|\varphi(x)|\,|\psi(y)|\,\mathrm{d}x\,\mathrm{d}y \leq \sqrt{MN}\,\|\varphi\|_{L^2(\mathbb{R}^d)}\,\|\psi\|_{L^2(\mathbb{R}^d)}\,. \tag{6.43}$$

(ii) *Write the expression on the left-hand side of* (**??**) *as*

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\Big[|a(x,y)|^{1/2}\,|\varphi(x)|\Big]\,\Big[|a(x,y)|^{1/2}\,|\psi(y)|\Big]\,\mathrm{d}x\,\mathrm{d}y\,. \tag{6.44}$$

Use the Cauchy-Schwarz inequality in $L^2(\mathbb{R}^d \times \mathbb{R}^d)$ in (**??**) *to deduce the claim.*

Exercise 6.6 (Uniqueness of solutions to the heat equation). *Suppose that $\varphi \in \mathcal{S}(\mathbb{R}^d)$ is given and suppose that for $t > 0$, u_t and v_t solve the heat equation with initial data φ, i.e.*

$$\Big(\frac{\partial}{\partial t} - \frac{\Delta}{2}\Big)\,u_t = 0\,, \quad u_0 = \varphi\,,$$
$$\Big(\frac{\partial}{\partial t} - \frac{\Delta}{2}\Big)\,v_t = 0\,, \quad v_0 = \varphi\,.$$

Suppose furthermore that $u_t, v_t \in \mathcal{S}(\mathbb{R}^d)$ for all $t > 0$. Show that $u_t = v_t$ for all $t > 0$.
HINT: Show that

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\|u_t - v_t\|^2_{L^2(\mathbb{R}^d)} \leq 0\,.$$

Here, one can justify all of the calculations rigorously by Remark 3.8.

Exercise 6.7. *Show the identity* (**??**) *and deduce that $\mu^{(t)}_{x,y}$ is not a probability measure.*
HINT: Set $f = 1$ in (**??**).

Exercise 6.8 (The Feynman-Kac formula and the heat equation). *In this exercise, we study solutions of the heat equation using the Feynman-Kac formula. Throughout, we fix $\varphi \in \mathcal{S}(\mathbb{R}^d)$ which is pointwise nonnegative.*

(i) *Suppose that* $u : \mathbb{R}^d \times [0,\infty) \to \mathbb{C}$ *solves the heat equation.*

$$\Big(\frac{\partial}{\partial t} - \frac{\Delta}{2}\Big)u(x,t) = 0\,, \qquad u(x,0) = \varphi(x)\,. \tag{6.45}$$

Show that $u \geq 0$ *pointwise.*

(ii) *More generally, suppose that* $V \in \mathcal{S}(\mathbb{R}^d)$ *is given and that* $v : \mathbb{R}^d \times [0,\infty) \to \mathbb{C}$ *solves the heat equation with external potential* V*.*

$$\Big(\frac{\partial}{\partial t} - \frac{\Delta}{2} + V\Big)v(x,t) = 0\,, \qquad v(x,0) = \varphi(x)\,. \tag{6.46}$$

Show that $v \geq 0$ *pointwise.*

Throughout the exercise, you may assume uniqueness results for (**??**)–(**??**) *without proof. (The assumption that* $V \in \mathcal{S}(\mathbb{R}^d)$ *is added so that we do not have to worry about questions concerning domains).*

Exercise 6.9 (Optional)**.** *We fix* $V \in \mathcal{S}(\mathbb{R}^d)$ *(and hence* $\mathcal{D}(V) = L^2(\mathbb{R}^d)$*). Let* $H = -\Delta/2 + V$*. Show that for all* $\varphi \in \mathcal{S}(\mathbb{R}^d)$*,* $t > 0$*, and* $x \in \mathbb{R}^d$*, we have*

$$\big(\mathrm{e}^{-\mathrm{i}tH}\varphi\big)(x) = \lim_{n\to\infty}\bigg(\frac{4\pi\mathrm{i}t}{n}\bigg)^{-\frac{nd}{2}} \int_{\mathbb{R}^{nd}} \exp\big(\mathrm{i}S_n(x_0,\ldots,,x_n,t)\big)\,, \tag{6.47}$$

where

$$S_n(x_0,x_1,\ldots,x_n,t) := \sum_{j=1}^{n}\frac{t}{n}\left[\frac{1}{4}\bigg(\frac{|x_j - x_{j-1}|}{t/n}\bigg)^2 - V(x_j)\right].$$

HINT: Use the result of Exercise **??***.*
For a physical interpretation of (**??**) *in terms of Feynman's original work, see the discussion in* [**?**, Pages 275–276].

6.4. **Acknowledgements.** Thanks to Daniel Ueltschi for sharing his notes [**?**] and to Peter Mühlbacher for the comments incorporated therein. Credits to Daniel Abraham, Georgios Athanasopolous, Hang Lok Edison Au-Yeung, which Ali Farahat, Benedetta Bertoli, Matt Dopson, Miles Foster-Nyamundanda, Ajan Sittampalam, and Chengyi Sun for their helpful comments on the current notes.

6.5. **A quote for the end.** *'Consider nothing impossible, then treat possibilities as probabilities'.*
–Charles Dickens, David Copperfield.

References

[1] R. B. Feynman, R. P. Leighton, M. Sands, *The Feynman Lectures in Physics, Volume III*, Addison-Wesley Publishing Company, 1965.
[2] G.B. Folland, *Real Analysis*, Wiley Interscience, 1999.
[3] G. Friedlander, M. Joshi, *Introduction to the Theory of Distributions*, Cambridge University Press, 2nd edition, 1998.
[4] S. J. Gustafson, I. M. Segal, *Mathematical Concepts of Quantum Mechanics*, Springer Universitext, 3rd edition, 2020.
[5] A. Messiah, *Quantum Mechanics*, Dover, 1999.
[6] P. Mörters, Y. Peres, *Brownian motion*, Cambridge Studies in Statistical and Probabilistic Mathematics, Cambridge University Press, 2010.
[7] M. Reed, B. Simon, *Methods of Modern Mathematical Physics*, Volume 1: Functional Analysis. Elsevier Academic Press, 1980.
[8] M. Reed, B. Simon, *Methods of Modern Mathematical Physics*, Volume 2: Fourier Analysis, Self-adjointness. Elsevier Academic Press, 1975.
[9] G. Strang, *Introduction to Linear Algebra*, Wellesley-Cambridge Press, Fifth Edition, 2016.
[10] T. Tao, *Hardy's uncertainty principle*, available from https://terrytao.wordpress.com/2009/02/18/hardys-uncertainty-principle, accessed Dec. 29, 2021.
[11] G. Teschl, *Mathematical Methods in Quantum Mechanics, With Applications to Schrödinger Operators*, AMS Graduate Studies in Mathematics, Volume 157, 2014.
[12] D. Ueltschi, *An introduction to Quantum Mechanics for mathematicians*, lecture notes for MA4A7, March 2021.

Index