Lecture Notes Summer 2026 14^{th} July, 2026

Mathematical Quantum Theory

Gandalf Lechner

FAU Erlangen-Nürnberg, Department of Mathematics

gandalf.lechner@fau.de

These lectures give an introduction to quantum mechanics from a mathematical perspective, emphasizing the general framework of observable *-algebras and states. The notes are divided into three chapters: Chapter 1 provides a brief account of classical Hamiltonian mechanics, which however includes a fairly detailed discussion of states given by probability measures on phase space and their correlations. Chapter 2 is devoted to "matrix mechanics", i.e. quantum mechanics on a finite-dimensional Hilbert space. The main structural elements of this theory are presented, including a discussion of entangled states and the Bell inequalities. In Chapter 3, the general Hilbert space framework of quantum mechanics is explored, assuming familiarity with functional analysis and spectral theory. This includes spectral analysis of Hamiltonians and the dynamics they induce, together with physically relevant examples such as the hydrogen atom. Several appendices recall relevant mathematical material.

FAU

Contents

About these lecture notes

These lecture notes are written to accompany the mathematical quantum theory (MQT) lectures at FAU Erlangen in summer 2026. The target audience consists of mathematics and physics students alike, but the focus is on the mathematics students who already know about functional analysis, but not so much about quantum physics. The focus will be on the mathematical and conceptual aspects of the field. We will also comment on the physical implications, but refer to the physics literature for a deeper discussion of them.

The design of these notes contains some self-explanatory coloured elements, such as grey for definitions, green for examples, etc. Another element is:

> ⚛ The boxes marked in dark blue and with an ⚛ are dedicated to questions of interpretation in terms of physics, applications in physics, etc.

At the end of each section, there are exercises and problems. The exercises are meant for the in person tutorials, and the problems are meant to be solved in homework.

These lecture notes are written as a service to the students to allow for a smooth learning process. It is still highly recommended to also consult suitable text books. Some books on quantum mechanics that are specifically written for a mathematical audience are [Hal13; Tak08; Tes09]. A more physics oriented book that is still readable by mathematicians is [GP90].

This document is updated during the term as we go along. In case you spot any inaccuracies or mistakes, or in case something is not clear, please let me know via email or our StudOn forum so that I can improve the notes accordingly.

Gandalf Lechner

Acknowledgements

Many thanks go to Johannes Große for producing many of the exercises in these notes, and all solutions (which are not included). Many thanks go also to the participants of the lecture who have helped to improve these notes by questions, pointing out typos, and generally speaking being an interested audience.

1 Aspects of classical mechanics

Quantum mechanics can only be understood with at least a minimal background in classical mechanics. Such a background will be provided in this chapter in a form suitable for later comparison with quantum mechanics.

Classical mechanics is the oldest and best understood field of theoretical physics. It accurately describes many phenomena observed in the real world, ranging from falling apples over the motion of the celestial bodies to pendulums, spinning tops and space rockets, to name just a few. Classical mechanics can be formulated in several ways (called Newtonian/Lagrangian/Hamiltonian mechanics), all of which are of an analytic nature and involve differential equations. In fact, calculus was strongly influenced and even developed in parallel with classical mechanics by Newton and Leibniz.

Given the great age and success of classical mechanics, it is no surprise that it has been extended to cover also other areas, such as relativistic mechanics, fluid dynamics, etc. But as any physical theory, also classical mechanics has only a finite range of validity. One important limitation is that classical mechanics does *not* correctly describe the microscopic world, i.e. atoms, molecules, light (photons), or other subatomic particles. This is the realm of *quantum* mechanics, and the main subject of this lecture. The word "classical" is used to distinguish classical mechanics from "quantum" mechanics.

We will present classical mechanics mostly in its Hamiltonian formulation, in which a Hamiltonian function defines a dynamics on the classical observable algebra, because this is closest to the quantum mechanical setting, in which a Hamiltonian operator defines a dynamics on the quantum observable algebra. We will also discuss a few instructive examples that will later be compared with corresponding examples of quantum systems. For more detailed literature on classical mechanics, see [Wal07, Chapter 1] or [AM87; Arn89]. This chapter also covers probabilistic/statistical scenarios described by probability measures that are typically not part of classical mechanics, but allow us to introduce various concepts such as expectation values, uncertainties, correlations, and entanglement, already in the classical setting.

1.1 Terminology of classical mechanics and Newton's Equation

Classical mechanics provides the tools for describing the motion of massive bodies under the influence of forces, i.e. their positions, momenta, etc. (*kinematics*), and the change of these quantities with time (*dynamics*). In a mathematical theory, all these physical quantities are modeled by objects from analysis, such as numbers, vectors, functions, vector fields, etc.

As a start, we present a vocabulary for translating some of the physical quantities into mathematical objects.

- **Time.** Time is a one-dimensional continuum, modeled by the line of real numbers $\mathbb{R}$. We will usually denote time by t.
- **Space and Position.** We live in a three-dimensional Euclidean space[1], modeled as $\mathbb{R}^3$. Often, it will also be necessary or convenient to consider spaces of different dimension, such as $\mathbb{R}^1$ (for motion confined to a straight line), $\mathbb{R}^2$ (for motion in a plane), or $\mathbb{R}^n$

[1] When taking into account relativity, one has to change this point of view and consider a four-dimensional Lorentzian manifold instead.

for some large $n \in \mathbb{N}$, for instance $n = N \cdot d$ for describing the positions of N particles in $\mathbb{R}^d$.

Sometimes we will also consider systems in which particles are confined to an (typically open) subset $U \subset \mathbb{R}^d$. The space U of all possible positions is called the *configuration space*.

The physical significance of vectors $x \in U$ is to specify the position of a particle (at a given time, see below). We will write $x = (x_1, \dots, x_d)$ for the *components* of d-dimensional vectors, and $\|x\| = \sqrt{x_1^2 + \dots + x_d^2}$ for their Euclidean norm (length).

- **Particles and curves.** A particle is an idealized object which has *no spatial extent* and hardly any structure. It thus occupies a specific position $x(t) \in U$ at any given time t. The function $t \mapsto x(t)$ is called the *trajectory* of the particle. This is a curve from a (bounded or unbounded) interval $I \subset \mathbb{R}$ into the configuration space $U \subset \mathbb{R}^d$ which will typically be smooth, i.e. an element of $C^\infty(I, U)$.

  In classical mechanics, the only additional datum that a particle has is its *mass*, a positive parameter $m > 0$. This has a dual role: It measures how much the particle resists to forces acting on it (see Newton's Equation below), and it also measures how much it is influenced by gravitational force. This latter aspect will not be considered in these lectures. In quantum mechanics, we will see that particles can also carry another datum, called *spin*, in addition to mass. Spin is not visible in classical mechanics.
- **Velocity, Momentum, and Acceleration.** When a particle moves along some trajectory (according to a law / ODE to be specified), there are several other interesting quantities ("observables", i.e. quantities you could measure by performing an experiment), prominent examples being the *velocity*, *momentum* and *acceleration* of the particle.

  The velocity is the temporal change of the position. In mathematical terms, this is the (first) derivative of $t \mapsto x(t)$, and it is common notation to write it as

$$\dot{x}(t) = \frac{dx(t)}{dt} = \left(\frac{dx_1(t)}{dt}, \dots, \frac{dx_d(t)}{dt} \right). \tag{1.1}$$

  Instead of the velocity, we will work with a closely related quantity, the *momentum* p. It also incorporates the mass m of the particle, and is simply

$$p(t) := m\, \dot{x}(t). \tag{1.2}$$

  Similarly, the acceleration is the temporal change of the velocity, and thus the second derivative of $t \mapsto x(t)$. It is also written as

$$\ddot{x}(t) = \frac{d^2x(t)}{dt^2} = \left(\frac{d^2x_1(t)}{dt^2}, \dots, \frac{d^2x_d(t)}{dt^2} \right). \tag{1.3}$$

  Clearly not *every* trajectory has velocity/acceleration (i.e., is once or twice differentiable), but in the situations we will encounter, these derivatives will always exist and hence also form curves in $\mathbb{R}^d$.

  Even if the position $x(t)$ is confined to a non-trivial configuration space $U \subset \mathbb{R}^d$, the *momentum space*, namely the space of all possible momenta $p(t)$, will always be $\mathbb{R}^d$.
- **Phase space.** In Hamiltonian mechanics, it will be convenient to collect the position and momentum variables in a single vector

$$\xi(t) := \begin{pmatrix} x(t) \\ p(t) \end{pmatrix} = \begin{pmatrix} x(t) \\ m\dot{x}(t) \end{pmatrix} \in U \times \mathbb{R}^d \subset \mathbb{R}^{2d}. \tag{1.4}$$

The space

$$M := U \times \mathbb{R}^d \subset \mathbb{R}^{2d} \tag{1.5}$$

in which these vectors live is called *phase space*. The letter M indicates that this is a manifold. General treatments of classical mechanics work with more general phase space manifolds, but we will always restrict to phase spaces of the form (1.5).

- **Force.** From everyday experience we know that a force has both, a direction (push/pull) and a "strength", both of which can vary with the position at which you experience the force. (Think of a wind that is blowing at different angles and speeds through a city.) We thus model a force as a *vector field*, i.e. a map $F : U \subset \mathbb{R}^d \to \mathbb{R}^d$ that associates to every point x in the configuration space U a vector $F(x) \in \mathbb{R}^d$, indicating the force[2] pulling at the point x in the direction of $F(x)$, with strength $\|F(x)\|$. We will write

$$\mathcal{X}(U) = C^\infty(U, \mathbb{R}^d) \tag{1.6}$$

for the space of all smooth vector fields on an open set $U \subset \mathbb{R}^d$.

These basic concepts allow us to talk about quantities of interest in mechanics in mathematical language and provide the tools to describe their time dependence.

Definition 1.1 (Newton's Equation). Let $U \subset \mathbb{R}^d$ be open, $F \in \mathcal{X}(U)$ a smooth force field and $m > 0$ (mass). *Newton's Equation* is the second order differential equation

$$m\,\ddot{x}(t) = F(x(t)) \tag{1.7}$$

for trajectories $t \mapsto x(t) \in U$. Equivalently, the *phase space formulation of Newton's Equation* is the first order differential equation

$$\dot{\xi}(t) = \mathcal{V}_F(\xi(t)), \qquad \mathcal{V}_F : M \to \mathbb{R}^{2d}, \qquad \mathcal{V}_F \begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} m^{-1}p \\ F(x) \end{pmatrix}, \tag{1.8}$$

for curves $t \mapsto \xi(t) \in M$ in phase space.

In Newton's Equation, the vector field F and the parameter m are given, and one is searching for smooth curves (trajectories) $x : I \to U$ (or $\xi : I \to M$) defined on a time interval I. The role of the force field F is to describe how the given force F pushes the particle, and the role of the mass m is to describe how inert the particle is, i.e. how much it resists to its motion being changed by forces acting on it.

Example 1.2 (Free particle). The simplest possible situation is that in which *no* force acts (also called the "free" particle), i.e. $F : \mathbb{R}^d \to \mathbb{R}^d$, $F(x) = 0$ for all $x \in \mathbb{R}^d$. Then Newton's Equation reads

$$m\ddot{x}(t) = 0 \iff \dot{\xi}(t) = 0. \tag{1.9}$$

[2] More generally, a force can also depend on time (wind picking up / slowing down) and/or on the velocity of the particle it is acting on (friction). For the purposes of this quick review of classical mechanics, we will ignore these more general situations.

By integration, we easily see that it has the unique solution

$$x(t) = x(0) + \frac{t}{m} \cdot p(0) \iff \xi(t) = \begin{pmatrix} 1 & \frac{t}{m}1 \\ 0 & 1 \end{pmatrix} \xi(0), \tag{1.10}$$

where 1 denotes the $(d \times d)$ unit matrix.

As Newton said: In the absence of forces, a particle at rest ($p(0) = 0$) remains at rest ($p(t) = 0$ for all $t \in \mathbb{R}$), and a particle moving at constant velocity remains moving at constant velocity ($\dot{x}(t) = 0$, i.e. moving in a straight line at constant speed).

This simple example shows that we can only expect to find a *unique* solution to Newton's Equation once we have fixed an *initial condition* of the form[3]

$$(x(0) = x_0, \quad m\dot{x}(0) = p_0) \iff \xi(0) = \xi_0 := (x_0, p_0) \in M \tag{1.11}$$

for some $x_0 \in U$, $p_0 \in \mathbb{R}^d$. The initial condition (1.11) fixes the position of the particle at time 0 to be x_0, and the momentum of the particle at time 0 to be p_0. Once these data are fixed, the whole trajectory is uniquely determined.

One therefore considers the ODE $\dot{\xi}(t) = \mathcal{V}_F(\xi(t))$ in combination with an initial condition $\xi(0) = \xi_0$. The general mathematical setting consists of an arbitrary vector field $\mathcal{V}$ defined on an open subset $\mathcal{O} \subset \mathbb{R}^n$. In the example of Newton's Equation, we have $\mathcal{V} = \mathcal{V}_F$ and $\mathcal{O} = M = U \times \mathbb{R}^d \subset \mathbb{R}^{2d}$.

Definition 1.3 (Initial value problem). Let $\mathcal{O} \subset \mathbb{R}^n$ be open, $\mathcal{V} \in \mathfrak{X}(\mathcal{O})$ a smooth vector field, and $\xi_0 \in \mathcal{O}$. The *initial value problem* consists in finding smooth curves $\xi : I \to \mathcal{O}$ defined on an open interval $I \subset \mathbb{R}$ containing $t = 0$, such that

$$\dot{\xi}(t) = \mathcal{V}(\xi(t)), \quad t \in I, \qquad \xi(0) = \xi_0. \tag{1.12}$$

These curves are called (*local*) *solutions*. A local solution $\xi : I \to \mathcal{O}$ is called *maximal* if there exists no local solution extending it to a larger interval $J \supsetneq I$. A *global solution* is a local solution with $I = \mathbb{R}$.

Every local solution can be extended to a maximal solution, which may be global or not depending on the initial condition ξ_0, see Exercise 1.1 for an example. This phenomenon can occur in two ways: either a solution runs into the boundary of $\mathcal{O}$, or it escapes to infinity in finite time.

Let us recall a main theorem from the theory of ordinary differential equations in the general setting of an arbitray smooth vector field on an arbitrary open set in $\mathbb{R}^n$. We formulate it for smooth vector fields, but note that there exist other versions for Lipschitz continuous vector fields or just continuous vector fields (Peano's Theorem, which only gives existence but not uniqueness) [Heu95; Chi99].

Theorem 1.4 (Picard-Lindelöf). *Let $\mathcal{O} \subset \mathbb{R}^n$ be open and $\mathcal{V} \in \mathfrak{X}(\mathcal{O})$ smooth.*

a) Existence of local solutions: *For every $\xi_0 \in \mathcal{O}$, there exists a local solution of the initial value problem given by $\mathcal{V}$ and ξ_0.*

b) Uniqueness of maximal local solutions: *Every local solution can be extended to a maximal local solution. Maximal local solutions are unique.*

[3] Of course, we could also use a different time $t_0 \in \mathbb{R}$ instead of $t_0 = 0$ to formulate the initial condition.

c) Global existence criterion: *If* $\mathcal{O} = \mathbb{R}^n$ *and there exist* $c_1, c_2 > 0$ *such that*

$$\|\mathcal{V}(\xi)\| \leq c_1 + c_2\|\xi\|, \qquad \xi \in \mathbb{R}^n, \tag{1.13}$$

then every local solution extends to a global solution.

d) Smooth dependence on initial data: *Let* I_{ξ_0} *be the interval on which the maximal solution* ξ *with* $\xi(0) = \xi_0$ *is defined. Then*

$$D := \bigcup_{\xi_0 \in \mathcal{O}} I_{\xi_0} \times \{\xi_0\} = \{(t, \xi_0) \in \mathbb{R} \times \mathcal{O} \ : \ t \in I_{\xi_0}\} \subset \mathbb{R} \times \mathcal{O} \tag{1.14}$$

is open, and the flow of $\mathcal{V}$*, i.e. the map*

$$\phi^{\mathcal{V}} : D \to \mathbb{R}^n, \qquad (t, \xi_0) \mapsto \phi^{\mathcal{V}}(t, \xi_0) := \phi^{\mathcal{V}}_t(\xi_0) := \xi(t) \tag{1.15}$$

that maps (t, ξ_0) *to the value* $\xi(t)$ *of the unique solution with initial condition* $\xi(0) = \xi_0$ *at time* t*, is smooth.*

e) Properties of the flow: *The flow* $\phi^{\mathcal{V}}$ *satisfies*

$$\phi^{\mathcal{V}}_0(\xi_0) = \xi_0, \qquad \xi_0 \in \mathcal{O}, \tag{1.16}$$

and for all $\xi_0 \in \mathcal{O}$ *and* $t, s \in \mathbb{R}$ *with* (t, ξ_0), (s, ξ_0), $(t + s, \xi_0) \in D$*,*

$$\phi^{\mathcal{V}}_t(\phi^{\mathcal{V}}_s(\xi_0)) = \phi^{\mathcal{V}}_{t+s}(\xi_0) = \phi^{\mathcal{V}}_s(\phi^{\mathcal{V}}_t(\xi_0)). \tag{1.17}$$

This theorem follows essentially from the Banach Fixed Point Theorem. In Exercise 1.2, you are asked to prove part of it.

- The sufficient condition in c) is satisfied for the vector field $\mathcal{V}_F$ coming from a force $F \in \mathcal{X}(\mathbb{R}^d)$ (1.8) that satisfies a bound of the form $\|F(x)\| \leq c_1' + c_2'\|x\|$. Thus, at least for force fields that are linearly bounded and defined on all of $\mathbb{R}^d$, all solutions are global. In this situation, one calls the flow *global*.
- A global flow defines a one-parameter group of diffeomorphisms $\phi^{\mathcal{V}}_t \in \mathrm{Diff}(\mathcal{O})$, i.e.

$$\phi^{\mathcal{V}} : \mathbb{R} \to \mathrm{Diff}(\mathcal{O}), \qquad t \mapsto \phi^{\mathcal{V}}_t \tag{1.18}$$

  is a group homomorphism. In detail, this means that every $\phi^{\mathcal{V}}_t$ is a diffeomorphism of $\mathcal{O}$ (i.e. smooth, bijective, with smooth inverse), and

$$\phi^{\mathcal{V}}_t \circ \phi^{\mathcal{V}}_s = \phi^{\mathcal{V}}_{t+s}. \tag{1.19}$$

  In particular, $\phi^{\mathcal{V}}_0 = \mathrm{id}_{\mathcal{O}}$ and $(\phi^{\mathcal{V}}_t)^{-1} = \phi^{\mathcal{V}}_{-t}$. All these properties follow directly from part e) of the theorem.
- The flow $\phi^{\mathcal{V}}$ of the vector field $\mathcal{V}$ can be seen as a time evolution: It maps the solution at time $t = 0$, i.e. the initial condition ξ_0, to the solution at time t, i.e. $\xi(t)$. Equation 1.17 then simply says that first evolving by time s and then by time t amounts to evolving by time $t + s$.
- For a vector field $\mathcal{V}_F$ (1.8) coming from a force field F and a point mass (particle) $m > 0$, we will attribute properties of solutions of $\dot{\xi}(t) = \mathcal{V}_F(\xi(t))$ to the particle described by it. For example, we will speak of the position, momentum, energy, etc., of the particle.

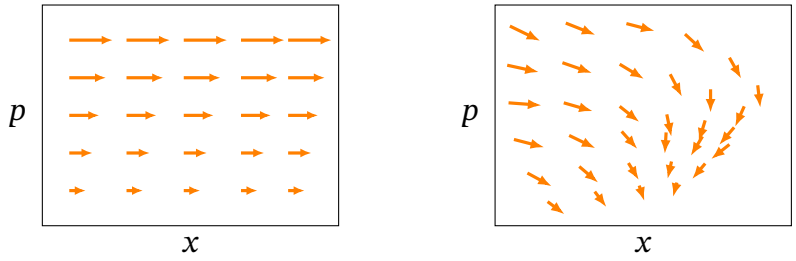


Figure 1: Two vector fields on the phase space $\mathbb{R}^2$. The left one is the linear zero force vector field (1.20), the right one a more interesting non-constant vector field. To solve the ODE $\dot{\xi}(t) = \mathcal{V}(\xi(t))$ means to find the curves tangential to the vector field. An initial condition $\xi(0) = \xi_0$ fixes the starting point of the curve, and the length of the vectors from the vector field indicates the velocity of the particle at that point.

- In Example 1.2, the vector field is linear,

$$\mathcal{V} : \mathbb{R}^{2d} \to \mathbb{R}^{2d}, \qquad \mathcal{V}\begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} m^{-1}p \\ 0 \end{pmatrix}, \tag{1.20}$$

  see Figure 1, and its (global) flow is

$$\phi_t^{\mathcal{V}}(\xi_0) = \begin{pmatrix} 1 & \frac{t}{m}1 \\ 0 & 1 \end{pmatrix} \xi_0. \tag{1.21}$$

  In this case, smoothness of $\phi^{\mathcal{V}}$ and the group property (1.17) are easily verified. For a more interesting example, see Problem 1.2.

Having specified the general setting, we now specialize to so-called conservative force fields, defined as follows.

Definition 1.5 (Conservative forces and potentials). A *conservative vector field* $F \in \mathcal{X}(U)$ is defined to be a gradient field, i.e. F is conservative if and only if there exists[ab] $V \in C^\infty_{\mathbb{R}}(U)$ such that

$$F(x) = -\nabla V(x) = -\left(\frac{\partial V(x)}{\partial x_1}, \dots, \frac{\partial V(x)}{\partial x_n}\right), \qquad x \in U. \tag{1.22}$$

The function V is called a *potential* of F.

[a] Attention with the notation: $\mathcal{V}$ denotes vector fields, i.e. vector-valued maps $\mathcal{O} \to \mathbb{R}^d$, and V denotes potentials, i.e. scalar (number-valued) maps $U \to \mathbb{R}$.

[b] A subscript $\mathbb{R}$ on $C_{\mathbb{R}}(U)$ or $C^\infty_{\mathbb{R}}(U)$ means that these are spaces of *real-valued* functions, whereas $C(U)$, $C^\infty(U)$ are the corresponding spaces of *complex-valued* functions.

Note that in case a vector field $F \in \mathcal{X}(U)$ has a potential V, then also $V + c$, with $c \in \mathbb{R}$ an arbitrary constant, is a potential for F. If U is connected, all potentials of F are of this form: Given two potentials V, W for F, we have $\nabla(V - W) = 0$ which implies $V(x) = W(x) + c$ by connectedness of U.

Writing $\partial_k = \frac{\partial}{\partial x_k}$ for partial derivatives, conservative fields can be characterized as follows.

Proposition 1.6 (Characterization of conservative fields). *Let $U \subset \mathbb{R}^d$ be open and simply connected. A vector field $F \in \mathcal{X}(U)$ is conservative if and only if*

$$\partial_k F_l = \partial_l F_k, \qquad k, l \in \{1, \dots, d\}. \tag{1.23}$$

In particular, in dimension $d = 1$ every vector field is conservative.

Proof. Let $V \in C^\infty_{\mathbb{R}}(U)$ be a potential for F. Then, by Schwarz' Theorem from analysis 2,

$$\partial_k F_l = -\partial_k \partial_l V = -\partial_l \partial_k V = \partial_l F_k. \tag{1.24}$$

Conversely, assume (1.23), pick a point $x_0 \in U$, and for every $x \in U$, a continuously differentiable curve $\gamma_x : [0,1] \to U$ with $\gamma_x(0) = x_0$ and $\gamma_x(1) = x$. The line integral

$$V(x) := \int_{\gamma_x} F \cdot d\vec{s} = \int_0^1 \langle F(\gamma_x(t)), \gamma_x'(t) \rangle \, dt \tag{1.25}$$

does not depend on the explicit choice of γ_x, and is a potential for F (see analysis 2, or analysis 3 with Stokes' Theorem). □

Example 1.7 (The Coulomb Force). This example is the vector field

$$F : \mathbb{R}^d \backslash \{0\} \to \mathbb{R}^d, \qquad F(x) = c \cdot \frac{x}{\|x\|^d}, \tag{1.26}$$

where $c \in \mathbb{R}$ is a constant which can either be positive (repulsive force) or negative (attractive force). It is readily verified by differentiation that

$$V : \mathbb{R}^d \backslash \{0\} \to \mathbb{R}, \qquad V(x) = \begin{cases} \frac{c}{d-2} \frac{1}{\|x\|^{d-2}} & d \neq 2 \\ -c \log \|x\| & d = 2 \end{cases} \tag{1.27}$$

is a potential for (1.26).

⚛ The importance of the force field (1.26) and its potential (1.27) derives from the fact that it describes two physically important situations: On the one hand, the (attractive) *gravitational* force between two massive particles (with $c = Gm_1m_2 > 0$, where m_1, m_2 are the masses of the particles, and G the gravitational constant), and on the other hand, the electrostatic attraction/repulsion between two charged particles (with $c = k_e q_1 q_2$, where q_1, q_2 are the positive or negative charges of the particles, and k_e Coulomb's constant).

The name *Coulomb* potential refers to its role in the context of electrostatic force. It is important in the description of the attraction/repulsion of electrons and protons in atoms, and will also appear in quantum mechanics later in this lecture.

For conservative forces, one can define an energy function (Hamiltonian function).

Definition 1.8 (Hamiltonian function). Let $F \in \mathcal{X}(U)$ be conservative with potential V. Then the *Hamiltonian function* $H \in C^\infty_{\mathbb{R}}(M)$, $M = U \times \mathbb{R}^d$, of F is defined as

$$H : U \times \mathbb{R}^d \to \mathbb{R}, \qquad H(x,p) := \frac{\|p\|^2}{2m} + V(x). \tag{1.28}$$

The first summand of the Hamiltonian, $\frac{1}{2m}\|p\|^2$, is called *kinetic energy* , and the second term, the potential $V(x)$, is called *potential energy* . Kinetic energy is the energy of a particle coming from its movement (it is zero for $p = 0$), and potential energy is the energy of a particle coming from its position in a force field.

A very important feature of conservative force fields is that the total energy, i.e. the Hamiltonian, is conserved in the following sense.

Theorem 1.9 (Conservation of energy for conservative forces). *Let $F \in \mathcal{X}(U)$ be conservative and H its Hamiltonian function (1.28). Let $I \ni t \mapsto x(t)$ be a solution of Newton's Equation. Then*

$$H(x(t), p(t)) = H(x(0), p(0)), \qquad t \in I. \tag{1.29}$$

Proof. We write $\langle x, y\rangle = \sum_{k=1}^d x_k y_k$ for the scalar product in $\mathbb{R}^d$, and calculate

$$\begin{aligned}
\frac{d}{dt}H(x(t),p(t)) &= \frac{d}{dt}\left(\frac{\|p(t)\|^2}{2m} + V(x(t))\right) \\
&= \frac{\langle \dot p(t),\ p(t)\rangle}{m} + \langle \dot x(t),\ \nabla V(x(t))\rangle \\
&= \langle \ddot x(t),\ p(t)\rangle + \frac{1}{m}\langle p(t), -F(x(t))\rangle \\
&= \frac{1}{m}\langle m\ddot x(t) - F(x(t)),\ p(t)\rangle \\
&= 0.
\end{aligned}$$

Hence the differentiable function $I \ni t \mapsto H(x(t),p(t))$ has derivative zero, i.e. it is constant. □

Note that thanks to energy conservation, every solution $I \ni t \mapsto x(t) \in \mathbb{R}^d$ of Newton's Equation has a fixed energy $E := H(x(t),p(t))$ which does not depend on t. This value is also referred to as the energy of the particle described by the solution.

Energy considerations are often useful when the solution can not be obtained explicitly (which is the general case). The following corollary shows how the sublevel sets of the potential,

$$\{V \le V_0\} := \{y \in \mathbb{R}^d \ :\ V(y) \le V_0\}, \qquad V_0 \in \mathbb{R}, \tag{1.30}$$

give information about solutions of Newton's Equation.

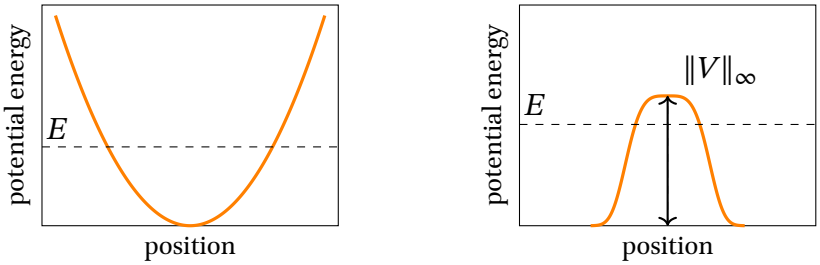


Figure 2: Two potential/energy graphs. The orange curve is the graph of the potential V. Left: all sublevel sets of the potential are bounded. Right: For $E < \|V\|_\infty$, the sublevel set $\{V \leq E\}$ has two connected components.

Corollary 1.10. *Let $U \subset \mathbb{R}^d$ be open, $F \in \mathcal{X}(U)$ conservative with potential V, and $I \ni t \mapsto x(t) \in U$ a solution of Newton's Equation with initial value $(x(0), p(0))$ and energy $E := H(x(0), p(0))$. Then for any $t \in I$, the position $x(t)$ lies in the connected component of $\{V \leq E\}$ that contains $x(0)$.*

Proof. Since the kinetic energy $\|p\|^2/2m$ is non-negative, we have $E \geq V(x(t))$, i.e. $x(t) \in \{V \leq E\}$, for all t. Suppose that $x(t)$ lies in a different connected component C than the component C_0 containing $x(0)$. Then x is a continuous curve connecting C and C_0, so there exists t such that $x(t) \notin \{V \leq E\}$, a contradiction. □

Example 1.11 (Potential/Energy considerations).

a) Let V be a potential that grows to infinity at infinity, $V(x) \to \infty$ for $\|x\| \to \infty$. Then every solution of Newton's Equation is bounded because all sublevel sets $\{V \leq E\}$ are bounded (Figure 2, left).

b) In dimension $d = 1$, let V be a potential "bump", namely a smooth positive function V of compact support (Figure 2, right). Suppose $x(0)$ lies to the left of the support of V. Then $V(x(0)) = 0$, and if $p(0) < \sqrt{2m\|V\|_\infty}$, the solution will remain to the left of the support of V for all times: Its energy is

$$E = \frac{p(0)^2}{2m} + V(x(0)) = \frac{p(0)^2}{2m} < \|V\|_\infty,$$

so the sublevel set $\{V \leq E\}$ has two connected components separated by a neighbourhood of x_* such that $V(x_*) = \|V\|_\infty$. If $p(0) > \sqrt{2m\|V\|_\infty}$, the sublevel set $\{V \leq E\}$ consists of all of $\mathbb{R}$, and Corollary 1.10 gives no restriction on $x(t)$. This can be summarized as: "The particle can travel over the hill only if its initial momentum is large enough".

See Problem 1.1 for related considerations.

These observations are in agreement with the intuition that the solutions $t \mapsto x(t)$ behave like a little ball rolling friction less on the graph of the potential V, i.e. accelerating when going downhill, and slowing down when going uphill. In contrast to that, our later investigations of quantum mechanics will lead to counterintuitve phenomena in quantum dynamical systems.

Another illustration of the usefulness of energy conservation is the following. In dimension $d = 1$, energy conservation implies

$$E = \frac{m}{2}\,\dot{x}(t)^2 + V(x(t)) \implies \frac{dx}{dt} = \pm\sqrt{\frac{2}{m}\,(E - V(x(t)))}\,. \tag{1.31}$$

This equation can be solved by separation of variables, resulting in

$$t = \int_{x(0)}^{x(t)} \frac{dy}{\pm\sqrt{\frac{2}{m}\,(E - V(y))}}\,. \tag{1.32}$$

The problem of solving Newton's Equation (1.7) has thus been simplified to calculating the above integral, and inverting the resulting equation to obtain $x(t)$.

1.2 Hamiltonian formulation

In view of the importance of energy conservation, we reformulate Newton's Equation for conservative systems in terms of the Hamiltonian. It will be helpful to introduce the antisymmetric $(2d) \times (2d)$ matrix (the canonical symplectic form on $\mathbb{R}^{2d}$)

$$\Sigma := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{1.33}$$

where every block is a $(d \times d)$-matrix. This will be fixed for all systems in the following. In a general treatment of Hamiltonian mechanics, the symplectic form is part of the description of the system.

Definition 1.12 (Hamiltonian vector fields, flows, and equations). Let $H \in C^\infty_{\mathbb{R}}(M)$.

a) The *Hamiltonian vector field* of H is the vector field $\mathcal{V}_H \in \mathcal{X}(M)$ defined by

$$\mathcal{V}_H := \Sigma \nabla H \iff \mathcal{V}_H(x,p) = \begin{pmatrix} \nabla_p H \\ -\nabla_x H \end{pmatrix}, \tag{1.34}$$

where ∇_x denotes the gradient of H w.r.t. the x-variables, ∇_p the gradient of H w.r.t. the p-variables, and ∇ the full gradient (w.r.t $\xi = (x,p)$).

b) The *Hamiltonian flow* is the flow defined by the Hamiltonian vector field via Theorem 1.4. It is denoted $\phi^H := \phi^{\mathcal{V}_H}$.

c) *Hamilton's Equation* is the first order ODE defined by the Hamiltonian flow,

$$\dot{\xi}(t) = \mathcal{V}_H(\xi(t)). \tag{1.35}$$

So far we have based our considerations on a force field $F \in \mathcal{X}(U)$ on configuration space U, and derived a Hamiltonian function H from it in case F was conservative. The above definition presents a change of perspective – it is based on a a general smooth real function $H \in C^\infty_{\mathbb{R}}(M)$ on phase space $M = U \times \mathbb{R}^d$. In case H comes from a conservative force $F = -\nabla V$ (1.28), we note that

$$\nabla_p H = \nabla_p \left(\frac{\|p\|^2}{2m} + V(x) \right) = m^{-1}\, p, \tag{1.36}$$

$$-\nabla_x H = -\nabla_x \left(\frac{\|p\|^2}{2m} + V(x) \right) = F(x). \tag{1.37}$$

Comparing with (1.8), we conclude that Newton's Equation is equivalent to Hamilton's Equation (1.35) in this case. Hence our new point of view, taking an arbitrary Hamiltonian function H as a starting point, is a generalization of the Newtonian point of view. We introduce some more terminology associated with this perspective now.

Definition 1.13 (Classical Hamiltonian system and observable algebra).
a) A (*classical*) *Hamiltonian system* is a pair (M, H) consisting of a phase space $M = U \times \mathbb{R}^d$ and a function $H \in C^\infty_{\mathbb{R}}(M)$, called the Hamiltonian function.
b) The (*classical*) *observable algebra* of a Hamiltonian system (M, H) is

$$\mathcal{A}_{\mathrm{cl}} := C^\infty(M). \tag{1.38}$$

Elements of $\mathcal{A}_{\mathrm{cl}}$ are all smooth functions on phase space $U \times \mathbb{R}^d$, for example the components of x (position), p (momentum), $\dot{x}$ (velocity), the potential energy V, the kinetic energy, the Hamiltonian H, etc. In order to avoid confusion between the variables x, p and the functions $(x, p) \mapsto x, (x, p) \mapsto p$, we write, $k \in \{1, \dots, d\}$

$$X_k, P_k : M \to \mathbb{R}, \qquad X_k(x, p) := x_k, \qquad P_l(x, p) := p_k, \tag{1.39}$$

and have $X_k, P_l \in C^\infty(M)$. Note that in (1.38), we consider complex smooth functions (no index $\mathbb{R}$). Only the real functions in $C^\infty(M)$ are actual observables, but we include the complex ones in the interest of a better comparison with quantum mechanics later, which will be based on complex functions. In our current setting of classical ODEs, complex functions do not cause any problems because we can split any complex function into its real and imaginary part.

We now discuss the structure of the classical observable algebra $C^\infty(M)$. Of course, $C^\infty(M)$ is a associative and commutative complex algebra w.r.t. the pointwise operations, and it is unital with unit element the constant function

$$1(\xi) = 1, \qquad \xi \in M. \tag{1.40}$$

It moreover is a *-algebra, namely pointwise complex conjugation,

$$f^*(\xi) := \overline{f(\xi)}, \tag{1.41}$$

is an antilinear involution $C^\infty(M) \to C^\infty(M)$ satisfying $(fg)^* = g^* f^*$. A function $f \in C^\infty(M)$ is real-valued if and only if $f = f^*$.

Another important structural element of $C^\infty(M)$ is its Poisson bracket.

Definition 1.14 (Poisson bracket). The *Poisson bracket* is the map

$$\{\cdot, \cdot\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M), \qquad \{f, g\} := \langle \nabla \bar{f}, \Sigma \nabla g \rangle, \tag{1.42}$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean scalar product on $\mathbb{C}^{2d}$. Explicitly,

$$\{f, g\} = \sum_{k=1}^{d} \left(\frac{\partial f}{\partial x_k} \frac{\partial g}{\partial p_k} - \frac{\partial g}{\partial x_k} \frac{\partial f}{\partial p_k} \right). \tag{1.43}$$

Proposition 1.15 (Properties of the Poisson bracket). *The Poisson bracket has the following properties*[a] *($f, g, h \in C^\infty(M)$):*

a) $\{\cdot,\cdot\}$ is bilinear and antisymmetric,

b) The Leibniz rule holds: $\{f, gh\} = \{f, g\}h + g\{f, h\}$,

c) The Jacobi Identity holds: $\{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} = 0$,

d) The Poisson bracket respects the *-operation: $\{f, g\}^* = \{f^*, g^*\}$.

[a] For those who know Lie algebras: Items a) and c) can be summarized in the statement that $(C^\infty(M), \{\cdot,\cdot\})$ is a (infinite-dimensional) Lie algebra. Besides the Lie algebra structure, $C^\infty(M)$ also has a commutative algebra structure which is compatible with the Lie algebra structure in the sense of b) – one speaks of a *Poisson algebra* in this case.

The proof follows by straightforward calculation (Problem 1.3).

Particularly simple examples of functions in $C^\infty(M)$ are the X_k, P_l (1.39). Their Poisson brackets are called *canonical Poisson brackets*, and given by

$$\{X_k, X_l\} = 0, \qquad \{P_k, P_l\} = 0, \qquad \{X_k, P_l\} = \delta_{k,l} \cdot 1, \tag{1.44}$$

where 1 denotes the constant function (1.40). We will see similar expressions in quantum mechanics later.

We next formulate the Hamiltonian flow on the level of the classical observable algebra and study its interplay with the Poisson bracket. For simplicity, we assume the flow to be global, i.e. all local solutions to extend to global ones.

Theorem 1.16 (Hamiltonian dynamics on the classical observable algebra). *Let (M, H) be a Hamiltonian system with global flow ϕ^H, and define*[a]

$$\alpha_t^H : C^\infty(M) \to C^\infty(M), \qquad \alpha_t^H(f) := f \circ \phi_t^H. \tag{1.45}$$

Then

a) $\alpha^H : \mathbb{R} \to \operatorname{Aut} C^\infty(M)$ is a one parameter group of Poisson *-algebra automorphisms. That is, for every $t \in \mathbb{R}$, the map α_t^H is an algebra homomorphism that moreover satisfies $\alpha_t^H(f)^* = \alpha_t^H(f^*)$,

$$\alpha_t^H(\{f, g\}) = \{\alpha_t^H(f), \alpha_t^H(g)\}, \tag{1.46}$$

and

$$\alpha_t^H \circ \alpha_s^H = \alpha_{t+s}^H, \qquad t, s \in \mathbb{R}. \tag{1.47}$$

b) The Hamiltonian generates the dynamics in the sense

$$\frac{d}{dt}\alpha_t^H(f) = \{\alpha_t^H(f), H\}, \qquad f \in C^\infty(M), \tag{1.48}$$

where the derivative is understood pointwise in $\xi \in M$.

c) *The Hamiltonian is invariant under the dynamics,*

$$\alpha_t^H(H) = H. \tag{1.49}$$

[a] When the Hamiltonian H is clear from the context, we will drop the superscript H and write α_t instead of α_t^H.

Proof. a) α_t^H is a *-algebra homomorphism by definition, and the group law (1.47) follows from Theorem 1.4 e). The proof of the preservation of the Poisson bracket (1.46) is left as an exercise (Problem 1.3).

b) Let $\xi_0 \in M$ and denote by $t \mapsto \xi(t)$ the unique solution of Hamilton's Equation (1.35) with initial value $\xi(0) = \xi_0$. Then

$$\begin{aligned}
\frac{d}{dt}\alpha_t^H(f)(\xi_0) &= \frac{d}{dt} f(\xi(t)) && \text{(definition of } \alpha^H \text{ and } \phi^H) \\
&= \langle \dot{\xi}(t), \nabla f(\xi(t)) \rangle && \text{(chain rule)} \\
&= \langle \mathcal{V}_H(\xi(t)), \nabla f(\xi(t)) \rangle && \text{(Hamilton's Equation)} \\
&= \langle \Sigma \nabla H(\xi(t)), \nabla f(\xi(t)) \rangle && \text{(definition of } \mathcal{V}_H) \\
&= \langle \nabla \bar{f}(\xi(t)), \Sigma \nabla H(\xi(t)) \rangle && \text{(symmetry of scalar product)} \\
&= \{f(\xi(t)), H(\xi(t))\} && \text{(definition of Poisson bracket)} \\
&= \alpha_t^H(\{f, H\})(\xi_0) && \text{(definition of } \alpha_t^H) \\
&= \{\alpha_t^H(f), \alpha_t^H(H)\}(\xi_0). && (1.46)
\end{aligned}$$

Setting $f = H$ in the last equation, we get $\frac{d}{dt}\alpha_t^H(H) = \alpha_t^H(\{H, H\}) = 0$ from the antisymmetry of the Poisson bracket. As $\alpha_0^H(H) = H$, c) follows. Inserting this back into the above calculation yields $\frac{d}{dt}\alpha_t^H(f) = \{\alpha_t^H(f), H\}$. □

This theorem shows that Poisson brackets allow us to neatly express the dynamics, and that Hamiltonian vector fields are special because they preserve the Poisson bracket. Another way to see that Hamiltonian vector fields are special vector fields is the following fact, known as Liouville's Theorem, that states that the flow of a Hamiltonian vector fields preserves phase space volume. Recall that the divergence of a vector field $\mathcal{V} \in \mathcal{X}(M), M \subset \mathbb{R}^{2d}$, is $\operatorname{div}\mathcal{V} = \sum_{k=1}^{2d} \partial_{\xi_k} \mathcal{V}_k = \operatorname{Tr}(D\mathcal{V})$, where $D\mathcal{V}$ is the Jacobi matrix of $\mathcal{V}$.

Proposition 1.17 (Liouville's Theorem). *The flow ϕ^H of a Hamiltonian vector field $\mathcal{V}_H$ preserves phase space volume in the sense that $\operatorname{div}\mathcal{V}_H = 0$ and $\det D\phi_t^H = 1$ for all $t \in \mathbb{R}$ for which the flow is defined.*

Proof. For the divergence of the vector field we compute

$$\begin{aligned}
\operatorname{div}\mathcal{V}_H &= \sum_{k=1}^{2d} \partial_{\xi_k} (\mathcal{V}_H)_k = \sum_{k=1}^{d} \partial_{x_k} (\mathcal{V}_H)_k + \sum_{k=1}^{d} \partial_{p_k} (\mathcal{V}_H)_{k+d} \\
&= \sum_{k=1}^{d} \partial_{x_k} \partial_{p_k} H + \sum_{k=1}^{d} \partial_{p_k} (-\partial_{x_k} H) = 0.
\end{aligned}$$

The proof of $\det D\phi_t^H = 1$ will be given in the exercises. □

It is worth noting that Theorem 1.16 says in particular that despite our generalization from Newtonian Mechanics with conservative force fields to the Hamiltonian setting, the important feature of energy conservation is still present in the more general Hamiltonian formalism: $H(\phi_t^H(\xi)) = H(\xi)$. That is, solutions $t \mapsto \xi(t)$ of Hamilton's Equation are confined to surfaces on which H is constant.

Part c) of the theorem suggests the following definition.

Definition 1.18 (Constants of motion). Let (M, H) be a Hamiltonian system. A *constant of motion* (or *conserved quantity*) is a function $f \in C^\infty(M)$ such that

$$\alpha_t^H(f) = f, \qquad t \in \mathbb{R}. \tag{1.50}$$

So a constant of motion is what the name suggests: A quantity that is invariant under time evolution.

Lemma 1.19 (Constants of motion and Poisson brackets). *Let (M, H) be a Hamiltonian system. A function $f \in C^\infty(M)$ is a constant of motion if and only if*

$$\{f, H\} = 0. \tag{1.51}$$

In particular, H is a constant of motion. Given two constants of motion f, g, also $\{f, g\}$ is a constant of motion[a].

[a] That is, the constants of motion form a Lie subalgebra of $C^\infty(M)$.

Proof. The first claim follows directly from Theorem 1.16 b). The Hamiltonian is a constant of motion because the Poisson bracket is antisymmetric. The last statement is a consequence of the Jacobi identity of the Poisson bracket. □

1.3 Symmetries

The analysis of a classical or quantum physical system often benefits from understanding its symmetries. Here we introduce the tools to talk about symmetries in Hamiltonian mechanics, but keep everything at an elementary level. Much more can be said with a background in Lie groups.

Definition 1.20 (Symmetries). A *symmetry* of a Hamiltonian system (M, H) is a diffeomorphism $\varphi \in \mathrm{Diff}(M)$ such that its pullback[a]

$$\varphi^* : C^\infty(M) \to C^\infty(M), \qquad \varphi^*(f) := f \circ \varphi^{-1}, \tag{1.52}$$

respects the Poisson bracket, namely

$$\varphi^*(\{f, g\}) = \{\varphi^*(f), \varphi^*(g)\}, \qquad f, g \in C^\infty(M), \tag{1.53}$$

and leaves the Hamiltonian invariant, namely

$$\varphi^*(H) = H. \tag{1.54}$$

[a]The upper index * is the usual notation for pullback, and unfortunately clashes with the equally usual notation for *-operations. It should always be clear from context what is meant.

By Theorem 1.16, the Hamiltonian flow maps ϕ_t^H are symmetries of H. Of course also the identity id_M is a symmetry of H. In case φ is a symmetry of H, also φ^{-1} is because for any $f, g \in C^\infty(M)$, we have $\{f \circ \varphi^{-1}, g \circ \varphi^{-1}\} = \{f \circ \varphi^{-1} \circ \varphi, g \circ \varphi^{-1} \circ \varphi\} = \{f, g\}$. So we conclude that the symmetries of H form a group (a subgroup of $\mathrm{Diff}(M)$), called the *symmetry group of* H.

We next show what can be expected from symmetries: They preserve the dynamics given by H. In preparation, recall that chain rule $D(a \circ b)(\xi) = (Da)(b(\xi))(Db)(\xi)$ (for differentiable function $b : \mathbb{R}^n \to \mathbb{R}^m$, $a : \mathbb{R}^m \to \mathbb{R}^k$). For scalar functions $f : \mathbb{R}^n \to \mathbb{R}$, we have $(Df) = (\nabla f)^T$, hence for $\phi : \mathbb{R}^n \to \mathbb{R}^n$ we get $\nabla(f \circ \phi)(\xi) = (D\phi)^T(\xi)(\nabla f)(\phi(\xi))$.

Proposition 1.21 (Symmetries commute with the dynamics). *Let φ be a symmetry of the Hamiltonian system (M, H) and assume the Hamiltonian flow ϕ^H is global. Then the Hamiltonian flow commutes with the symmetry,*

$$\varphi(\phi_t^H(\xi)) = \phi_t^H(\varphi(\xi)), \qquad t \in \mathbb{R},\ \xi \in M, \tag{1.55}$$

and the dynamics commutes with the pullback of the symmetry,

$$\varphi^* \circ \alpha_t^H = \alpha_t^H \circ \varphi^*, \qquad t \in \mathbb{R}. \tag{1.56}$$

Proof. Let $f, g \in C^\infty_{\mathbb{R}}(M)$. Since φ respects the Poisson bracket, we have for any $\xi \in M$

$$\begin{aligned}
\{f \circ \varphi, g \circ \varphi\}(\xi) &= \langle \nabla(f \circ \varphi), \Sigma \nabla(g \circ \varphi)\rangle(\xi) \\
&= \langle (D\varphi)(\xi)^T (\nabla f)(\varphi(\xi)), \Sigma (D\varphi)(\xi)^T (\nabla g)(\varphi(\xi))\rangle \\
&= \{f, g\}(\varphi(\xi)) \\
&= \langle (\nabla f)(\varphi(\xi)), \Sigma (\nabla g)(\varphi(\xi))\rangle.
\end{aligned}$$

As f, g were arbitrary, $(D\varphi)(\xi)\Sigma(D\varphi)(\xi)^T = \Sigma$ follows. The invariance of the Hamiltonian, $H \circ \varphi^{-1} = H$, leads by differentiation to

$$\begin{aligned}
(\nabla H)(\xi) &= \nabla(H \circ \varphi^{-1})(\xi) \\
&= (D\varphi^{-1})(\xi)^T (\nabla H)(\varphi^{-1}(\xi)) = ((D\varphi)(\xi)^T)^{-1} (\nabla H)(\varphi^{-1}(\xi)).
\end{aligned}$$

Combining these two statements, we get the invariance of the Hamiltonian vector field,

$$\begin{aligned}
(D\varphi)(\xi) \mathcal{V}_H(\xi) &= (D\varphi)(\xi)\Sigma\nabla H(\xi) = \Sigma((D\varphi)(\xi)^T)^{-1}\nabla H(\xi) = \Sigma \nabla H(\varphi)(\xi) \\
&= \mathcal{V}_H(\varphi(\xi)).
\end{aligned}$$

Now consider the function $\eta(t) := \varphi(\phi_t^H(\xi))$. At time $t = 0$, it agrees with $\varphi(\xi)$, and its time derivative is

$$\dot\eta(t) = (D\varphi)(\phi_t^H(\xi)) \mathcal{V}_H(\phi_t^H(\xi)) = \mathcal{V}_H(\varphi(\phi_t^H(\xi))) = \mathcal{V}_H(\eta(t)).$$

So by uniqueness (Theorem 1.4), it must agree with the solution with initial condition $\varphi(\xi)$, namely $t \mapsto \phi_t^H(\varphi(\xi))$. Hence we have shown that the Hamiltonian flow commutes with the symmetry.

The second statement follows immediately by pullback. □

We now consider the connection between symmetries and dynamics. To that end, we consider a function $A \in C^\infty_{\mathbb{R}}(M)$ as a Hamiltonian in its own right, with associated vector field $\mathcal{V}_A := \Sigma\nabla A$, flow ϕ^A, and pulled back flow $\alpha^A_s := (\phi^A_s)^*$. The physical interpretation of such a flow is not that of time evolution when a Hamiltonian $H \neq A$ generating time evolution by its flow α^H_t is already fixed. This will become clear in the examples below. To indicate the difference in interpretation, we typically write s for the real parameter of these flows.

Theorem 1.22 (Hamiltonian form of Noether's Theorem). *Let (M, H) be a Hamiltonian system. A function $A \in C^\infty_{\mathbb{R}}(M)$ is a conserved quantity if and only if its pulled back flow (assumed to be global) α^A_s, $s \in \mathbb{R}$, consists of symmetries.*

Proof. We have seen in Theorem 1.16 that the flow maps α^A_s, $s \in \mathbb{R}$, always preserve the Poisson bracket. Hence the α^A_s are symmetries if and only if they preserve the Hamiltonian, i.e. $\alpha^A_s(H) = H$, $s \in \mathbb{R}$. Thinking of A as the Hamiltonian and A as the observable, this is equivalent to H being a constant of motion for the dynamics generated by A. According to Lemma 1.19, this is equivalent to $0 = \{H, A\} = -\{A, H\}$. □

Noether's Theorem provides a link between symmetries of Hamiltonian systems and conserved quantities. The simplest example of this is the Hamiltonian itself, i.e. $A = H$. Then the corresponding symmetries are the symmetries given by the Hamiltonian flow, i.e. the *time translations*.

Example 1.23 (Spatial translation symmetry and momentum conservation). Let $k \in \{1, \dots, d\}$. We consider the k-th component P_k of momentum (1.39), on the phase space $M = \mathbb{R}^{2d}$. Its Hamiltonian vector field is constant, namely

$$\mathcal{V}_{P_k}(x, p) = \Sigma\nabla p_k = \begin{pmatrix} e_k \\ 0 \end{pmatrix}, \tag{1.57}$$

with e_k the k-th canonical basis vector of $\mathbb{R}^d$. By solving the differential equation $\dot{x}(s) = e_k$, $\dot{p}(s) = 0$, the flow of this vector field is seen to be

$$\phi^{P_k}_s\left(\begin{pmatrix} x \\ p \end{pmatrix}\right) = \begin{pmatrix} x + se_k \\ p \end{pmatrix}, \qquad s \in \mathbb{R}. \tag{1.58}$$

So we conclude that the k-th component of momentum generates the symmetry of *spatial translations in the direction of e_k*. Given a Hamiltonian H, this flow is a symmetry of H if and only if H does not depend on x_k (momentum conservation).

In particular, for the free Hamiltonian $H(x, p) = \frac{1}{2m}\|p\|^2$, this is a symmetry for all k, meaning that all components of momentum are constants of motion.

A more interesting example is given by a multi particle phase space of the form $U = \mathbb{R}^{Nd}, M = \mathbb{R}^{Nd} \times \mathbb{R}^{Nd}$ (with $N \in \mathbb{N}, N > 1$). We interpret its elements

$$\xi = (x^{(1)}, \dots, x^{(N)}, p^{(1)}, \dots, p^{(N)}) \tag{1.59}$$

as describing the positions $x^{(k)} \in \mathbb{R}^d$ and momenta $p^{(k)} \in \mathbb{R}^d$ of N particles moving in $\mathbb{R}^d$, $k = 1, \dots, N$. To model a situation in which the only forces acting in the

system act between the particles (like attraction/repulsion due to an electrostatic or gravitational potential V, for example), one considers a Hamiltonian of the form

$$H(x^{(1)}, \dots, x^{(N)}, p^{(1)}, \dots, p^{(N)}) = \sum_{j=1}^{N} \frac{\|p^{(j)}\|^2}{2m} + \sum_{i<j} V(x^{(i)} - x^{(j)}). \tag{1.60}$$

Here $V \in C^\infty_{\mathbb{R}}(\mathbb{R}^d)$ is a potential for the force between the particles, m the mass of the particles (here all masses are assumed to be the same), and the second sum runs over all pairs $(i, j) \in \{1, \dots, N\}^2$ with $i < j$.

In this example, the spatial translations moving the individual positions $x^{(j)}$ independently do not preserve H (unless V is constant). However, the collective "center of mass" translation

$$x^{(j)} \mapsto x^{(j)} + s\, e_k, \qquad s \in \mathbb{R}, \tag{1.61}$$

that moves all coordinates in the same direction $e_k \in \mathbb{R}^d$, *does* preserve H because the potential term depends only on differences of positions. This symmetry corresponds to the k-th component of the *total momentum*, defined as

$$P^{\text{tot}}_k(x^{(1)}, \dots, x^{(N)}, p^{(1)}, \dots, p^{(N)}) = \sum_{j=1}^{N} p^{(j)}_k. \tag{1.62}$$

To see that the flow of (1.62) really generates the symmetries (1.61), we note

$$\mathcal{V}_{P^{\text{tot}}_k}(x^{(1)}, \dots, x^{(N)}, p^{(1)}, \dots, p^{(N)}) = \Sigma \begin{pmatrix} 0 \\ \vdots \\ 0 \\ e_k \\ \vdots \\ e_k \end{pmatrix} = \begin{pmatrix} e_k \\ \vdots \\ e_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

A comparison with (1.57) and (1.58) shows that P^{tot}_k generates translations in the $(e_k, \dots, e_k)$ direction (leaving all momenta invariant), i.e. (1.61).

For our next example, we argue conversely and start with the symmetries.

Example 1.24 (Rotation symmetry and angular momentum). Let $U \subset \mathbb{R}^d$ be a rotationally symmetric configuration space, i.e. $RU = U$ for all $R \in \mathrm{SO}(d)$ (for example $U = \mathbb{R}^d$ or $U = \{x \in \mathbb{R}^d \,:\, \|x\| < r\}$ an open ball). Then $x \mapsto Rx$ is a diffeomorphism of U, and

$$R \oplus R : \begin{pmatrix} x \\ p \end{pmatrix} \mapsto \begin{pmatrix} Rx \\ Rp \end{pmatrix} \tag{1.63}$$

is a diffeomorphism of phase space $M = U \times \mathbb{R}^d$ that preserves Poisson brackets because $R \oplus R$ commutes with Σ. If H is a rotationally invariant Hamiltonian, i.e. $H(x, p) = H(Rx, Rp)$ for all $R \in \mathrm{SO}(d)$, then each $R \oplus R$ is a symmetry of (M, H). A

typical form of such a Hamiltonian is

$$H(x,p) = \frac{\|p\|^2}{2m} + V(x), \qquad V(x) = v(\|x\|). \tag{1.64}$$

A potential of the form $V(x) = v(\|x\|)$ for some smooth $v : \mathbb{R}_+ \to \mathbb{R}$ is clearly rotationally symmetric and called *central potential*. A prominent example is the Coulomb potential (1.27).

We now restrict to dimension $d = 3$ and define angular momentum.

Definition 1.25 (Angular momentum). The function

$$L : \mathbb{R}^6 \to \mathbb{R}^3, \qquad L(x,p) := x \times p = \begin{pmatrix} x_2p_3 - x_3p_2 \\ x_3p_1 - x_1p_3 \\ x_1p_2 - x_2p_1 \end{pmatrix}, \tag{1.65}$$

is called *angular momentum*.

The significance of the components L_k of angular momentum is that they are the Hamiltonian functions generating the flows $(R_k(s) \oplus R_k(s))^*$ on $C^\infty(\mathbb{R}^6)$ given by rotations $R_k(s) \in \mathrm{SO}(3)$ around the e_k-axis by angle s, i.e.

$$R_1(s) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos s & \sin s \\ 0 & -\sin s & \cos s \end{pmatrix}, \qquad R_2(s) = \begin{pmatrix} \cos s & 0 & \sin s \\ 0 & 1 & 0 \\ -\sin s & 0 & \cos s \end{pmatrix},$$
$$R_3(s) = \begin{pmatrix} \cos s & \sin s & 0 \\ -\sin s & \cos s & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

This will be shown in Exercise 1.4.

These are the three most important applications of Noether's Theorem: Time translation invariance corresponds to conservation of energy, spatial translation invariance corresponds to conservation of momentum, and rotational invariance corresponds to conservation of angular momentum.

The following example illustrates the usefulness of angular momentum conservation.

Example 1.26 (Angular momentum conservation and planar trajectories). On the configuration space $U = \mathbb{R}^3$, let $V : \mathbb{R}^3 \to \mathbb{R}$, $V(x) = v(\|x\|)$, be a central potential. Then every solution $t \mapsto x(t)$ of Newton's Equation with force $F = -\nabla V$ is planar, i.e. lies in a two-dimensional subspace of $\mathbb{R}^3$.

Proof: Let $(x_0, p_0) \in \mathbb{R}^6$ be an arbitrary initial condition, and let $\mathcal{P} \subset \mathbb{R}^3$ be a two-dimensional subspace containing x_0 and p_0 (this is unique unless x_0 and p_0 are parallel). By definition of angular momentum as a cross product, $L(x(t), p(t))$ is orthogonal to $x(t)$ and $p(t)$, so $L := L(x(0), p(0))$ is orthogonal to $\mathcal{P}$. Since L is conserved, it follows that $x(t)$ and $p(t)$ are orthogonal to L for every t, i.e. the trajectory lies in $\mathcal{P}$.□

Further information on the trajectory can be extracted from the fact that the last

component L_3 of L is conserved: At any time t, we have

$$x_1(t)p_2(t) - x_2(t)p_1(t) = L_3(x(0), p(0)), \tag{1.66}$$

where the right hand side is independent of t.

This explains in particular why the earth revolves around the sun in a (approximately) planar orbit: It is a consequence of the centrality of the gravitational force and the resulting conservation of angular momentum.

1.4 Probability measures and states

In the theory of classical mechanics that we considered up to this point, all quantities take definite values: The initial condition $\xi_0 = (x_0, p_0) \in M$ consists of a single position x_0 and a single momentum p_0, and at every instance of time, the solution of Hamilton's Equation likewise gives a single point $\xi(t)$ in phase space. Also every other observable, like the energy or second component of angular momentum, takes a definite value $H(\xi(t))$ or $L_2(\xi(t))$.

We now want to extend the theory in such a way that we can describe situations in which we have *incomplete information* about the system. For example, we want to be able to describe initial conditions in which the position of particle is x_0 with probability $1/4$ and y_0 with probability $3/4$, or randomly distributed over U in some way, and analogously for momenta. In such a situation, to be formalized below, observables $f \in C^\infty(M)$ will no longer have definite values. Instead they will become random variables. Their values $f(\xi)$, $\xi \in M$, will be randomly distributed over the *spectrum of* f, defined as the closure of its range,

$$\sigma(f) := \overline{f(M)}. \tag{1.67}$$

There are several motivations for considering such probabilistic (or statistical) situations.

a) In real life scenarios, a position/momentum of a particle is never known with certainty, but subject to error and imperfections in the experiment used to measure it. To measure an observable f accurately, one would typically set up a series of experiments on identically prepared systems (a so-called statistical ensemble, say consisting of N experiments), measure the value of f in each instance of the experiment, and obtain measurement data $f_1, \dots, f_N \in \mathbb{R}$. One would then perform a statisical analysis of these data, i.e. consider their mean value $\langle f \rangle := N^{-1}(f_1 + \dots + f_N)$, their variance $N^{-1}\sum_{k=1}^{N}(f_k - \langle f \rangle)^2$, etc., and compare these numbers with predictions from theory.

b) Another motivation is that when considering a system consisting of many particles, like for example a gas of atoms, one is not interested in the positions and momenta of the individual atoms, but rather in their average values over the whole gas. Statistical averages describe the macroscopic/collective properties of the gas, such as its pressure.

c) In quantum mechanics, probabilistic aspects are unavoidable. Hence a proba-

bilistic version of classical mechanics will lead to a clearer comparison with quantum mechanics.

These probabilistic scenarios are adequatly described by measure theory, see Appendix M.1 for some reminders and notation. We fix a Borel probability measure

$$\mu : \mathfrak{B}(M) \to [0,1] \tag{1.68}$$

on a phase space M, and write $\mathcal{P}(M)$ for the set of all these measures. For every observable $f \in C^\infty_{\mathbb{R}}(M)$, we consider the pushforward measure $f_*\mu$ (Def. M.5) of μ by f. This is a Borel probability measure on $\mathbb{R}$ with support

$$\operatorname{supp}(f_*\mu) \subset \sigma(f), \tag{1.69}$$

see Proposition M.6. As these measures describe how the values of the observables are distributed, we may call $f_*\mu$ also the *probability distribution of f under μ*. This probability measure determines the probabilities that the value of f lies in some (Borel) set. Other relevant information contained in $f_*\mu$ are the expectation value and uncertainty of f (cf. M.4). We collect the necessary definitions here.

Definition 1.27 (Expectation values and uncertainties of observables). Let M be a phase space, $\mu \in \mathcal{P}(M)$, and $f \in C^\infty_{\mathbb{R}}(M)$ an observable.

a) The *probability* that the value of f lies in the Borel set $B \subset \mathbb{R}$ is

$$\Pr_\mu(f \in B) := (f_*\mu)(B). \tag{1.70}$$

b) If the first moment of $f_*\mu$ exists, the *expectation value* of f in μ is defined as

$$\omega_\mu(f) := \mathbb{E}(f_*\mu) = \mathbb{M}_1(f_*\mu). \tag{1.71}$$

c) If the second moment of $f_*\mu$ exists, the *uncertainty* of f in μ is defined as

$$\Delta_\mu(f) := \sqrt{\mathbb{V}(f_*\mu)}, \tag{1.72}$$

the standard deviation of $f_*\mu$.

d) If μ has compact support, $\omega_\mu(f)$ exists for all $f \in C^\infty_{\mathbb{R}}(M)$. The map

$$\omega_\mu : C^\infty_{\mathbb{R}}(M) \to \mathbb{R}, \qquad f \mapsto \omega_\mu(f) \tag{1.73}$$

is called the *state given by μ*.

The notation $\omega_\mu(f)$ and $\Delta_\mu(f)$ is typical in quantum physics and hence already introduced here. We note that for $\mu \in \mathcal{P}_c(M)$, the map (1.73) is linear. We extend it to all of $C^\infty(M)$ in a complex linear manner, so ω_μ is a complex linear functional on $C^\infty(M)$ (typically not an algebra homomorphism). It is clearly *normalized* in the sense that

$$\omega_\mu(1) = 1, \tag{1.74}$$

and *positive* in the sense that[4]

$$f \geq 0 \implies \omega_\mu(f) \geq 0. \tag{1.75}$$

[4] Here $f \geq 0$ means $f(\xi) \geq 0$ for all $\xi \in M$.

In particular, $\omega_\mu(f^* f) \geq 0$ and $\omega_\mu(f^*) = \overline{\omega_\mu(f)}$ for all $f \in C^\infty(M)$ (real functions have real expectation values). For the uncertainty, it is useful to note $\Delta_\mu(f + c \cdot 1) = \Delta_\mu(f)$, $c \in \mathbb{R}$.

Let us unpack these definitions: For an observable $f \in C^\infty_{\mathbb{R}}(M)$ and a Borel probability measure μ on M, the probability that the value of f lies in the Borel set $B \subset \mathbb{R}$ is

$$\Pr_\mu(f \in B) = \int_{\mathbb{R}} \chi_B \, d(f_*\mu) = \int_M \chi_B(f(\xi)) \, d\mu(\xi) = \int_{f^{-1}(B)} d\mu(\xi). \tag{1.76}$$

The expectation value and uncertainty of f in μ are (if they exist)

$$\omega_\mu(f) = \mathbb{E}(f_*\mu) = \int_{\sigma(f)} \lambda \, d(f_*\mu)(\lambda) = \int_{\operatorname{supp}\mu} f(\xi) \, d\mu(\xi), \tag{1.77}$$

$$\Delta_\mu(f)^2 = \mathbb{V}(f_*\mu) = \int_{\sigma(f)} (\lambda - \omega_\mu(f))^2 \, d(f_*\mu)(\lambda) = \int_{\operatorname{supp}\mu} (f(\xi) - \omega_\mu(f))^2 \, d\mu(\xi).$$

where we have used the change of variables formula for pushforward measures (Prop. M.6). In the first integral in (1.77), we could also integrate over $\mathbb{R}$ instead of $\sigma(f)$, and in the second over M instead of $\operatorname{supp}\mu$, but this does not change anything because $f_*\mu$ is supported on $\sigma(f)$ and μ is supported on $\operatorname{supp}\mu$.

In the language of L^p-spaces, we have

$$\omega_\mu(f) = \int_M f \, d\mu, \tag{1.78}$$

$$\Delta_\mu(f) = \sqrt{\omega_\mu(f^2) - \omega_\mu(f)^2} = \|f - \omega_\mu(f)1\|_{L^2(M,\mu)}. \tag{1.79}$$

The expectation value $\omega_\mu(f)$ exists if $f \in L^1(M,\mu)$, and the uncertainty $\Delta_\mu(f)$ exists if $f \in L^2(M,\mu)$.

These formulas clearly express that the expectation value $\omega_\mu(f)$ is the mean value of $f(\xi) \in \sigma(f)$, averaged with the measure μ. The uncertainty $\Delta_\mu(f)$ captures how much $f(\xi)$ deviates from its expectation value $\omega_\mu(f)$. The integral expression for $\Delta_\mu(f)$ makes also sense for complex $f \in C^\infty(M)$, and we take it as our definition of uncertainty for complex f. Note that the expectation value can be any real number (or any complex number for complex f), but the uncertainty is always non-negative for real $f = f^*$.

So far we have considered the pushforward measures $f_*\mu$ for single observables $f : M \to \mathbb{R}$. We can however easily extend this idea to several observables.

Definition 1.28 (Joint probability measures). Let M be a phase space and $\mu \in \mathcal{P}(M)$.

a) Let $f_1, \ldots, f_n \in C^\infty_{\mathbb{R}}(M)$ and

$$f_1 \times \ldots \times f_n : M \to \mathbb{R}^n, \qquad \xi \mapsto \begin{pmatrix} f_1(\xi) \\ f_2(\xi) \\ \vdots \\ f_n(\xi) \end{pmatrix}. \tag{1.80}$$

The *joint probability distribution of* $f_1, \ldots, f_n$ *under* μ is the Borel probability measure $(f_1 \times \ldots \times f_n)_*\mu$ on $\mathbb{R}^n$.

b) The *probability measure for position* under μ is the joint probability measure for the components $X_1, \dots, X_d$ of position $X : M \to U, (x,p) \mapsto x$ (1.39), i.e.

$$(X_1 \times \dots \times X_d)_* \mu = X_* \mu, \tag{1.81}$$

and the *probability measure for momentum* under μ is the joint probability measure for the components $P_1, \dots, P_d$ of momentum $P : M \to \mathbb{R}^d, (x,p) \mapsto p$, i.e.

$$(P_1 \times \dots \times P_d)_* \mu = P_* \mu. \tag{1.82}$$

The probability measures for position/momentum under μ satisfy by definition

$$(X_* \mu)(A) = \mu(X^{-1}(A)) = \mu(A \times \mathbb{R}^d), \qquad A \in \mathfrak{B}(U), \tag{1.83}$$

$$(P_* \mu)(B) = \mu(P^{-1}(B)) = \mu(U \times B), \qquad B \in \mathfrak{B}(\mathbb{R}^d), \tag{1.84}$$

so $X_*\mu, P_*\mu$ are precisely the *marginals* of μ.

If $f_1, \dots, f_n \in C^\infty_{\mathbb{R}}(M)$ are observables and $B_1, \dots, B_n \in \mathfrak{B}(\mathbb{R})$ Borel sets, their joint probability measure under μ gives on the product set $B_1 \times \dots \times B_n$

$$\Pr_{(f_1 \times \dots \times f_n)_* \mu}(B_1 \times \dots \times B_n) = \mu(\{\xi \in M \ : \ f_1(\xi) \in B_1, \dots, f_n(\xi) \in B_n\}),$$

the probability that f_1 takes a value in B_1 *and* f_2 takes a value in B_2, ..., *and* f_n takes a value in B_n. By the change of variables formula (M.7), we have for any $(f_1 \times \dots \times f_n)_* \mu$-integrable function $g : \mathbb{R}^n \to \mathbb{C}$

$$\int_{\mathbb{R}^n} g \, d(f_1 \times \dots \times f_n)_* \mu = \int_M g(f_1(\xi), \dots, f_n(\xi)) \, d\mu(\xi). \tag{1.85}$$

In particular, for a momonomial $g(y_1, \dots, y_n) = y_1^{m_1} \dots y_n^{m_n}$, with $m_1, \dots, m_n \in \mathbb{N}_0$, the integrals

$$\int_{\mathbb{R}^n} y_1^{m_1} \cdots y_n^{m_n} \, d(f_1 \times \dots \times f_n)_* \mu(y) = \int_M f_1(\xi)^{m_1} \cdots f_n(\xi)^{m_n} \, d\mu(\xi) \tag{1.86}$$

are called the *mixed moments* of $(f_1 \times \dots \times f_n)_* \mu$. These are precisely the expectation values of products of the observables f_j. They encode not only the expectation values of the individual f_j's, but also their correlations (see Section 1.5).

Example 1.29. Let $\rho : M \to \mathbb{R}_+$ be a Lebesgue-integrable non-negative density with $\int_M \rho(\xi) d\xi = 1$, and consider the measure $d\mu(\xi) = \rho(\xi) d\xi$ (i.e. $\mu = \rho \ell$ with ℓ_M Lebesgue measure on M, see Def. M.7). Then, $f \in C^\infty_{\mathbb{R}}(M)$,

$$\Pr_\mu(f \in B) = \int_{f^{-1}(B)} \rho(\xi) d\xi, \qquad B \in \mathfrak{B}(\mathbb{R}),$$

$$\omega_\mu(f) = \int_M f(\xi) \rho(\xi) \, d\xi,$$

$$\Delta_\mu(f)^2 = \int_M (f(\xi) - \omega_\mu(f))^2 \rho(\xi) \, d\xi.$$

The probability measures for position and momentum are in this case given by $X_*\rho\ell_M = \rho_X\ell_U$ and $P_*\rho\ell_M = \rho_P\ell_{\mathbb{R}^d}$ with the densities

$$\rho_X(x) = \int_{\mathbb{R}^d} \rho(x,p)\,dp, \qquad \rho_P(p) = \int_U \rho(x,p)\,dx. \tag{1.87}$$

It is also very instructive to consider phase space measures concentrated at a single point.

Example 1.30 (Dirac measures). Let $\xi_0 \in M = U \times \mathbb{R}^d$, and consider the Borel probability measure $\mu := \delta_{\xi_0}$ (Dirac measure). Then, for every $f \in C^\infty_{\mathbb{R}}(M)$ and $B \in \mathfrak{B}(\mathbb{R})$

$$\Pr_{\delta_{\xi_0}}(f \in B) = \int\limits_{f^{-1}(B)} d\delta_{\xi_0}(\xi) = \begin{cases} 1 & f(\xi_0) \in B \\ 0 & f(\xi_0) \notin B \end{cases}, \tag{1.88}$$

i.e.

$$f_*\delta_{\xi_0} = \delta_{f(\xi_0)}. \tag{1.89}$$

Writing $\xi_0 = (x_0, p_0)$, the position and momentum probability measures are in this case

$$X_*\delta_{\xi_0} = \delta_{x_0}, \qquad P_*\delta_{\xi_0} = \delta_{p_0}. \tag{1.90}$$

Expectation values and uncertainties of observables are given by

$$\omega_{\delta_{\xi_0}}(f) = \int_M f(\xi)\,d\delta_{\xi_0}(\xi) = f(\xi_0), \tag{1.91}$$

$$\Delta_{\delta_{\xi_0}}(f)^2 = \int_M (f(\xi) - f(\xi_0))^2\,d\delta_{\xi_0}(\xi) = 0. \tag{1.92}$$

Hence in this example every observable f has the definite value $f(\xi_0)$ with zero uncertainty. This example shows how the previously considered non-probabilistic situation of initial conditions $\xi_0 \in M$ fits into the more general probabilistic setting.

We have just seen that Dirac measures are very special; they lead to zero uncertainty for all observables. Also the converse is true: Zero uncertainty states are given by Dirac measures. This and other characterizations of these so-called pure states will be given later in Theorem 1.35.

We have now learned how to evaluate observables in states given by probability distributions, generalizing their evaluations at phase space points $\xi \in M$. So far no Hamiltonian or dynamics has entered. We now comment on the Hamiltonian dynamics when the initial condition $\xi_0 \in M$ is generalized to a Borel probability measure μ.

Proposition 1.31 (Dynamics of probabilistic Hamiltonian systems). *Let (M,H) be a Hamiltonian system with Hamiltonian flow ϕ^H and μ a Borel probability measure on M. Then $\mu^H_t := (\phi^H_t)_*\mu$ is a one parameter family of Borel probability measures on M, and*

for any observable $f \in C^\infty(M)$,

$$\omega_\mu(\alpha_t^H(f)) = \omega_{\mu_t^H}(f), \qquad t \in \mathbb{R}. \tag{1.93}$$

These expectation values exist if and only if $\alpha_t^H(f) \in L^1(M, d\mu)$.

Proof. The pushforward measures $\mu_t^H = (\phi_t^H)_*\mu$ are Borel probability measures because μ is (Proposition M.6). The claimed identity (1.93) follows directly from the change of variables formula (M.7):

$$\omega_\mu(\alpha_t^H(f)) = \int_M f(\phi_t^H(\xi))\, d\mu(\xi) = \int_M f(\xi)\, d((\phi_t^H)_*\mu)(\xi) = \omega_{(\phi_t^H)_*\mu}(f).$$

□

The role of any physical theory is to give a mathematical model determining the expectation values of its observables and how they change with time. Formula (1.93) expresses that in our current setting of classical Hamiltonian mechanics, there are two equivalent ways of doing that: Either we can consider the state ω_μ as time-independent, and the observables f as changing with time according to the Hamiltonian flow $\alpha_t^H(f)$, or we can consider the observables as time-independent, and the state (probability measure) as changing with time according to $(\phi_t^H)_*\mu$. Both points of view give the same result (1.93). In quantum mechanics, the first point of view is called *Heisenberg picture* and the second one *Schrödinger picture*.

For $\mu = \delta_{\xi_0}$, both sides of (1.93) agree with $f(\xi(t))$, with $\xi(t) = \phi_t^H(\xi)$ the unique solution of Hamilton's equation with initial condition $\xi(0) = \xi_0$. In this case, we have $(\delta_{\xi_0})_t^H = \delta_{\phi_t^H(\xi_0)} = \delta_{\xi(t)}$, so a Dirac measure δ_{ξ_0} remains a Dirac measure under time evolution, concentrated at the phase space point $\xi(t)$.

For a general measure, the time evolution can however look more complicated. The following example illustrates this.

Example 1.32 (Spread of probability distributions under free dynamics). We consider the two-dimensional phase space $M = \mathbb{R}^2$ with the free dynamics from Example 1.2, i.e. the Hamiltonian $H(x, p) = \frac{p^2}{2m}$ with flow

$$\phi_t^H(x, p) = (x + \frac{t}{m}p, p). \tag{1.94}$$

For the initial condition, we consider a probability measure that has a density of product form, namely $\rho(x, p) = \rho_1(x)\rho_2(p)$ w.r.t. Lebesgue measure $dx\, dp$ on $\mathbb{R}^2$. Here ρ_1, ρ_2 are non-negative compactly supported functions in $L^1(\mathbb{R})$ with integral 1.

The expectation value and uncertainty of position at time $t = 0$ are

$$\omega_\mu(X) = \int_{\mathbb{R}^2} x\rho_1(x)\rho_2(p)\, dx\, dp = \int_{\mathbb{R}} x\rho_1(x)\, dx,$$
$$\omega_\mu(P) = \int_{\mathbb{R}^2} p\rho_1(x)\rho_2(p)\, dx\, dp = \int_{\mathbb{R}} p\rho_2(p)\, dp.$$

For simplicty, let us assume $\omega_\mu(X) = 0$. Then we find for general time $t \in \mathbb{R}$

$$\begin{aligned}\omega_\mu(\alpha_t^H(X)) &= \omega_\mu(X + \tfrac{t}{m}P) = \tfrac{t}{m}\omega_\mu(P),\\ \omega_\mu(\alpha_t^H(X^2)) &= \omega_\mu((X + \tfrac{t}{m}P)^2)\\ &= \omega_\mu(X^2) + \tfrac{2t}{m}\omega_\mu(XP) + \tfrac{t^2}{m^2}\omega_\mu(P^2)\\ &= \omega_\mu(X^2) + \tfrac{t^2}{m^2}\omega_\mu(P^2).\end{aligned}$$

In the last step, we have used that X and P are *uncorrelated* (see Section 1.5) in μ, namely $\omega_\mu(XP) = \omega_\mu(X)\omega_\mu(P)$: This follows from the product form of ρ and leads to $\omega_\mu(XP) = 0$ because $\omega_\mu(X) = 0$ by assumption.

As a consequence, (note $\omega_\mu(X) = 0 \implies \Delta_\mu(X)^2 = \omega_\mu(X^2)$)

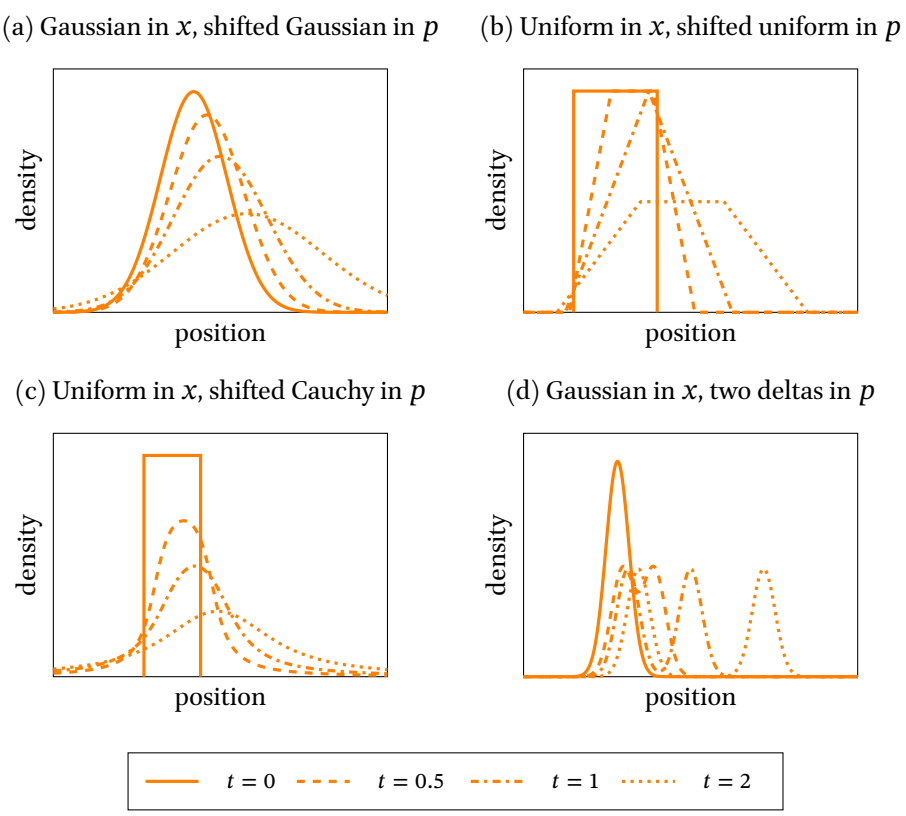
$$\Delta_\mu(\alpha_t^H(X)) = \sqrt{\Delta_\mu(X) + \tfrac{t^2}{m^2}\omega_\mu(P^2)} \to \infty \quad \text{for } t \to \pm\infty.$$

This matches our intuition coming from a statistical ensemble: The expectation value of position moves with the average velocity $\omega_\mu(P)/m$, and since the statisical ensemble of particles described by μ has varying velocities (momenta), the distribution of X will spread out as time progresses.

In the plots below, you see the density of the position probability measure, i.e. the function $\rho_{1,t}$ such that $X_*(\phi_t^H)_*\mu = \rho_{1,t}\,d\ell$, with ℓ Lebesgue measure on $\mathbb{R}$.

The four initial densities in these examples are

$$\begin{aligned}&\text{a) } \rho(x,p) &&= \frac{1}{2\pi\sigma_x\sigma_p}\exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{(p-p_0)^2}{2\sigma_p^2}\right) && \text{(Gaussian in } x \text{ and } p),\\ &\text{b) } \rho(x,p) &&= \frac{1}{4ab}\,\chi_{[-a,a]}(x)\,\chi_{[p_0-b,p_0+b]}(p) && \text{(uniform in } x \text{ and } p),\\ &\text{c) } \rho(x,p) &&= \frac{1}{2a}\chi_{[-a,a]}(x)\,\frac{1}{\pi\big(1+(p-p_0)^2\big)} && \text{(uniform in } x, \text{ Cauchy in } p),\\ &\text{d) } d\mu_0(x,p) &&= \frac{\exp\left(-\frac{x^2}{2\sigma_x^2}\right)}{2\sqrt{2\pi}\sigma_x}\big(d\delta_{p_0}(p) + d\delta_{p_1}(p)\big) && \text{(Gaussian in } x, \text{ two Diracs in } p),\end{aligned}$$



If we instead look at momentum P, we find that the probability measure $(\phi_t^H)_*(P_*\mu)$ for momentum does not depend on time because P is a constant of motion of the free Hamiltonian H. In particular, $\omega_\mu(\alpha_t^H(P))$ and $\Delta_\mu(\alpha_t^H(P))$ are time-independent.

In quantum mechanics, the phenomenon described above is called "spreading of the wave packet", but we see that it occurs in the same way in classical (statistical) mechanics. See Exercise 1.8 and Problems 1.7, 1.8,1.10 for more examples of dynamics with probability measures.

1.5 Composite systems and correlations

In this section we have a closer look at states and how they behave across subsystems with respect to their correlations. The basic definition is the following.

Definition 1.33 (Subsystems and correlations). Let $\mu \in \mathcal{P}_c(M)$ and $f, g \in C^\infty(M)$ be real ($f = f^*, g = g^*$). The *covariance of* f, g *in* μ is

$$\mathrm{Cov}_\mu(f,g) := \omega_\mu(fg) - \omega_\mu(f)\omega_\mu(g). \tag{1.95}$$

If $\Delta_\mu(f) \neq 0$ and $\Delta_\mu(g) \neq 0$, the *correlation of* f, g *in* μ is defined as

$$\mathrm{Corr}_\mu(f,g) := \frac{\mathrm{Cov}_\mu(f,g)}{\Delta_\mu(f)\Delta_\mu(g)} = \frac{\omega_\mu(fg) - \omega_\mu(f)\omega_\mu(g)}{\Delta_\mu(f)\Delta_\mu(g)}. \tag{1.96}$$

Two observables are called *uncorrelated in* μ if $\mathrm{Cov}_\mu(f,g) = 0$.

In this definition, we have restricted ourselves to compactly supported measures $\mu \in \mathcal{P}_c(M)$ in order to guarantee existence of the expectation values and uncertainties that occur. But as for states (1.73), $\mathrm{Cov}_\mu(f,g)$ and $\mathrm{Corr}_\mu(f,g)$ also exist for non-compactly supported measures μ for suitable f, g. For instance $\mathrm{Cov}_\mu(f,g)$ is well-defined for $\mu \in \mathcal{P}(M)$ and $f, g, fg \in L^1(M, d\mu)$.

Two remarks are $\mathrm{Cov}_\mu(f,f) = \Delta_\mu(f)^2$ and $\mathrm{Corr}_\mu(f,g) \in [-1,1]$. To prove the latter, we use the Cauchy-Schwarz Inequality:

$$\begin{aligned} |\mathrm{Cov}_\mu(f,g)|^2 &= |\omega_\mu((f-\omega_\mu(f)1)(g-\omega_\mu(g)1))|^2 \\ &= \left| \int_M (f(\xi)-\omega_\mu(f))(g(\xi)-\omega_\mu(g))\, d\mu(\xi) \right|^2 \\ &\leq \int_M |f(\xi)-\omega_\mu(f)|^2\, d\mu(\xi) \cdot \int_M |g(\xi)-\omega_\mu(g)|^2\, d\mu(\xi) \\ &= \Delta_\mu(f)^2 \Delta_\mu(g)^2, \end{aligned}$$

this implies $\mathrm{Corr}_\mu(f,g) \in [-1,1]$.

Any observable f is perfectly correlated with itself in the sense that $\mathrm{Corr}_\mu(f,f) = 1$. The following example illustrates partial correlations in a physical system.

Example 1.34 (Correlations of harmonically coupled particles). We consider the phase space $M = \mathbb{R}^4$ and interpret its elements (x_1, x_2, p_1, p_2) as describing two particles, one with position/momentum (x_1, p_1) and another one with position/momentum (x_2, p_2). As state, we take

$$d\mu(x_1, x_2, p_1, p_2) = \rho(x_1)\rho(x_2)\, dx_1\, dx_2\, d\delta_0(p_1)\, d\delta_0(p_2),$$

with the density ρ satisfying $\rho \geq 0$, $\int_{\mathbb{R}} \rho = 1$, $\int_{\mathbb{R}} x_1 \rho(x_1) dx_1 = 0$, $\int_{\mathbb{R}} x_1^2 \rho(x_1) dx_1 > 0$. This corresponds to a situation in which the momenta p_1 and p_2 are zero with certainty, and the position variables x_1, x_2 are distributed with mean value 0 and non-zero uncertainty.

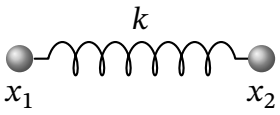
Because of the product form of the density, we see that X_1 and X_2 are uncorrelated,

$$\mathrm{Corr}_\mu(X_1, X_2) = 0.$$

We now consider the Hamiltonian function $H = \frac{1}{2m}(P_1^2 + P_2^2) + \frac{k}{2}(X_1 - X_2)^2$, that is

$$H(x_1, x_2, p_1, p_2) = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + \frac{k}{2}(x_1 - x_2)^2. \tag{1.97}$$

Here $m > 0$ is the mass, and $k > 0$ a parameter measuring the strength of the force. The first two terms are the kinetic energies of the two particles, and the quadratic potential corresponds to a linear force (i.e. a harmonic oscillator, see Problem 1.2) that pulls the particles towards each other. So one may visualize this as two particles that can move on a line and that are harmonically coupled by an elastic spring connecting them.



Computing the dynamics, one finds (Problem 1.11)

$$\mathrm{Corr}_\mu(\alpha_t^H(X_1)\alpha_t^H(X_2)) = \frac{\sin^2(\nu t)}{2}, \qquad \nu := \sqrt{\frac{k}{m}}. \tag{1.98}$$

The correlation depends on time and is always non-negative in this example. One can interpret this as follows: If the position x_1 is positive/negative, i.e. it deviates from its mean in the positive/negative direction, then also x_2 tends to deviate from its mean in positive/negative direction because of the force pulling the two particles together. This leads to a positive correlation, indicating that measuring $x_1(t)$ provides information about $x_2(t)$.

See also Exercise 1.7 for an exercise on correlations.

We will often be concerned with the question whether two observables are uncorrelated in a state, and with correlations of whole families of observables (subsystems). As a preparation for discussing this, we first make the easy observation that for a state ω_μ with a Dirac measure $\mu = \delta_{\xi_0}$, all correlations vanish. Hence Dirac measures are not only special w.r.t. uncertainties, but also w.r.t. correlations. To collect all their properties, we note that $\mathcal{P}(M)$ is a *convex set*, i.e. with $\mu_1, \mu_2 \in \mathcal{P}(M)$ and $0 \leq c \leq 1$, also

$$\mu := c\mu_1 + (1-c)\mu_2 \in \mathcal{P}(M) \tag{1.99}$$

is a Borel probability measure on M. By induction, it is clear that given $\mu_1, \dots, \mu_n \in \mathcal{P}(M)$ and numbers $0 \leq c_1, \dots, c_n \leq 1$ such that $c_1 + \dots + c_n = 1$, the convex combination $\sum_{k=1}^n c_k\mu_k$ is an element of $\mathcal{P}(M)$ (Exercise 1.10).

The measure $\sum_{k=1}^n c_k\mu_k$ is called a *mixture* of the measures $\mu_1, \dots, \mu_n$. We call a measure $\mu \in \mathcal{P}(M)$ *pure* if it is not a mixture, i.e. if $\mu = c\mu_1 + (1-c)\mu_2$ for $\mu_1, \mu_2 \in \mathcal{P}(M)$ and $c \in [0,1]$ is only possible in a trivial way, namely either $c \in \{0,1\}$ or $\mu_1 = \mu_2$. A measure that is not pure is called *mixed*.

With this terminology, we obtain the following characterization of pure measures.

Theorem 1.35 (Pure states on $C^\infty(M)$).

a) Let $\mu \in \mathcal{P}(M)$. The following are equivalent:

i) $\mu = \delta_\xi$ for some $\xi \in M$.

ii) ω_μ is completely uncorrelated: $\mathrm{Cov}_\mu(f,g) = 0$ for all $f, g \in C_c^\infty(M)$.

iii) $\Delta_\mu(f) = 0$ for every $f \in C_c^\infty(M)$.

iv) μ is pure.

b) Let $\mu \in \mathcal{P}_c(M)$. Then there exists a sequence $(\mu_n)_{n\in\mathbb{N}} \in \mathcal{P}_c(M)$, such that every μ_n is a convex combination of Dirac measures, and

$$\lim_{n\to\infty} \omega_{\mu_n}(f) = \omega_\mu(f), \qquad f \in C^\infty(M). \tag{1.100}$$

Proof. a) will be proven in the exericses.

b) For each $n \in \mathbb{N}$, choose a finite partition

$$\operatorname{supp}\mu = \bigsqcup_{j=1}^{N_n} B_{j,n}$$

of the compact set $\operatorname{supp}\mu$ into disjoint nonempty Borel sets $B_{j,n}$ of diameter at most $1/n$. For each j, pick a point $\xi_{j,n} \in B_{j,n}$, and define $\mu_n := \sum_{j=1}^{N_n} \mu(B_{j,n})\,\delta_{\xi_{j,n}}$, a convex combination of Dirac measures. Then we have for any $f \in C^\infty(M)$

$$\begin{aligned}
|\omega_{\mu_n}(f) - \omega_\mu(f)| &= \left| \sum_{j=1}^{N_n} \left(\mu(B_{j,n}) f(\xi_{j,n}) - \int_{B_{j,n}} f(\xi)\, d\mu(\xi) \right) \right| \\
&\le \sum_{j=1}^{N_n} \int_{B_{j,n}} \underbrace{|f(\xi_{j,n}) - f(\xi)|}_{\le \|\nabla f\|_\infty \|\xi_{j,n} - \xi\| \le \frac{1}{n}\|\nabla f\|_\infty} d\mu(\xi) \\
&\le \frac{1}{n} \|\nabla f\|_\infty,
\end{aligned}$$

which converges to 0 for $n \to \infty$. □

The first part of the theorem says in particular that Dirac measures are precisely pure measures, i.e. the extremal points of the convex set $\mathcal{P}(M)$. The second part of the theorem says that while not every $\mu \in \mathcal{P}_c(M)$ is a convex combination of extremal points (convex combinations, just as linear combinations, are always finite, so convex combinations of Dirac measures only produce measures with finite support), we recover all of $\mathcal{P}_c(M)$ when we allow for "infinite convex combinations", i.e. limits of convex combinations.

In Example 1.34, we considered a phase space of product form $M = M_1 \times M_2$ (in that case, with $M_1 = M_2 = \mathbb{R}^2$), with the phase space variables (x_j, p_j) referring to M_j, $j = 1, 2$. Hence we may regard M_1, M_2 as subsystems of M, and M as the product of M_1 and M_2. The observable algebras $C^\infty(M_j)$ can then be regarded as subalgebras of $C^\infty(M)$ by promoting a function $f_1 \in C^\infty(M_1)$ to a function in $C^\infty(M)$ that is constant in the (x_2, p_2) variables.

This motivates the following definitions.

Definition 1.36 (Subsystems and composite systems). Let M, M_1, M_2 be phase spaces.

a) A *subsystem* is a unital *-subalgebra $\mathcal{A} \subset C^\infty(M)$.
b) Given two subsystems $\mathcal{A}, \mathcal{B} \subset C^\infty(M)$, the *bipartite system generated by $\mathcal{A}$ and $\mathcal{B}$* is the smallest subsystem $(\mathcal{A}, \mathcal{B})$ of $C^\infty(M)$ that contains $\mathcal{A}$ and $\mathcal{B}$.
c) The *product* (or *composite system*) of $C^\infty(M_1)$ and $C^\infty(M_2)$ is $C^\infty(M_1 \times M_2)$.
d) The *uncoupled product* (or *interaction-free product*) of two Hamiltonian systems (M_1, H_1) and (M_2, H_2) is $(M_1 \times M_2, H_1 + H_2)$, with the Hamiltonian

$$(H_1 + H_2)(\xi_1, \xi_2) := H_1(\xi_1) + H_2(\xi_2). \tag{1.101}$$

Several remarks are in order here.

Remark 1.37.

a) Our notion of subsystem is quite general. Note however that we demand sub*algebras* and not just subsets or subspaces. For example, if an observable f belongs to a subsystem $\mathcal{A}$, then also f^n belongs to $\mathcal{A}$ for all $n \in \mathbb{N}_0$ (this is necessary in order to formulate higher moments of the measures $f_*\mu$ entirely in terms of functions in $\mathcal{A}$), with $f, g \in \mathcal{A}$, also linear combinations $c_1 f + c_2 g$ and products fg lie in $\mathcal{A}$ (this is necessary for correlations), and with f, also f^* lies in $\mathcal{A}$ (this is necessary to be able to split $f \in \mathcal{A}$ into its real and imaginary parts in $\mathcal{A}$).

b) The name *bipartite* system indicates that it consists of two parts, i.e. two subsystems. These parts can relate to phase space variables of different particles as in Example 1.34, or to two different observers (typically called Alice and Bob in (quantum) information theory) who have access only to certain subalgebras of observables. For example, Alice/Bob might only have access to observables in certain regions of configuration space (see Example 1.40).

c) The uncoupled product of two Hamiltonian systems is the most boring way in which two Hamiltonian systems can be combined into a single one. In the uncoupled product, one can solve Hamilton's Equation in both subsystems separately, i.e. the two dynamics do not interfere in any way. Physically speaking, this means that the total energy of the product system is simply the sum of the two subsystem energies (1.101). In Example 1.34, we have a more interesting coupling in which the total energy depends on the relative position $x_1 - x_2$ of the two particles.

Let us consider the product of two systems $C^\infty(M_1)$ and $C^\infty(M_2)$ from the point of view of (algebraic) *tensor products*[5].

Lemma 1.38. *Let M_1, M_2 be phase spaces. The algebraic tensor product $C^\infty(M_1) \otimes C^\infty(M_2)$ embeds into $C^\infty(M_1 \times M_2)$, i.e.*

$$\iota : C^\infty(M_1) \otimes C^\infty(M_2) \to C^\infty(M_1 \times M_2), \tag{1.102}$$

$$\iota(f \otimes g)(\xi_1, \xi_2) := f(\xi_1) g(\xi_2) \tag{1.103}$$

extends linearly to an injective homomorphism of unital *-algebras.

Proof. Recalling that the tensor product is bilinear, and that products and conjugates in $C^\infty(M_1) \otimes C^\infty(M_2)$ are given by $(f_1 \otimes g_1) \cdot (f_2 \otimes g_2) = f_1 f_2 \otimes g_1 g_2$ and $(f \otimes g)^* = f^* \otimes g^*$, respectively, we see directly that ι is a *-algebra homomorphism. Furthermore, it is injective. We may choose the (f_n) as linearly independent, and the (g_m) as linearly independent. If

$$\sum_{n=1}^{N} f_n(\xi_1) g_n(\xi_2) = 0 \qquad \forall \xi_1 \in M_1,\ \xi_2 \in M_2,$$

then, for fixed but arbitrary $\xi_1 \in M_1$, we have $0 = \sum_{n,m} c_{nm} f_n(\xi_1) g_m \in C^\infty(M_2)$. As the (g_m) are linearly independent and ξ_1 was arbitrary, we conclude $0 = \sum_n c_{nm} f_n$ for all m. Hence

$$\sum_{n,m=1}^{N} c_{nm}\, f_n \otimes g_m = \sum_m \left(\sum_n c_{nm} f_n \right) \otimes g_m = 0.$$

[5] $C^\infty(M_1) \otimes C^\infty(M_2)$ is a unital *-algebra in the following manner: The linear structure is given by bilinearity of $\otimes$, the product by $(f_1 \otimes g_1) \cdot (f_2 \otimes g_2) = f_1 f_2 \otimes g_1 g_2$, the *-operation by $(f \otimes g)^* = f^* \otimes g^*$, and the unit is $1_{M_1} \otimes 1_{M_2}$.

□

An unsatisfactory feature of this definition is that the product $C^\infty(M_1 \times M_2)$ of $C^\infty(M_1)$ and $C^\infty(M_2)$ is different from the bipartite system generated by the subsystems $C^\infty(M_1)$ and $C^\infty(M_2)$. This is the case because the term "generated" refers to being generated as an algebra, whereas the product is not formulated in terms of the observable algebras $C^\infty(M_j)$, but rather in terms of the underlying spaces M_j. This problem disappears when equipping our observable algebras with suitable topologies and demanding that all systems (observable algebras) are complete in these topologies[6].

The smallest subalgebra of $C^\infty(M_1 \times M_2)$ containing $C^\infty(M_1)$ and $C^\infty(M_2)$ is the algebraic tensor product $C^\infty(M_1) \odot C^\infty(M_2)$, consisting of all polynomials in functions that depend only on the variables of M_1 or M_2, respectively. Closure in a suitable topology gives the topological tensor product $C^\infty(M_1) \otimes C^\infty(M_2) = C^\infty(M_1 \times M_2)$.

Given two subsystems, i.e. a bipartite system, it is often interesting to study the correlations between these subsystems, generalizing the notion of correlations of individual observables.

Definition 1.39 (Product states, classically correlated states, entangled states). Let $(\mathcal{A}, \mathcal{B})$ be a bipartite system in $C^\infty(M)$, and ω_μ a state on $C^\infty(M)$.

a) ω_μ is called *product state* over $(\mathcal{A}, \mathcal{B})$, or *uncorrelated* in $(\mathcal{A}, \mathcal{B})$, if f is uncorrelated to g in μ for any $f \in \mathcal{A}$ and any $g \in \mathcal{B}$. States that are not product states are called *correlated*.

b) A correlated state ω_μ is called *classically correlated* (or *separable*) over $(\mathcal{A}, \mathcal{B})$ if it is a pointwise limit of convex combinations of product states.

c) A correlated state ω_μ is called *entangled* over $(\mathcal{A}, \mathcal{B})$ if it is not separable over $(\mathcal{A}, \mathcal{B})$.

The name *product state* is easy to understand: By definition, ω_μ is a product state over $(\mathcal{A}, \mathcal{B})$ if

$$\omega_\mu(fg) = \omega_\mu(f)\omega_\mu(g), \qquad f \in \mathcal{A},\ g \in \mathcal{B}. \tag{1.105}$$

According to Theorem 1.35, pure states of $C^\infty(M)$ are uncorrelated for any bipartite system; they are even uncorrelated for $\mathcal{A} = \mathcal{B} = C^\infty(M)$.

Example 1.40 (Alice and Bob in classical mechanics). In the phase space $M = \mathbb{R}^d \times \mathbb{R}^d$ we consider two open disjoint subsets $U_A, U_B \subset \mathbb{R}^d$ of position space. We think of U_A and U_B as the laboratories of two scientists called Alice and Bob that are far apart. As subsystems of $C^\infty(M)$, we take

$$\mathcal{A} = \{f \in C^\infty(M) \ :\ f(x,p) = 0 \text{ if } x \notin U_A\} + \mathbb{C}1, \tag{1.106}$$
$$\mathcal{B} = \{f \in C^\infty(M) \ :\ f(x,p) = 0 \text{ if } x \notin U_B\} + \mathbb{C}1. \tag{1.107}$$

[6] On $C^\infty(M)$, consider the family of seminorms $p_{K,n}$ indexed by compact subsets $K \subset M$ and $n \in \mathbb{N}_0$,

$$p_{K,n} : C^\infty(M) \to \mathbb{R}_+, \qquad p_{K,n}(f) := \max_{|\alpha| \le n} \sup_{\xi \in K} |\partial^\alpha f(\xi)|. \tag{1.104}$$

Here $\alpha \in \mathbb{N}_0^{2d}$ is a multiindex, $|\alpha| = \alpha_1 + \ldots + \alpha_{2d}$, and $\partial^\alpha = \partial_{\xi_1}^{\alpha_1} \cdot \partial_{\xi_{2d}}^{2d}$. The initial topology of these seminorms is a good choice of topology on $C^\infty(M)$.

So functions in $\mathcal{A}$ are supported in U_A in x up to a possible constant part, and analogously for $\mathcal{B}$ and U_B. The constant functions are included in order to make sure that $\mathcal{A}, \mathcal{B}$ are subsystems in the sense of Definition 1.36.

In Exercise 1.11, you will study the correlations of such subsystems. In particular, you will find that $\mathcal{A}$ and $\mathcal{B}$ are correlated.

Whereas correlations naturally appear in classical bipartite systems, entanglement does not appear.

Proposition 1.41 (Absence of entanglement in classical mechanics). *Let M be a phase space, $(\mathcal{A}, \mathcal{B})$ a bipartite system in $C^\infty(M)$, and $\mu \in \mathcal{P}_c(M)$. Then ω_μ is separable over $(\mathcal{A}, \mathcal{B})$.*

Proof. According to Theorem 1.35 b), the state ω_μ can be written as a pointwise (in observables f) limit of convex combinations ω_n of states given by Dirac measures. According to Theorem 1.35 a), each Dirac measure is completely uncorrelated, so in particular a product state for $(\mathcal{A}, \mathcal{B})$. Hence ω_μ is separable over $(\mathcal{A}, \mathcal{B})$. □

This proposition says that entanglement is an empty concept for classical mechanics. In quantum mechanics, we will however find many examples of entangled states and discuss this phenomenon.

To conclude our brief tour of classical mechanics, we take another look at the expectation value functionals of classical mechanics. We have seen that given any compactly supported $\mu \in \mathcal{P}_c(M)$, the linear functional

$$\omega_\mu : C^\infty(M) \to \mathbb{C} \tag{1.108}$$

is normalized (1.74) and positive (1.75). It is natural to ask whether there exist more positive normalized linear functionals on $C^\infty(M)$ than those given by compactly supported Borel probability measures. Riesz' Representation Theorem (Theorem M.2) tells us that positive linear functionals $\omega : C(M) \to \mathbb{C}$ on the algebra $C(M)$ of *continuous* functions $M \to \mathbb{C}$ are in bijection with compactly supported Borel measures via

$$\omega(f) = \int_M f \, d\mu, \tag{1.109}$$

and normalization of ω clearly corresponds to μ being a probability measure. This result can be extended to the smaller algebra $C^\infty(M)$ (larger dual): The positive linear normalized functionals of $C^\infty(M)$ are also in bijection with compactly supported Borel measures via (1.109) because any such functional can be extended from $C^\infty(M)$ to $C(M)$ [BW00]. We record this result here without proof.

Theorem 1.42 (The states of $C^\infty(M)$). *The positive linear normalized functionals on $C^\infty(M)$ are in bijection with the compactly supported Borel probability measures on M via (1.109).*

Conceptually, this is an important step: We now realize that the states (measures) are a structure that is intrinsic to our observable algebra and does not have to be put in by hand.

Review of structure of classical mechanics

Let us review the structure of the theory of classical mechanics: A Hamiltonian system (M, H) consists of a phase space M and Hamiltonian function H. The role of the phase space is to define the classical algebra of observables $C^\infty(M)$, which contains in particular the components of position and momentum, etc. Due to the splitting of M into position and momentum space, we have a symplectic form (the antisymmetric matrix Σ), which defines the Poisson bracket on $C^\infty(M)$.

The Hamiltonian H is a real function that is picked from $C^\infty_{\mathbb{R}}(M)$. Typical examples are sums of kinetic and potential energy terms, but in principle, any choice works. The role of the Hamiltonian is to generate a dynamics: It defines a one-parameter group of automorphisms $(\alpha^H_t)_{t\in\mathbb{R}}$ of $C^\infty(M)$ that preserves all algebraic structures and in particular the Poisson bracket. The equation of motion, Hamilton's Equation, takes the simple form $\frac{d}{dt}\alpha^H_t(f) = \{\alpha^H_t(f), H\}$ for any $f \in C^\infty(M)$. Its solution describes how the observable f changes with time, i.e. how f varies along the flow defined by the Hamiltonian.

In case f does not change at all with time, we call it a constant of motion; this is equivalent to $\{f, H\} = 0$. Constants of motion are in bijection with one-parameter groups of symmetries preserving the Hamiltonian (Noether's Theorem).

In a probabilistic setting, one considers states of $C^\infty(M)$, i.e. positive linear normalized functionals $\omega : C^\infty(M) \to \mathbb{C}$. The space of all states (state space) can be identified with the space $\mathcal{P}_c(M)$ of all compactly supported Borel probability measures on M, and $\omega(f)$ has the meaning of expectation value of f in the state with a given measure. This also leads to the concept of uncertainties, and the special pure states that have no uncertainties at all (Dirac measures). These pure states are also free of correlations, whereas other states (which are always realized as limits of convex combinations of pure ones) do exhibit correlations. Across subsystems, all correlations that we find are classical, the concept of entangled states is not realized in classical mechanics.

1.6 Exercises and Problems

Exercise 1.1 (Local solutions). Solve the initial value problem for the one-dimensional smooth vector field $\mathcal{V} \in \mathcal{X}(\mathbb{R})$, $\mathcal{V}(\xi) = \xi^2$, i.e. find all smooth curves (functions) $\xi : I \to \mathbb{R}$ such that $\dot\xi(t) = \xi(t)^2$ and $\xi(0) = \xi_0$ for some prescribed $\xi_0 \in \mathbb{R}$. Show that only one initial condition yields a global solution.

Exercise 1.2 (Picard-Lindelöf Theorem). Prove part a) and b) of Theorem 1.4. Let $\mathcal{V}$ be smooth vector field on an open subset $U \subset \mathbb{R}^n$ and $\xi_0 \in \mathbb{R}^n$ and proceed in the following steps:

a) Show that for a smooth curve $\xi : I \to \mathbb{R}^n$ with I an open interval containing 0, the following are equivalent:

$$\dot\xi(t) = \mathcal{V}(\xi(t)), \xi(0) = \xi_0 \quad \text{and} \quad \xi(t) - \xi_0 = \int_0^t \mathcal{V}(\xi(s))ds, \quad \text{for all } t \in I. \tag{1.110}$$

Consider a compact subset $K \subset \mathbb{R}^n$ s.t. $\xi_0 \in K$. Define the set

$$C(I,K,\xi_0) := \{\gamma : I \to K \ : \ \gamma \text{ is continuous and } \gamma(0) = \xi_0\} \tag{1.111}$$

of continuous curves from I to K which take the value ξ_0 at 0. Equip this space with the metric

$$d : C(I,K,\xi_0) \times C(I,K,\xi_0) \to [0,\infty), \quad d(\gamma_1,\gamma_2) := \sup_{t\in I} \|\gamma_1(t) - \gamma_2(t)\|. \tag{1.112}$$

You may use in the following that $(C(I,K,\xi_0), d)$ is a complete metric space. Let moreover $\epsilon, \delta > 0$ and set $I := (-\epsilon, \epsilon)$ and $K := \overline{B_\delta(\xi_0)} \subset U$ for δ sufficiently small.

b) Show that there exist $C, D > 0$ s.t. $\|\mathcal{V}(x) - \mathcal{V}(y)\| < C\|x-y\|$ and $\|\mathcal{V}(x)\| < D$ for all $x, y \in K$. Then show for $\epsilon < \frac{\delta}{D}$ that the map

$$L : C(I,K,\xi_0) \to C(I,K,\xi_0), \quad L(\gamma)(t) := \xi_0 + \int_0^t \mathcal{V}(\gamma(s))ds \tag{1.113}$$

is well-defined.

c) Let now $\epsilon < \min\{\frac{\delta}{D}, \frac{1}{2C}\}$. Show that L is a contraction.

d) Recall the Banach Fixed Point Theorem and show part a) of Theorem 1.4.

e) Prove part b) of Theorem 1.4 by the Lemma of Zorn.

Exercise 1.3 (Stationary Particle). Consider $M = U \times \mathbb{R}^n$ and $H = \frac{P^2}{2m} + V(X)$ for a potential $V : U \to \mathbb{R}$. Show that a solution of Hamilton's equation $\xi : I \to M$ satisfies $0 = \dot{\xi}(t)$ for all $t \in I$ if and only if $x(t)$ is constant and $-\nabla V(x(t)) = 0$ for all $t \in I$.

Exercise 1.4 (Angular momentum). Show that the component L_k, $k = 1,2,3$, of angular momentum L (Def. 1.25) is the Hamiltonian function of rotations around the e_k-axis, i.e. of the flow $(R_k(s) \oplus R_k(s))^*$ on $C^\infty(\mathbb{R}^6)$.

Exercise 1.5 (Symmetries of the Coulomb potential). Consider the phase space $M = (\mathbb{R}^3 \backslash \{0\}) \times \mathbb{R}^3$ and the Hamilton function

$$H : M \to \mathbb{R}, \quad H(x,p) = \frac{\|p\|^2}{2m} - \frac{\alpha}{\|x\|}$$

for some $m, \alpha > 0$. Consider the vector valued functions $L, R : M \to \mathbb{R}^3$ given by

$$L(x,p) := x \times p = \begin{pmatrix} x_2p_3 - x_3p_2 \\ x_3p_1 - x_1p_3 \\ x_1p_2 - x_2p_1 \end{pmatrix}, \quad R(x,p) := \frac{x}{\|x\|} + \frac{1}{\alpha m} L(x,p) \times p \tag{1.114}$$

called angular momentum and Runge-Lenz-vector with components $L_i, R_i, i = 1, 2, 3$.

a) Compute the Poisson brackets

$$\{L_i, H\}, \quad \{R_i, H\}, \quad \{L_i, L_j\}, \quad \{L_i, R_j\}, \quad \{R_i, R_j\} \tag{1.115}$$

and show that L and R are constants of motion.

b) Show that $\mathfrak{l} := \mathrm{Span}_{\mathbb{R}}\{L_1, L_2, L_3\} \subset C^\infty(M)$ form a Lie-subalgebra of $C^\infty(M)$ i.e. $\{f, g\} \in \mathfrak{l}$ for all $f, g \in \mathfrak{l}$. Does the same hold for $\mathrm{Span}_{\mathbb{R}}\{R_1, R_2, R_3\}$?

c) Let $\xi(t) = (x(t), p(t))$ be a solution of Hamilton's equation for H with $R(\xi(t)) \neq 0$. Denote the angle between $x(t)$ and $R(\xi(t))$ by θ, i.e. $\cos(\theta(t)) := -\frac{\langle x(t), R(x(t))\rangle}{\|R(x(t))\|\|x(t)\|}$. Show that it holds

$$\|x(t)\| = \frac{1}{\alpha m} \frac{L_0^2}{1 + \frac{R_0}{mk}\cos(\theta(t))}, \tag{1.116}$$

where $L_0 := \|L(\xi(0))\|$ and $R_0 := \|R(\xi(0))\|$. What happens in the case $R(\xi(t)) = 0$?

Hint: It might help to introduce the Kronecker delta δ_{ij} and the Levi-Civita symbol

$$\varepsilon_{jkl} = \begin{cases} +1, & \text{if } jkl \text{ is an even permutation} \\ -1, & \text{if } jkl \text{ is an even permutation} \\ 0, & \text{else} \end{cases},$$

which satisfies the relations $(a \times b)_j = \varepsilon_{jkl} a_k b_l$ for $a, b \in \mathbb{R}^3$ and $\sum_{k=1}^3 \varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$.

Exercise 1.6 (Integral formula of pullback measure). Let μ be a Borel probability measure on X_1 and $\varphi : X_1 \to X_2$ Borel measurable. Show the following (Proposition M.6):

a) $\varphi_*\mu$ is a Borel probability measure on X_2.

b) $\mathrm{supp}(\varphi_*\mu) \subset \overline{\varphi(\mathrm{supp}\,\mu)}$. And if ϕ is continuous, show that equality holds.

c) A measurable function $g : X_2 \to \overline{\mathbb{R}}$ is integrable w.r.t. $\varphi_*\mu$ if and only if $g \circ \varphi$ is integrable w.r.t. μ. In this case,

$$\int_{X_2} g \, d(\varphi_*\mu) = \int_{X_1} (g \circ \varphi) \, d\mu. \tag{1.117}$$

d) Show that

$$\varphi : [0,1] \to [0,1], \quad \varphi(x) := \begin{cases} 0, & x \notin \mathbb{Q} \\ x, & x \in \mathbb{Q} \end{cases} \tag{1.118}$$

is a Borel measurable function and that $\varphi_*\ell = \delta_0$ for the Lebesgue measure. Show that $\mathrm{supp}(\varphi_*\mu) \neq \overline{\varphi(\mathrm{supp}\,\mu)}$.

Exercise 1.7 (Correlated Gaussian measures). Consider the phase space $M = \mathbb{R}^2$ and a measure $d\mu = \rho_A(\xi)d\xi$ with density $\rho_A = e^{-\langle \xi, A\xi\rangle}$ for a symmetric positive-definite matrix $A \in \mathrm{M}_2(\mathbb{R})$ with $\det(A) = \pi^2$.

a) Show that μ is a probability measure.
b) Compute the probability measures $X_*\mu$ and $P_*\mu$ and their first and second moments.
c) Compute the probability measure for $\|X\|$ and $\|P\|$.
d) Compute the probability measure $L_*\mu$ for $L(x, p) = xp$ and its first and second moment.
e) Compute the covariances in μ of the following pairs

$$X, P \quad \text{and} \quad X, L \quad \text{and} \quad P, L. \tag{1.119}$$

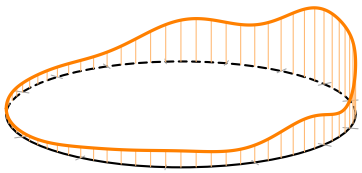
Hint: You may use that the function

$$Z(A) := \int_{\mathbb{R}^n} e^{-\xi^T A\xi}\, d\xi \tag{1.120}$$

on the set $\{A \in \mathrm{M}_n(\mathbb{R}) \ : \ \text{A symmetric and positive definite}\}$ is smooth and its derivatives w.r.t. the entries of A are given by

$$\frac{\partial Z}{\partial A_{ij}} = \begin{cases} -\int_{\mathbb{R}^n} 2\xi_i\xi_j\, e^{-\langle \xi, A\xi\rangle} d\xi, & i \neq j, \\ -\int_{\mathbb{R}^n} \xi_i\xi_j\, e^{-\langle \xi, A\xi\rangle} d\xi, & i = j. \end{cases}$$

Exercise 1.8 (Particle density on a circle). Consider a particle on a circle, i.e. the phase space $M = (-L, L) \times \mathbb{R}$, with $L > 0$ a length parameter, and periodic boundary conditions identifying L and $-L$. That is, all functions of position x are $2L$-periodic in this example.



The free Hamiltonian flow on the circle is

$$\phi_t^H(x, p) = \left(x + \frac{t}{m}p, p\right),$$

understood as a $2L$-periodic function in its first argument. We consider a phase space density $\rho \in C_c((-L, L) \times \mathbb{R})$ (also understood as a $2L$-periodic function in its first argument), i.e. $\rho(x, p) \geq 0$ for all (x, p), and $\int_{-L}^{L}\int_{\mathbb{R}} \rho(x, p)\, dx\, dp = 1$, and the phase space Borel probability measure $d\mu(x, p) = \rho(x, p)\, dx\, dp$.

Given an arbitrary observable $f \in C^\infty((-L, L) \times \mathbb{R})$ ($2L$-periodic in its first argument), discuss the long time asymptotics of $\omega_\mu(\alpha_t^H(f))$, and explain the result.
Hint: Expand $f(\cdot, p)$ and $\rho(\cdot, p)$ in Fourier series.

Exercise 1.9 (Probability distributions and dynamics). Let (M, H) be a Hamiltonian system and $f \in C^\infty_{\mathbb{R}}(M)$ be an observable. Show that for any $\mu \in \mathcal{P}(M)$ it holds

$$f_* \mu_t^H = (\alpha_t^H(f))_* \mu, \tag{1.121}$$

where $\mu_t^H := (\Phi_t^H)_* \mu$. Conclude that if f is a constant of motion then

$$f_* \mu_t^H = f_* \mu. \tag{1.122}$$

Exercise 1.10 (Convex combinations). Let V be a real vector space. A subset $C \subset V$ is called *convex* if

$$v_1, v_2 \in C, \lambda \in [0,1] \implies cv_1 + (1-c)v_2 \in C.$$

Show that a convex set $C \subset V$ satisfies for all $n \in \mathbb{N}$

$$v_1, \dots, v_n \in C, \lambda_1, \dots, \lambda_n \geq 0, \sum_{k=1}^n \lambda_k = 1 \implies \sum_{k=1}^n \lambda_k v_k \in C.$$

Exercise 1.11 (Alice, Bob, and classical correlations). (See Example 1.40.)

Consider the phase space $M = \mathbb{R}^{2d}$ and let $U \subset \mathbb{R}^d$ be an open subset. Define

$$\mathcal{A}(U) = \{f \in C^\infty(M) \mid f(x,p) = 0 \text{ if } x \notin U\} + \mathbb{C}1 \tag{1.123}$$

and show that $\mathcal{A}(U) \subset C^\infty(M)$ is a subsystem.

Let now $U_A, U_B \subset \mathbb{R}^d$ be disjoint open bounded subsets and consider the subsystems $\mathcal{A} := \mathcal{A}(U_A)$ and $\mathcal{B} := \mathcal{A}(U_B)$ and the bipartite system $(\mathcal{A}, \mathcal{B})$.

a) For $\mu \in \mathcal{P}_c(M)$ characterize whether ω_μ is uncorrelated in $(\mathcal{A}, \mathcal{B})$ in terms of $\operatorname{supp} \mu$. Show that there exist states ω_μ that are correlated in $(\mathcal{A}, \mathcal{B})$.

b) Let ω_μ be an $(\mathcal{A}, \mathcal{B})$-correlated state. Let $\epsilon > 0$ and show there exist $\chi_A \in \mathcal{A}$ s.t.

$$\omega_\mu(\chi_A) > 1 - \epsilon \implies \Pr_\mu(X \in U_A) > 1 - \epsilon \quad \text{and} \quad \Pr_\mu(X \in U_B) < \epsilon. \tag{1.124}$$

Interpret this result.

c) Let ω_μ be a correlated state over $(\mathcal{A}, \mathcal{B})$ and choose $\chi_A \in \mathcal{A}$ and $\chi_B \in \mathcal{B}$ as in $b)$. Show that

$$\operatorname{Cov}_\mu(\chi_A, \chi_B) < 0 \quad \text{and} \quad \operatorname{Corr}_\mu(\chi_A, \chi_B) < 0 \quad \text{if it exists.}$$

Interpret this result.

Problem 1.1 (Particle bouncing off a wall). In $d = 1$ dimension, consider the poten-

tial

$$V : (0,\infty) \to \mathbb{R}, \qquad V(x) := \frac{1}{x^2} - \frac{1}{x}. \tag{1.125}$$

Sketch the graph of V and then prove the following for an arbitrary solution $\mathbb{R} \ni t \mapsto x(t) \in \mathbb{R}$ of Newton's equation with energy E.

a) $-\frac{1}{4} \leq E < \infty$.
b) For $E < 0$, the solution is bounded.
c) For $E > 0$, the solution is unbounded.
d) What happens for $E = 0$?
e) In all cases, discuss $\lim_{t\to\infty} x(t)$.

It might be useful to investigate points with vanishing momentum. It is not necessary to solve Newton's or Hamilton's equations.

Problem 1.2 (Harmonic oscillator). In $d = 1$ dimension, consider the potential

$$V : \mathbb{R} \to \mathbb{R}, \qquad V(x) := \frac{k}{2} x^2, \tag{1.126}$$

where $k > 0$ is a parameter. Explicitly compute the flow $\phi^H : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$, $(t, x_0, p_0) \mapsto (x(t), p(t)) = \xi(t)$ of the Hamiltonian vector field

$$\mathcal{V}_H(x,p) = \begin{pmatrix} m^{-1}p \\ -\nabla_x V \end{pmatrix}.$$

Problem 1.3 (Properties of the Poisson bracket). The Poisson bracket is defined as

$$\{\cdot,\cdot\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M), \qquad \{f,g\} := \langle \nabla \bar{f}, \Sigma \nabla g \rangle. \tag{1.127}$$

Prove the properties of the Poisson bracket stated in Proposition 1.15 and Theorem 1.16 (1.46).

Problem 1.4 (Dirac measures on $C^\infty(M)$). Let $\mu \in \mathcal{P}(M)$. Show that the following are equivalent:

a) $\mu = \delta_\xi$ is a Dirac measure for some $\xi \in M$.
b) $\Delta_\mu(f) = 0$ for every $f \in C_c^\infty(M)$.

Problem 1.5 (Liouville's Theorem). Let (M, H) be a Hamiltonian system. Show that $\det D\phi_t^H = 1$ for all $t \in \mathbb{R}$ for which the flow is defined.

Problem 1.6 (Center of $C^\infty(M)$). Let A be in the center of $C^\infty(M)$ w.r.t. the Poisson bracket, i.e. $\{A, f\} = 0$ for all $f \in C^\infty(M)$. Show that A is a constant function.

Hint: Consider the Hamiltonian system (M, A).

Problem 1.7 (Measure and densities). Let (M, H) be a Hamiltonian system. Consider $\mu \in \mathcal{P}(M)$ given by the density $\rho \in L^1(M, d\xi)$ (w.r.t. the Lebesgue measure $d\xi$). Show that the measure $(\phi_t^H)_*\mu$ has density $(\phi_t^H)^*\rho$.

Let now $\rho(\xi) = F(H(\xi))$ for a smooth compactly supported function $F : \mathbb{R} \to [0, \infty)$. Show that $(\phi_t^H)_*\mu = \mu$ for all $t \in \mathbb{R}$ s.t. the flow ϕ_t^H exists.

Problem 1.8 (Tunnel effect?). Consider the Hamilton function

$$H : \mathbb{R}^2 \to \mathbb{R}, \quad H(x, p) = \frac{p^2}{2m} + V(x),$$

where the potential V is a non-negative smooth function with compact support and global maximum $V(0) = V_0 > 0$ with $V''(0) < 0$. Consider $\mu \in \mathcal{P}(\mathbb{R}^2)$ with density ρ, satisfying $\rho(x, p) = 0$ for all $(x, p) \in \mathbb{R}^2$ with $H(x, p) \geq V_0$. Show that

$$\Pr_{(\phi_t^H)_*\mu}(X \in (0, \infty)) = ((\phi_t^H)_*\mu)(\{X > 0\}) \tag{1.128}$$

is independent of t.

Problem 1.9 (Pure and mixed states on $C^\infty(M)$ (2)). Prove the remaining parts of Theorem 1.35 a).

Problem 1.10 (Classical Ehrenfest Theorem). Consider the Hamilton function

$$H_N : \mathbb{R}^2 \to \mathbb{R}, \quad H_N(x, p) = \frac{p^2}{2m} + V_N(x), \quad V_N(x) := x^N, \quad N \in \mathbb{N}_0.$$

Let $\mu \in \mathcal{P}_c(\mathbb{R}^2)$ and denote $F_N(x) := -(\nabla V_N)(x)$.

a) Using the algebraic properties of the Poisson bracket derive

$$\frac{d}{dt}\alpha_t^H(P) = \alpha_t^H(F_N(X)) \quad \text{and} \quad \frac{d}{dt}\alpha_t^H(X) = \frac{1}{m}\alpha_t^H(P).$$

b) Show that the functions

$$f_X^\mu(t) := \omega_\mu(\alpha_t^H(X)) \quad \text{and} \quad f_P^\mu(t) := \omega_\mu(\alpha_t^H(P)) \tag{1.129}$$

are differentiable with derivatives

$$\frac{d}{dt}f_X^\mu(t) = \omega_\mu(\alpha_t^H(P)) \quad \text{and} \quad \frac{d}{dt}f_P^\mu(t) = \omega_\mu(\alpha_t^H(F_N(X))). \tag{1.130}$$

c) Show that for $N \in \{0, 1, 2\}$

$$f : \mathbb{R} \to \mathbb{R}^2, \quad f(t) = (f_X^\mu(t), f_P^\mu(t)) \tag{1.131}$$

solves Hamilton's equation. Why does the same not happen for $N \geq 3$ for general $\mu \in \mathcal{P}_c(\mathbb{R}^2)$?

Problem 1.11 (Harmonically coupled particles). Let $U, V \subset \mathbb{R}^n$ be open sets. A diffeomorphism $\phi : U \times \mathbb{R}^n \to V \times \mathbb{R}^n$ is called canonical transformation if

$$(D\phi)\Sigma(D\Phi)^T = \Sigma$$

for the canonical symplectic form Σ.

Consider the Hamiltonian function

$$H : \mathbb{R}^4 \to \mathbb{R}, \quad H = \frac{1}{2m}(P_1^2 + P_2^2) + \frac{k}{2}(X_1 - X_2)^2$$

and the state $d\mu(x_1, x_2, p_1, p_2) = \rho(x_1)\rho(x_2)\, dx_1\, dx_2\, d\delta_0(p_1)\, d\delta_0(p_2)$ with the density ρ satisfying $\rho \geq 0$, $\int_{\mathbb{R}} \rho = 1$, $\int_{\mathbb{R}} x\rho(x)dx = 0$, $\int_{\mathbb{R}} x^2\rho(x)dx > 0$.

a) Consider $X_\pm(x_1, x_2, p_1, p_2) := \frac{1}{\sqrt{2}}(x_1 \pm x_2)$ and $P_\pm(x_1, x_2, p_1, p_2) := \frac{1}{\sqrt{2}}(p_1 \pm p_2)$ and show that

$$\phi : \mathbb{R}^4 \to \mathbb{R}^4, \quad \phi = (X_+, X_-, P_+, P_-) \tag{1.132}$$

is a canonical transformation. Show moreover that any canonical transformation satisfies

$$\{\phi^*(f), \phi^*(g)\} = \phi^*(\{f, g\}), \quad f, g \in C^\infty(V \times \mathbb{R}^n). \tag{1.133}$$

b) Show that $(\mathbb{R}^4, H \circ \phi)$ is the uncoupled product of $(\mathbb{R}^2, H_0)$ and $(\mathbb{R}^2, H_1)$, where

$$H_0(x, p) = \frac{p^2}{2m} \quad \text{and} \quad H_1(x, p) = \frac{p^2}{2m} + \frac{k}{2}x^2. \tag{1.134}$$

c) Show that

$$\mathrm{Corr}_\mu(\alpha_t^H(X_1)\alpha_t^H(X_2)) = \frac{\sin^2(\nu t)}{2}, \qquad \nu := \sqrt{\frac{k}{m}}. \tag{1.135}$$

2 Finite quantum mechanical systems (matrix mechanics)

In the late 19th and early 20th century, various problems with classical physics were noted: For some experimental observations, like the black body radiation or the photoelectric effect, classical physics gave wrong predictions. Other phenomena like the stability of atoms or discrete lines in atomic absorption spectra could also not be explained with classical physics, and the search for a new theory going beyond classical mechanics began. This theory was coined "quantum mechanics" because several observable quantities that could vary continuously in classical mechanics occured only in discrete smallest portions or quanta in the new theory.

As in classical mechanics (and as in any theory of physics), also quantum mechanics talks about states and observables, modelling the states a given system can be in and its observable (measurable) quantities. Given a state ω and an observable A, one is still interested in describing the expectation value $\omega(A)$ and uncertainty $\Delta_\omega(A)$ of A in the state ω, the probability $\Pr_\omega(A \in S)$ that the value of A lies in some (Borel) subset $S \subset \mathbb{R}$, etc. When a dynamics is specified, one also still wants to describe how these quantities change with time.

In contrast to classical mechanics, states, observables and dynamics will no longer be given by probability measures and smooth functions on phase space, but rather take a quite different form that however also share structural similarities. We will therefore stick to the familiar symbol ω for states, A for observables (and e.g. X for position, P for momentum, H for Hamiltonian, L for angular momentum), and $\omega(A)$ for the expectation value of an observable A in a state ω. Typographically this looks exactly like in classical mechanics, but the mathematical objects ω, A, etc are of a different nature to be specified. From context, it will however always be clear what is meant.

The passage from classical mechanics to quantum mechanics can of course only be accomplished by comparison with experimental findings – to find the "right" mathematical model for quantum mechanics means to find a model that is build in such a way that its predictions match the experimental findings. It is therefore impossible to give mathematical proofs that a given framework is "the right quantum mechanics". In the mathematical approach taken here, we will take as a guiding principle the so-called *Heisenberg uncertainty relation*, according to which the uncertainties of position and momentum (say, in one dimension for simplicity) satisfy a bound of the form

$$\Delta_\omega(X) \cdot \Delta_\omega(P) \geq \frac{\hbar}{2} > 0, \tag{2.1}$$

where $\hbar \cong 10^{-34}$ Js is a constant of nature (called *Planck's constant*), a small[7] but positive number. The remarkable feature is that this inequality should hold in *every state* ω, and the bound on the right hand side is independent of ω.

A mathematical setting in which such an inequality holds can therefore in particular not admit any zero uncertainty states. Clearly classical mechanics is not compatible with (2.1).

In this chapter, we will first briefly look at a vast generalization of the setting of classical mechanics to general unital *-algebras, replacing the specific unital *-algebra $\mathcal{A} = C^\infty(M)$ of classical mechanics. As explained in Section 2.1, this setting naturally leads to uncertainty

[7] For dimensionful quantities like $\hbar$, the number value depends on the choice of physical units and hence "small" has only meaning in comparison to other quantities of the same physical dimension. $\hbar$ has the units of action, and is small in comparison to actions that we are used to from everyday experience.

relations. The rest of the chapter will then be spent by looking at a simple but instructive example, the algebra $M_N(\mathbb{C})$ of complex $(N \times N)$-matrices: In Section 2.2 we study the states of $M_N(\mathbb{C})$ and learn how they connect probability measures with eigenvalue problems, and in Section 2.3 we learn what the right substitute for Poisson brackets and Hamiltonians are in this context. In particular, we will find Heisenberg's and Schrödinger's Equations as the quantum replacements for Hamilton's Equation. Section 2.4 introduces spin, partially in analogy to angular momentum in classical mechanics, and in Section 2.5, we look at composite systems and find that for states on matrix algebras, entanglement does exist.

While this "matrix mechanics" is insufficient to fully describe quantum mechanics, it provides a simplified and instructive framework for learning the fundamentals of quantum mechanics before moving on to a deeper and more general analysis in the next chapter.

2.1 *-Algebras and their state spaces

The basic concept in the following is that of a unital *-algebra.

Definition 2.1 (*-algebra). A *-algebra is a complex associative algebra $\mathcal{A}$ together with an antilinear involution $A \mapsto A^*$ such that $(AB)^* = B^*A^*$ for all $A, B \in \mathcal{A}$. If $\mathcal{A}$ contains a unit element 1 for multiplication, $\mathcal{A}$ is called *unital*.

We emphasize that $\mathcal{A}$ is *not* required to be commutative, i.e. in general the *commutator*

$$[A, B] := AB - BA \tag{2.2}$$

of two elements $A, B \in \mathcal{A}$ will not vanish. We will extensively work with noncommutative algebras in quantum mechanics. Mathematically speaking, the transition from the commutative algebra $C^\infty(M)$ to a noncommutative one is the most important difference between classical and quantum mechanics.

Example 2.2.

a) Let $\mathcal{H}$ be a complex Hilbert space. Then the algebra of all bounded operators, $\mathcal{A} := B(\mathcal{H})$, is a unital *-algebra with $1 = \mathrm{id}_\mathcal{H}$ and A^* the adjoint of A, i.e.

$$\langle \psi, A\phi \rangle = \langle A^*\psi, \phi \rangle, \qquad \psi, \phi \in \mathcal{H}. \tag{2.3}$$

b) For $\mathcal{H} = \mathbb{C}^N$, the first example is isomorphic to $B(\mathcal{H}) \cong M_N(\mathbb{C})$, the unital *-algebra of complex $(N \times N)$-matrices. Here the unit 1 is the $(N \times N)$ identity matrix, and the *-operation is conjugate transposition, $(A^*)_{kl} = \overline{A_{lk}}$.

c) As in the previous chapter, $C^\infty(M)$ for a phase space M, or $C(X)$ for a topological space X, are unital *-algebras with the pointwise operations.

In the case of $\mathcal{A} = C^\infty(M)$, the actual observables were the real functions $f \in C^\infty_\mathbb{R}(M)$, characterized by $f = f^*$. This will be the same in general unital *-algebras, and we introduce some standard vocabulary now.

Definition 2.3. Let $\mathcal{A}$ be a unital *-algebra, and $A \in \mathcal{A}$.

a) A is called *selfadjoint* iff $A = A^*$.

b) A is called *unitary* iff A is invertible and $A^{-1} = A^*$.

c) A is called (*orthogonal*) *projection* iff $A = A^*$ and $A^2 = A$.

d) A is called *partial isometry* iff A^*A is an orthogonal projection.
e) A is called *normal* iff $AA^* = A^*A$.
f) A is called *trivial* iff $A = c1, c \in \mathbb{C}$.

The commutator $[\cdot,\cdot]$ of a *-algebra has properties analogous to that of the Poisson bracket of $C^\infty(M)$ (see Proposition 1.15, also for the definition of the Jacobi Identity and the Leibniz rule):

Lemma 2.4 (Properties of the commutator). *Let $\mathcal{A}$ be a *-algebra. Then*

$$-i[\cdot,\cdot] : \mathcal{A} \times \mathcal{A} \to \mathcal{A}, \qquad (A,B) \mapsto -i[A,B] \tag{2.4}$$

a) is bilinear and antisymmetric,
b) satisfies the Jacobi Identity,
c) satisfies the Leibniz rule,
d) respects the *-operation: $(-i[A,B])^* = -i[A^*,B^*], A,B \in \mathcal{A}$.

The simple proof is carried out in the exercises. We have here rescaled the commutator by a factor $-i$ in order to get compatability with the *-operation. Also any other purely imaginary factor would do this, but $-i$ will turn out to be convenient later. The first three properties listed in the lemma also hold without the prefactor $-i$.

For $\mathcal{A} = C^\infty(M)$, the expectation value functions (1.73) induced by states on M were a crucial aspect of the theory. In Theorem 1.42 we have seen that they can be identified with the positive linear normalized functionals on $C^\infty(M)$. This suggests the following definition.

Definition 2.5 (States on unital *-algebras). Let $\mathcal{A}$ be a unital *-algebra. A *state* is a linear functional

$$\omega : \mathcal{A} \to \mathbb{C} \tag{2.5}$$

that is *normalized* in the sense $\omega(1) = 1$ and *positive*[a] in the sense $\omega(A^*A) \geq 0$ for all $A \in \mathcal{A}$. The set of all states of $\mathcal{A}$ is denoted $S(\mathcal{A})$.

[a] This purely algebraic definition of positivity is not adequate for all *-algebras, but will be the right choice for the ones that we consider in these lectures.

While states on a general *-algebra are defined completely independently of probability measures, still various concepts that we learned in classical mechanics carry over immediately:

As in the special case of $\mathcal{A} = C^\infty(M)$, one sees directly that the state space $S(\mathcal{A})$ of a unital *-algebra is a convex set. Hence also the concept of pure and mixed states carries over.

Definition 2.6 (Pure and mixed states on unital *-algebras). Let $\mathcal{A}$ be a unital *-algebra. A state $\omega \in S(\mathcal{A})$ is called *pure* if it cannot be written as a non-trivial convex combination of other states, i.e. if $\omega = c\omega_1 + (1-c)\omega_2$ with $\omega_1, \omega_2 \in S(\mathcal{A})$ and $0 \leq c \leq 1$ implies either $c \in \{0,1\}$ or $\omega_1 = \omega_2$. The set of all pure states is denoted $S_{\text{pure}}(\mathcal{A}) \subset S(\mathcal{A})$. A state that is not pure is called *mixed*.

Also the definitions of uncertainties, correlations and covariances of selfadjoint elements transfer to the general setting without changes.

Definition 2.7 (Uncertainties, covariances and correlations). Let $\omega \in S(\mathcal{A})$ be a state on a unital *-algebra, and $A = A^*, B = B^* \in \mathcal{A}$.

a) The *uncertainty* of A in ω is

$$\Delta_\omega(A) := \sqrt{\omega(A^2) - \omega(A)^2} = \sqrt{\omega((A - \omega(A)1)^2)}. \tag{2.6}$$

b) The *covariance of* A, B *in* ω is

$$\mathrm{Cov}_\omega(A, B) := \omega(AB) - \omega(A)\omega(B). \tag{2.7}$$

c) If $\Delta_\omega(A) \neq 0$ and $\Delta_\omega(B) \neq 0$, the *correlation of* A, B *in* ω is defined as

$$\mathrm{Cov}_\omega(A, B) := \frac{\mathrm{Cov}_\omega(A, B)}{\Delta_\omega(A)\Delta_\omega(B)} = \frac{\omega(AB) - \omega(A)\omega(B)}{\Delta_\omega(A)\Delta_\omega(B)}. \tag{2.8}$$

Two observables are called *uncorrelated* if $\mathrm{Cov}_\omega(A, B) = 0$.

These formulae are essentially identical to the ones from classical mechanics, up to the notational difference that we now index uncertainties, covariances and correlations with the state ω instead of the phase space measure μ because there is no phase space measure here.

The name uncertainty suggests that this quantity measures a kind a statistical deviation from the mean value. We will see later that this is indeed the case. However, at first sight it is not even clear that uncertainties as defined here are non-negative. To show that we need the Cauchy-Schwarz Inequality.

Proposition 2.8 (General Cauchy-Schwarz Inequality). *Let $\mathcal{A}$ be a unital *-algebra, $\omega \in S(\mathcal{A})$ a state, and $A, B \in \mathcal{A}$. The map*

$$s : \mathcal{A} \times \mathcal{A} \to \mathbb{C}, \qquad s(A, B) := \omega(A^*B) \tag{2.9}$$

is a positive semi-definite sesquilinear form on $\mathcal{A}$. In particular, the general Cauchy-Schwarz Inequality *holds,*

$$|\omega(A^*B)|^2 \leq \omega(A^*A)\omega(B^*B), \tag{2.10}$$

holds, and ω is hermitian in the sense

$$\omega(A^*) = \overline{\omega(A)}. \tag{2.11}$$

Proof. It is clear that s is sesquilinear and positive semi-definite. Let $A, B \in \mathcal{A}$ and $\lambda, \mu \in \mathbb{C}$. Then the positivity of s implies

$$\begin{aligned} 0 &\leq \omega((\lambda A + \mu B)^*(\lambda A + \mu B)) \\ &= |\lambda|^2\omega(A^*A) + \overline{\lambda}\mu\omega(A^*B) + \lambda\overline{\mu}\omega(B^*A) + |\mu|^2\omega(B^*B) \\ &= \left\langle \begin{pmatrix}\lambda\\ \mu\end{pmatrix}, M \begin{pmatrix}\lambda\\ \mu\end{pmatrix} \right\rangle, \qquad M := \begin{pmatrix}\omega(A^*A) & \omega(A^*B)\\ \omega(B^*A) & \omega(B^*B)\end{pmatrix}, \end{aligned}$$

with $\langle \cdot, \cdot \rangle$ the canonical scalar product on $\mathbb{C}^2$. This shows that the matrix M is positive semidefinite. In particular, $M = M^*$ is selfadjoint, which implies (2.11) (set $B = 1$). Fur-

thermore, M has non-negative determinant,

$$0 \leq \det M = \omega(A^*A)\omega(B^*B) - \omega(A^*B)\omega(B^*A) = \omega(A^*A)\omega(B^*B) - |\omega(A^*B)|^2.$$

This shows the Cauchy-Schwarz Inequality (2.10). □

As a consequence of this proposition, we see that also in the general setting of unital *-algebras, selfadjoint elements $A = A^*$ have *real* expectation values and *non-negative* uncertainties: For $A = A^*$, we have $(A - \omega(A)1)^* = A - \omega(A)1$, and therefore

$$\Delta_\omega(A) = \omega((A - \omega(A)1)^*(A - \omega(A)1))^{1/2} \geq 0.$$

What is also true in the general setting is the shift invariance of the uncertainty,

$$\Delta_\omega(A - c1) = \Delta_\omega(A), \qquad A \in \mathcal{A}, c \in \mathbb{C}, \omega \in S(\mathcal{A}). \tag{2.12}$$

Another consequence of the Cauchy-Schwarz Inequality is that the correlation coefficient still lies in $[-1, 1]$:

$$\begin{aligned} |\operatorname{Cov}_\omega(A, B)| &= |\omega((A - \omega(A))^*(B - \omega(B)))| \\ &\leq \left(\omega(|A - \omega(A)|^2)\right)^{1/2} \left(\omega(|B - \omega(B)|^2)\right)^{1/2} = \Delta_\omega(A)\Delta_\omega(B). \end{aligned}$$

The most important consequence is however the uncertainty relation.

Proposition 2.9 (General uncertainty relation). *Let $\mathcal{A}$ be a unital *-algebra, $\omega \in S(\mathcal{A})$ a state, and $A = A^*, B = B^* \in \mathcal{A}$ two selfadjoint elements. Then the* general uncertainty relation

$$\Delta_\omega(A) \cdot \Delta_\omega(B) \geq \frac{1}{2}|\omega([A, B])| \tag{2.13}$$

holds.

Proof. Both the left and right hand side of (2.13) does not change when shifting A and B by multiples of the identity, hence we may assume $\omega(A) = \omega(B) = 0$ without loss of generality. Then

$$\Delta_\omega(A)^2\Delta_\omega(B)^2 = \omega(A^2)\omega(B^2) = \omega(A^*A)\omega(B^*B) \geq |\omega(A^*B)|^2 = |\omega(AB)|^2,$$

where the inequality follows from the Cauchy-Schwarz Inequality. Using that $|z| \geq |\operatorname{Im} z|$ and $\operatorname{Im} z = \frac{1}{2i}(z - \overline{z})$ (for $z \in \mathbb{C}$) as well as $\overline{\omega(AB)} = \omega(BA)$ (2.11), one concludes

$$|\omega(AB)|^2 \geq |\frac{1}{2i}\omega(AB - BA)|^2 = \frac{1}{4}|\omega([A, B])|^2.$$

The uncertainty relation follows by taking the square root. □

This proposition tells us that in any noncommutative algebra $\mathcal{A}$, there are lots of uncertainty relations. Whenever A, B are two non-commuting elements, there is a lower bound on the product of their uncertainties. If $\mathcal{A}$ has sufficiently many states so that for each non-zero element $C \in \mathcal{A}$ there exists a state ω with $\omega(C) \neq 0$, then we also see that $S(\mathcal{A})$ does not contain any zero uncertainty states as in the classical case.

Regarding the more specific Heisenberg Uncertainty Relation (2.1), we see that a unital *-algebra will realize it once we find elements $X, P \in \mathcal{A}$ such that $[X, P] = c\hbar 1$ with $c \in \mathbb{C}$, $|c| = 1$. We will come back to these so-called canonical commutation relations later.

2.2 The state space of matrix mechanics

In this section we study the states of the specific finite-dimensional unital *-algebra $\mathcal{A} = M_N(\mathbb{C})$ (Example 2.2 b)) of $(N \times N)$-matrices (with $N \in \mathbb{N}$ fixed in the following). On the one hand, this is mathematically quite simple and only involves linear algebra. On the other hand, this example is already rich enough to illustrate various phenomena and furthermore of direct relevance to certain physical systems (spin and lattice systems).

We refer to Appendix M.2 for some reminders about linear algebra and our notation. For example, for our first definition we rely on the concept of a positive matrix (Prop. M.10) and the matrix trace (M.33), see Problems 2.1 and 2.2 for some exercises relating to these concepts.

Definition 2.10 (Density matrices). A *density matrix in* $M_N(\mathbb{C})$ is a selfadjoint positive matrix $\rho \in M_N(\mathbb{C})$, $\rho \geq 0$, such that $\operatorname{Tr}\rho = 1$. The set of all density matrices in $M_N(\mathbb{C})$ is denoted $\mathcal{D}_N$.

The significance of density matrices is that they parameterize all states on $M_N(\mathbb{C})$.

Proposition 2.11 (The state space of $M_N(\mathbb{C})$). *Let $N \in \mathbb{N}$. The map*

$$\mathcal{D}_N \to S(M_N(\mathbb{C})), \qquad \rho \mapsto \omega_\rho, \qquad \omega_\rho(A) := \operatorname{Tr}(\rho A), \tag{2.14}$$

is a bijection.

Proof. We first check that ω_ρ is a state on $M_N(\mathbb{C})$: Indeed, $A \mapsto \operatorname{Tr}(\rho A)$ is linear and normalized because $\omega_\rho(1) = \operatorname{Tr}\rho = 1$. It is positive because for $A = B^*B \geq 0$,

$$\omega_\rho(A) = \operatorname{Tr}(\rho B^*B) = \operatorname{Tr}\big(B\sqrt{\rho}\sqrt{\rho}B^*\big) = \operatorname{Tr}\big(|\sqrt{\rho}B^*|^2\big) \geq 0,$$

where we have used the positivity of the trace. So for any $\rho \in \mathcal{D}_N$, the functional $\omega_\rho \in S(M_N(\mathbb{C}))$ is a state of $M_N(\mathbb{C})$.

To understand that $\rho \mapsto \omega_\rho$ is a bijection, we have to show that this map is surjective (every state ω is of the form $\omega = \omega_\rho$ for some $\rho \in \mathcal{D}_N$) and injective (different density matrices define different states). To see that, we consider the canonical matrix units $E_{kl} = |e_k\rangle\langle e_l| \in M_N(\mathbb{C})$ (Def. M.11). Their expectation values are

$$\omega_\rho(E_{kl}) = \operatorname{Tr}(\rho E_{kl}) = \sum_{j=1}^{N} \langle e_j, \rho E_{kl} e_j\rangle = \langle e_l, \rho e_k\rangle = \rho_{lk},$$

the matrix entries of ρ. From this we learn that different density matrices define different states (if $\rho \neq \rho'$, they differ in at least one entry kl, and then the expectation values $\omega_\rho(E_{kl}) \neq \omega_{\rho'}(E_{kl})$ differ). We also learn that every state is of density matrix form: Given $\omega \in S(M_N(\mathbb{C}))$, *define* $\rho \in M_N(\mathbb{C})$ by $\rho_{lk} := \omega(E_{kl})$. Then

$$\omega(A) = \sum_{k,l} A_{kl}\omega(E_{kl}) = \sum_{k,l} A_{kl}\rho_{lk} = \sum_l (\rho A)_{ll} = \operatorname{Tr}(\rho A) = \omega_\rho(A).$$

It remains to check that the ρ defined by $\rho_{lk} := \omega(E_{kl})$ is a density matrix. We see that $\operatorname{Tr}\rho = \sum_l \rho_{ll} = \sum_l \omega(E_{ll}) = \omega(1) = 1$, so ρ has unit trace. To check that it is positive, let

$\psi \in \mathbb{C}^N$ and consider the rank one projection onto $\mathbb{C}\psi$, i.e. $P_\psi = |\psi\rangle\langle\psi|$. Then $P_\psi \geq 0$, so

$$0 \leq \omega(P_\psi) = \mathrm{Tr}\big(\rho P_\psi\big) = \sum_l \langle e_l, \rho\langle\psi, e_l\rangle\psi\rangle = \langle\psi, \rho\psi\rangle.$$

Thus $\rho \geq 0$. □

This bijection between states of $M_N(\mathbb{C})$ and density matrices replaces the bijection between states of the classical observable algebra $C^\infty(M)$ and compactly supported Borel probability measures on M (Theorem 1.42). Hence we may view density matrices as the matrix analogues of probability measures. In classical mechanics, we may index states by measures $\mu \in \mathcal{P}_c(M)$ and hence write $\omega = \omega_\mu$, Δ_μ, etc. In matrix mechanics, we instead write $\omega = \omega_\rho$, Δ_ρ, etc., i.e. we use density matrices $\rho \in \mathscr{D}_N$ to label the states.

To explain the relation between probability measures and density matrices further, we compare finite measure spaces with diagonal matrices. Consider the finite set $X_N = \{1, \dots, N\}$ with σ-algebra 2^{X_N} (its power set). A natural measure on X_N is the counting measure,

$$\mu_{\mathrm{count}} := \delta_1 + \dots + \delta_N, \qquad \mu_{\mathrm{count}}(S) = |S|, \qquad S \subset \{1, \dots, N\}. \tag{2.15}$$

Here δ_k is the Dirac measure supported at $k \in X$, and $|S|$ is the cardinality of S (number of elements in S). All functions $f : X_N \to \mathbb{C}$ are μ_{count}-integrable, with integral

$$\int_{X_N} f \, d\mu_{\mathrm{count}} = f(1) + \dots + f(N). \tag{2.16}$$

Lemma 2.12 (Finite measure spaces and diagonal density matrices).

a) There is a bijection between functions $f : X_N \to \mathbb{C}$ and diagonal matrices in $M_N(\mathbb{C})$ given by $f \mapsto D(f)$, $D(f)_{kl} := \delta_{kl} f(k)$.

b) All probability measures on X_N are of the form $r\mu_{\mathrm{count}}$ with normalized densities $r \geq 0$, $\int_{X_N} r \, d\mu_{\mathrm{count}} = 1$. There is a bijection between probability measures on X_N and diagonal density matrices $\rho \in M_N(\mathbb{C})$, given by $r\mu_{\mathrm{count}} \mapsto D(r) =: \rho$.

c) For $f : X_N \to \mathbb{C}$ and a density r, one has

$$\int_{X_N} f \, d(r\mu_{\mathrm{count}}) = \omega_{D(r)}(D(f)). \tag{2.17}$$

d) Let r be a density and $\rho := D(r)$ the corresponding density matrix. Then

$$D(\chi_{\mathrm{supp}\, r}) = P_{\mathrm{Ran}\, \rho} = P^\perp_{\ker \rho}. \tag{2.18}$$

Proof. a) is clear. For b), we note that a measure ν on the power set of $\{1, \dots, N\}$ is given by assigning non-negative numbers $r(k) := \nu(\{k\})$ to $k \in \{1, \dots, N\}$, and ν is a probability measure if $1 = \nu(X_N) = \sum_k r(k)$. The bijection with the diagonal matrices in $\mathscr{D}_N$ is now clear. For c), we compute

$$\int_{X_N} f \, d(r\mu_{\mathrm{count}}) = \sum_{k=1}^N f(k)(r\mu_{\mathrm{count}})(\{k\}) = \sum_{k=1}^N f(k) r(k) = \mathrm{Tr}(D(r)D(f)) = \omega_{D(r)}(D(f)).$$

d) Let r be a density. Then $P := D(\chi_{\operatorname{supp} r})$ is the diagonal matrix with diagonal entries only 0 and 1, with $P_{kk} = 1$ if and only if $k \in \operatorname{supp} r$, and $\rho := D(r)$ is a diagonal matrix with $\rho_{kk} > 0$ if and only if $k \in \operatorname{supp} r$. This implies $P^{\perp}_{\ker \rho} = P$, and $P^{\perp}_{\ker \rho} = P_{\operatorname{Ran} \rho}$ holds for any selfadjoint matrix. □

As the bijection in the last part maps the density r to the density matrix $D(r)$, this explains the name "density matrix". Some important remarks are:

- In this example, we started from the special measure μ_{count}, corresponding to the trace Tr on the matrix algebra[8]. This has no analogue on our classical observable algebra $C^\infty(M)$ because we considered unbounded phase spaces. For compact $M \subset \mathbb{R}^{2d}$, Lebesgue measure ℓ_M would be the right analogue of the trace, and normalized Lebesgue measure $\ell_M/\ell_M(M)$ the right analogue of the normalized trace N^{-1} Tr.
- The direct correspondence between densities and density matrices explained above works for *diagonal* matrices. While any normal matrix (in particular, any selfadjoint matrix) can be diagonalized, two noncommuting normal matrices can not be diagonalized simultaneously, so that $\omega_\rho(A)$ can not be brought into the simple form $\sum_k \rho_{kk} A_{kk}$ in case ρ and A do not commute. There is no bijection between probability measures on X_N and the set $\mathscr{D}_N$ of all density matrices.

The last item d) allows us to capture the support of a probability measure in matrix language:

Definition 2.13 (Support projections of density matrices). The *support projection of a density matrix* $\rho \in \mathscr{D}_N$ is the orthogonal projection $P_{\operatorname{Ran} \rho} = P^{\perp}_{\ker \rho}$.

In classical mechanics, the states given by Dirac measures had many special features relating to uncertainties, correlations, purity, and minimal support (Theorem 1.35). We now want to to study these properties analogously for states on $M_N(\mathbb{C})$, and begin by comparing purity and support projections.

Theorem 2.14 (The pure states of $M_N(\mathbb{C})$). *Let $\omega \in S(M_N(\mathbb{C}))$ be a state with density matrix ρ. The following are equivalent:*

a) ω is pure.
b) ρ is a rank one projection.
c) ω is a vector state, i.e. there exists $\psi \in \mathbb{C}^N$, $\|\psi\| = 1$, such that

$$\omega(A) = \langle \psi, A\psi \rangle, \qquad A \in M_N(\mathbb{C}). \tag{2.19}$$

Every state $\omega \in S(M_N(\mathbb{C}))$ is a convex combination of pure states.

Proof. The bijection $\mathscr{D}_N \to S(M_N(\mathbb{C}))$ established in Proposition 2.11 respects convex combinations: $\omega_{c\rho_1 + (1-c)\rho_2} = c\omega_{\rho_1} + (1-c)\omega_{\rho_2}$, $0 \le c \le 1$. Thus ω_ρ is an extreme point of $S(M_N(\mathbb{C}))$ if and only if ρ is an extreme point of $\mathscr{D}_N$. We have to show that the extreme points of $\mathscr{D}_N$ are the rank one projections.

Let $\rho \in \mathscr{D}_N$ and $(b_k)_k$ an orthonormal basis of $\mathbb{C}^N$ in which ρ is diagonal (with eigenvalues λ_k). Then $\rho = \sum_k \lambda_k |b_k\rangle\langle b_k|$ gives a non-trivial convex decomposition of ρ into the

[8] The normalized trace is a distinguished state in $S(M_N(\mathbb{C}))$ – it is the only state ω that has the cyclicity property $\omega(AB) = \omega(BA)$.

projections $|b_k\rangle\langle b_k| \in \mathscr{D}_N$ unless all but one eigenvalue vanishes, i.e. $\lambda_k = \delta_{k,k_0}$. Thus ρ not being a rank one projection implies ρ not being pure.

To prove the converse, we consider a rank one projection P and write it as a convex combination, $P = c\rho + (1-c)\sigma$ with $\rho, \sigma \in \mathscr{D}_N$ and $0 < c < 1$. Since $P = P^2$, this yields

$$1 = \mathrm{Tr}\big(P^2\big) = c^2\,\mathrm{Tr}\big(\rho^2\big) + (1-c)^2\,\mathrm{Tr}\big(\sigma^2\big) + 2c(1-c)\,\mathrm{Tr}(\rho\sigma). \tag{2.20}$$

We have $\mathrm{Tr}\big(\rho^2\big) \leq 1$, $\mathrm{Tr}\big(\sigma^2\big) \leq 1$, and by Cauchy-Schwarz

$$\mathrm{Tr}(\rho\sigma) = N\omega_{1/N}(\rho^*\sigma) \leq N\sqrt{\omega_{1/N}(\rho^2)\omega_{1/N}(\sigma^2)} = \sqrt{\mathrm{Tr}(\rho^2)\,\mathrm{Tr}(\sigma^2)} \leq 1.$$

Inserting this into (2.20) gives

$$1 = c^2\,\mathrm{Tr}\big(\rho^2\big) + (1-c)^2\,\mathrm{Tr}\big(\sigma^2\big) + 2c(1-c)\,\mathrm{Tr}(\rho\sigma) \leq c^2 + (1-c)^2 + 2c(1-c) = 1,$$

and we conclude that we must have $\mathrm{Tr}\big(\rho^2\big) = 1$, $\mathrm{Tr}\big(\sigma^2\big) = 1$ and $\mathrm{Tr}(\rho\sigma) = 1$. The first two equations imply that ρ, σ are one-dimensional projections, say $\rho = P_\xi, \sigma = P_\eta$. Then the last equation reads $1 = \mathrm{Tr}\big(P_\xi P_\eta\big) = |\langle\xi,\eta\rangle|^2$, which implies $\xi = \alpha\eta$ for some $\alpha \in \mathbb{C}$, $|\alpha| = 1$, and hence $P_\xi = P_\eta$. So we conclude that the convex decomposition of P_ψ is necessarily trivial, i.e. P_ψ is an extreme point of $\mathscr{D}_N$.

A this point we have proven the equivalence of a) and b). A rank one projection P is always of the form $P = P_\psi = |\psi\rangle\langle\psi|$, and

$$\omega_{P_\psi}(A) = \mathrm{Tr}\big(P_\psi A\big) = \sum_{k=1}^{N} \langle e_k, \psi\rangle\langle\psi, Ae_k\rangle = \langle\psi, A\psi\rangle \tag{2.21}$$

shows that the states given by rank one projections as density matrices are precisely the vector states.

It remains to show that every state is a convex combination of pure states. But this is easy: Given a density matrix ρ, we choose an orthonormal basis $(\psi_k)_k$ of eigenvectors with eigenvalues $0 \leq \lambda_k \leq 1$, $\sum_k \lambda_k = 1$, and have $\rho = \sum_k \lambda_k P_{\psi_k}$, i.e. $\omega_\rho = \sum_k \lambda_k \omega_{P_{\psi_k}}$ is a convex combination of the pure states $\omega_{P_{\psi_k}}$. □

We remark that every rank one projection P is of the form $P = P_\psi$ for some normalized vector ψ, but ψ is not uniquely determined by P: Multiplying ψ by a complex phase $\alpha \in \mathbb{C}$, $|\alpha| = 1$, gives the same projection $P_{\alpha\psi} = P_\psi$ and hence the same state. In the interest of light notation, we will often abbreviate

$$\omega_\psi := \omega_{P_\psi}, \qquad \Delta_\psi := \Delta_{P_\psi}, \qquad \psi \in \mathbb{C}^N,\ \|\psi\| = 1, \tag{2.22}$$

when it is clear from context that ψ is a unit vector.

Example 2.15 (Some examples of density matrices). Let $N = 3$ and consider the maximally mixed density matrix $\rho = \frac{1}{3}\cdot 1$ and $\sigma = \frac{1}{2}(|e_1\rangle\langle e_1| + |e_2\rangle\langle e_2|)$ and $\gamma = \frac{1}{2}|e_1 - e_2\rangle\langle e_1 - e_2|$, i.e.

$$\rho = \frac{1}{3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \sigma = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \gamma = \frac{1}{2}\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.23}$$

Note that ω_γ is pure, but ω_ρ and ω_σ are not. A non-trivial decomposition of σ into rank one projections is clearly

$$\sigma = \frac{1}{2}\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}0&0&0\\0&1&0\\0&0&0\end{pmatrix} = \frac{1}{2}(|e_1\rangle\langle e_1| + |e_2\rangle\langle e_2|).$$

See Problem 2.6 for some more training with density matrices.

Let us compare Theorem 2.14 with its classical analogue, Theorem 1.35. In both cases, an explicit characterization of all pure states (by Dirac measures and rank one projections, respectively) is provided. In both cases, pure states are states of minimal support: In the classical setting, this means that the support of the probability measure giving a pure classical state is minimal as a one-point set, and in the quantum setting this means that the support projection of a density matrix giving a pure quantum state is minimal as a projection of rank one. In both cases, there is also a statement that a general mixed state arises as a mixture (or limits of mixtures) of pure states.

However, Theorem 1.35 has two more characterizations of classical pure states in terms of correlations and uncertainties that are missing in Theorem 2.14. The reason for that is that in the quantum case, pure states still exhibit uncertainties and correlations.

Proposition 2.16 (Uncertainties in states on $M_N(\mathbb{C})$). *Let $\rho \in \mathscr{D}_N$ and $A = A^* \in M_N(\mathbb{C})$.*

a) $\Delta_\rho(A) = 0$ if and only if $(A - \omega_\rho(A))P_\rho = 0$. If ω_ρ is pure and $\rho = P_\psi$, then $\Delta_\psi(A) = \|A\psi - \omega_\psi(A)\psi\|$, which vanishes if and only if ψ is an eigenvector of A, with eigenvalue $\omega_\psi(A)$.

b) The only selfadjoint $A \in M_N(\mathbb{C})$ such that $\Delta_\psi(A) = 0$ for all pure states (vector states) are the multiples of the identity.

c) There exists a pure state ω_ψ (depending on ρ and A) such that

$$\Delta_\psi(A) \leq \Delta_\rho(A). \tag{2.24}$$

Proof. a) Let $B := A - \omega_\rho(A)1$. By definition of the support projection P_ρ, we have $\rho = P_\rho\rho = \rho P_\rho$. Hence

$$\Delta_\rho(A)^2 = \mathrm{Tr}\big(\rho B^2\big) = \mathrm{Tr}\big(\rho^{1/2}P_\rho B^2 P_\rho \rho^{1/2}\big). \tag{2.25}$$

As the matrix in the last expression is positive, we see that $\Delta_\rho(A) = 0$ is equivalent to $\rho^{1/2}P_\rho B^2 P_\rho\rho^{1/2} = 0$. This matrix is equal to zero on the range of $P_\rho^\perp$, and on the range of P_ρ, the density matrix ρ and hence $\rho^{1/2}$ is invertible. We conclude that $\Delta_\rho(A) = 0$ is equivalent to $0 = P_\rho B^2 P_\rho = |(A - \omega_\rho(A))P_\rho|^2$, i.e. to $(A - \omega_\rho(A))P_\rho = 0$.

In case $\rho = P_\psi$ is pure, we have

$$\Delta_\psi(A) = \langle\psi, (A - \omega_\psi(A))^2\psi\rangle^{1/2} = \|(A - \omega_\psi(A))\psi\|. \tag{2.26}$$

It is clear that this norm vanishes if and only if ψ is an eigenvector of A, with eigenvalue $\omega_\psi(A) = \langle\psi, A\psi\rangle$.

b) For $A = c1$ we clearly have $\Delta_\rho(A) = 0$ for all $\rho \in \mathscr{D}_N$ and in particular for all pure states. If A is not a multiple of the identity, there exists a unit vector ψ that is not an eigenvector of A. For such ψ, we have $\Delta_\psi(A) = \|A\psi - \omega_\psi(A)\psi\| > 0$.

c) Let $\rho = \sum_k \lambda_k P_k$ be a convex decomposition of ρ into rank one projections. Then

$$\begin{aligned}
\Delta_\rho(A)^2 = \omega_\rho(A^2) - \omega_\rho(A)^2 &= \sum_k \lambda_k \omega_{P_k}(A^2) - \omega_\rho(A)^2 \\
&= \sum_k \lambda_k (\omega_{P_k}(A^2) - \omega_{P_k}(A)^2) + \sum_k \lambda_k (\omega_{P_k}(A)^2 - \omega_\rho(A)^2) \\
&= \sum_k \lambda_k \Delta_{P_k}(A)^2 + \sum_k \lambda_k (\omega_{P_k}(A) - \omega_\rho(A))^2 \\
&\geq \sum_k \lambda_k \Delta_{P_k}(A)^2 \\
&\geq \min_k \Delta_{P_k}(A)^2.
\end{aligned}$$

If k_0 is an index realizing the above minimum, the uncertainty of A in the pure state $\omega_{P_{k_0}}$ is at most as large as the uncertainty of A in the mixed state ω_ρ. □

These facts represent a fundamental difference between matrix mechanics and classical mechanics: Uncertainty free states do not exist. As in classical mechanics, the pure states are states of minimal uncertainty as described above, but even in pure states, non-zero uncertainties show up. The same holds for correlations: $M_N(\mathbb{C})$ does not have product states (Problem 2.3).

Hidden variable theories. Some properties of the states of $M_N(\mathbb{C})$ are radically different from classical physics: Whereas in classical mechanics, all probabilistic aspects were due to an "unsharp" initial condition and arise by mixing uncertainty free states to model partial ignorance about the system, in quantum mechanics even in a perfectly (pure) prepared state, our knowledge of the system appears to be incomplete. This has led to an intense debate which is partially still active even today. In particular Einstein initially opposed this with his famous quote that "God does not play dice", critizing the probabilistic interpretation of quantum mechanics.

A possible argument is that quantum mechanics (or rather matrix mechanics at this point) is "incomplete" in the sense that there is another underlying model of classical mechanics, but some variables are not accessible to us and we only observe statistical mixtures due to this ignoarance. This is roughly speaking the idea of a "hidden variable theory". As we will partially see later, there are very good reasons not to believe in hidden variable theories [Lan17, Chapter 6]. This however leads us to believe that quantum mechanics (which will in many ways be similar to the matrix mechanics discussed here) is "complete", i.e. that the uncertainties inherent in the quantum states are not a consequence of ignorance or experimental imperfections, but a matter of principle. Although this leads to predictions that contradict our common sense trained on macroscopic phenomena (we will see examples of this later), this is the widely accepted point of view today.

From a mathematical perspective, Proposition 2.16 establishes an interesting link between uncertainties and eigenvalue (spectral) considerations: The pure states (vector states)

in which a given observable A has zero uncertainty are precisely the ones given by the eigenvectors of A. To pursue this connection further, recall that in classical mechanics, a classical state ω_μ (with compactly supported Borel probability measure μ on M) induced measures $f_*\mu$ on $\overline{f(M)} \subset \mathbb{R}$ such that $\int_{\mathbb{R}} \lambda^n \, d(f_*\mu)(\lambda) = \omega_\mu(f^n)$ for all $n \in \mathbb{N}_0$: The measure $f_*\mu$ describes all moments and the probability distribution of f in the state ω_μ.

In matrix mechanics, the analogy is the following. (We write E^A for the spectral measure of normal $A \in M_N(\mathbb{C})$, see Appendix M.2).

Proposition 2.17 (Probabilistic interpretation of matrix mechanics and the Born rule). *Let $A \in M_N(\mathbb{C})$ and $\rho \in \mathcal{D}_N$. Then there exists a unique Borel probability measure μ_ρ^A such that $\int_{\mathbb{R}} \lambda^n \, d\mu_\rho^A(\lambda) = \omega_\rho(A^n)$, $n \in \mathbb{N}_0$. This measure is*

$$\mu_\rho^A = \sum_{\lambda \in \sigma(A)} \omega_\rho(E^A(\lambda)) \, \delta_\lambda, \tag{2.27}$$

the spectral measure of A in the state ω_ρ. *Its support is contained in the spectrum $\sigma(A)$ of A, and for any function $f : \mathbb{R} \to \mathbb{C}$,*

$$\omega_\rho(f(A)) = \int_{\mathbb{R}} f(\lambda) \, d\mu_\rho^A(\lambda) = \int_{\sigma(f)} f(\lambda) \, d\mu_\rho^A(\lambda). \tag{2.28}$$

In particular,

$$\Delta_\rho(A)^2 = \int_{\sigma(f)} (\lambda - \omega_\rho(A))^2 \, d\mu_\rho^A(\lambda), \tag{2.29}$$

$$\Pr_\rho(A \in S) = \omega_\rho(E^A(S)) = \int_S d\mu_\rho^A(\lambda). \qquad \text{(Born rule)} \tag{2.30}$$

Proof. The expression (2.27) is a Borel measure on $\mathbb{R}$ because it is a linear combination of Dirac measures and the coefficients $\omega_\rho(E^A(\lambda))$ are non-negative because $E^A(\lambda) > 0$ and ω_ρ is a positive functional. Furthermore,

$$\mu_\rho^A(\mathbb{R}) = \sum_{\lambda \in \sigma(A)} \omega_\rho(E^A(\lambda)) = \omega_\rho\left(\sum_{\lambda \in \sigma(A)} E^A(\lambda) \right) = \omega_\rho(1) = 1,$$

so μ_ρ^A is a probability measure. Its moments are

$$\begin{aligned} \mathbb{M}_n(\mu_\rho^A) = \int_{\mathbb{R}} s^n \, d\mu_\rho^A(s) &= \sum_{\lambda \in \sigma(A)} \omega_\rho(E^A(\lambda)) \int_{\mathbb{R}} s^n \, d\delta_\lambda(s) \\ &= \sum_{\lambda \in \sigma(A)} \omega_\rho(E^A(\lambda)) \lambda^n = \omega_\rho\left(\sum_{\lambda \in \sigma(A)} \lambda^n \, E^A(\lambda) \right) = \omega_\rho(A^n), \end{aligned}$$

as claimed.

Let now μ be any Borel probability measure on $\mathbb{R}$ such that $\int_{\mathbb{R}} \lambda^n \, d\mu(\lambda) = \omega_\rho(A^n)$, $n \in \mathbb{N}_0$. Then $\int_{\mathbb{R}} \lambda^n \, d\mu(\lambda) = \int_{\mathbb{R}} \lambda^n \, d\mu_\rho^A(\lambda)$ for all n, so integrating any polynomial p

against μ or μ_ρ^A gives the same result. We consider the polynomial

$$p(s) := \prod_{\lambda\in\sigma(A)} (s-\lambda)^2 \geq 0.$$

Then $\int_{\mathbb{R}} p\, d\mu = \int_{\mathbb{R}} p\, d\mu_\rho^A = 0$, so $p = 0$ μ-almost everywhere because p is non-negative. The polynomial p takes values larger than zero on the set $\mathbb{R}\backslash\sigma(A)$, which ist therefore a μ-zero set, $\mu(\mathbb{R}\backslash\sigma(A) = 0$. Thus $\operatorname{supp}\mu \subset \sigma(A)$. As in Lemma M.3, this implies $\mu = \sum_{\lambda\in\sigma(A)} c_\lambda \delta_\lambda$ with coefficients $c_\lambda \geq 0$. Fix $\lambda \in \sigma(A)$ and consider a polynomial q_λ with $q_\lambda(\lambda') = \delta_{\lambda\lambda'}$, $\lambda' \in \sigma(A)$. Then $c_\lambda = \int q_\lambda\, d\mu = \int q_\lambda\, d\mu_\rho^A = \omega_\rho(E^A(\lambda))$, which shows $\mu = \mu_\rho^A$.

The statement about the support of μ_ρ^A is clear. The other claims follow by integrating against the measure μ_ρ^A of by inserting the functional calculus of A, namely $f(A) = \sum_{\lambda\in\sigma(A)} f(\lambda)E^A(\lambda)$, into ω_ρ, for instance for the expectation value

$$\omega_\rho(f(A)) = \sum_{\lambda\in\sigma(A)} f(\lambda)\omega_\rho(E^A(\lambda)) = \int_{\mathbb{R}} f(s)\, d\mu_\rho^A(s).$$

□

Further information about the connection between the spectrum and expectation values and uncertainties is contained in the following lemma.

Lemma 2.18. *Let $A = A^* \in M_N(\mathbb{C})$. Denote by $\lambda^A_{\max/\min} \in \mathbb{R}$ the maximal/minimal eigenvalue of A, by $\operatorname{diam}\sigma(A) := \frac{1}{2}(\lambda^A_{\max} - \lambda^A_{\min})$ the diameter of the spectrum of A, and by $\operatorname{Conv}\sigma(A) := [\lambda^A_{\min}, \lambda^A_{\max}]$ the convex hull of the spectrum of A.*

a) The possible expectation values of A in arbitrary states are

$$\{\omega_\rho(A) \,:\, \rho \in \mathcal{D}_N\} = \{\omega_\psi(A) \,:\, \psi \in \mathbb{C}^N,\ \|\psi\| = 1\} = \operatorname{Conv}\sigma(A). \tag{2.31}$$

b) The uncertainties satisfy

$$\sup_{\rho\in\mathcal{D}_A} \Delta_\rho(A) = \operatorname{diam}(\sigma(A)) \leq \|A\|. \tag{2.32}$$

The proof will be given in Exercise 2.2.

Example 2.19 (A particle on a one-dimensional lattice). In this example we discuss a toy model of a particle moving on a one-dimensional finite lattice consisting of N sites. The position of the particle can take the N distinct sharp values $1, \dots, N$,

and hence we choose the diagonal matrix

$$X := \operatorname{diag}(1, \dots, N) = \sum_{k=1}^{N} kE_{kk} \tag{2.33}$$

as the observable describing the position of the particle. The canonical orthonormal basis of $\mathbb{C}^N$ consists of eigenvectors (e_k) of X, namely $Xe_k = ke_k$, and we have $E^X(k) = E_{kk}$. Hence the spectral representation of X is $X = \sum_{k=1}^N k\, E_{kk}$.

Given a density matrix $\rho \in \mathscr{D}_N$, the probability measure for position in this state is

$$\mu_\rho^X = \sum_{k=1}^N \operatorname{Tr}(\rho E_{kk})\delta_k = \sum_{k=1}^N \rho_{kk}\, \delta_k.$$

In the special case of a pure state, i.e. $\rho = P_\psi$ for some unit vector $\psi \in \mathbb{C}^N$, we have $\rho_{kk} = |\langle \psi, e_k\rangle|^2$. So in this case, the absolute squares of the basis expansion coefficients of ψ are the weights of the delta measures.

See problem 2.4 for further investigations of this example.

Example 2.20 (Projection observables and transition probabilities). Let us consider an orthogonal projection $P = P^* = P^2 \in M_N(\mathbb{C})$. Then $\sigma(P) \subset \{0,1\}$, with the trivial cases $\sigma(P) = \{0\}$ and $\sigma(P) = \{1\}$ only occuring for the trivial projections $P = 0$ and $P = 1$, respectively.

Hence the possible expectation values of non-trivial P form the unit interval $[0,1]$, and the possible sharp expectation values (those with zero uncertainty) are the end points 0 and 1. The observable P may be interpreted as a "true/false" observable, like a coin flip or logical proposition.

As the spectral projections of P are $E^P(1) = P$ and $E^P(0) = 1 - P =: P^\perp$,

$$\begin{aligned} \Pr_\omega(P \in \{1\}) &= \omega(E^P(1)) = \omega(P) \quad \text{"true"}, \\ \Pr_\omega(P \in \{0\}) &= \omega(E^P(0)) = \omega(P^\perp) \quad \text{"false"}, \end{aligned} \tag{2.34}$$

are the probabilities that the proposition described by P is true/false in the state ω.

In particular, given a unit vector $\psi \in \mathbb{C}^N$, we may consider the rank one projection $P = P_\psi$. As a proposition, this projection has the meaning "Is the state of the system ψ?". In a (say, pure) state ω with density projection P_ϕ, the *transition probability*

$$\Pr_{P_\phi}(P_\psi \in \{1\}) = \langle \phi, P_\psi \phi\rangle = |\langle \phi, \psi\rangle|^2 \tag{2.35}$$

is the probability that the state of the system is ψ when it was prepared in the state ϕ.

Up to this point, the probabilistic setting of classical and matrix mechanics exhibit strong similarities: Given a state ω and an observable A, there exists a unique Borel probability measure on the spectrum of A that describes the distribution (expectation values, uncertainties, probabilities) of the values of A in this state. A difference between classical and matrix mechanics is that in matrix mechanics, we do not have any zero uncertainty states.

Another difference shows up when considering several observables instead of a single one. Recall that in classical mechanics, given $f_1, \dots, f_n \in C^\infty_{\mathbb{R}}(M)$ and any state, we had found joint probability measures describing the joint distribution of these observables (Def. 1.28). This had the property that for any $m_1, \dots, m_n \in \mathbb{N}_0$, the multinonomial $q_m(\lambda_1, \dots, \lambda_n) :=$

$\lambda_1^{m_1} \cdots \lambda_n^{m_n}$ integrates against $f_*\mu$ to the expectation value of the product $f_1^{m_1} \cdots f_n^{m_n}$,

$$\int_{\mathbb{R}^n} q_m \, d(f_*\mu) = \omega_\mu(f_1^{m_1} \cdots f_n^{m_n}). \tag{2.36}$$

The analogue of this result in quantum mechanics is best formulated in terms of joint spectra (Proposition M.16).

Proposition 2.21 (Incommensurable observables). *Let $A_1, \dots, A_n \in M_N(\mathbb{C})$ be self-adjoint. The following are equivalent:*

a) For every density matrix $\rho \in \mathscr{D}_N$, there exists a Borel probability measure μ_ρ on $\mathbb{R}^n$ such that

$$\int_{\mathbb{R}^n} s_1^{m_1} \cdots s_n^{m_n} \, d\mu_\rho(s_1, \dots, s_n) = \omega_\rho(A_1^{m_1} \cdots A_n^{m_n}), \qquad m_1, \dots, m_n \in \mathbb{N}_0, \tag{2.37}$$

b) $[A_k, A_l] = 0$ for all $k, l \in \{1, \dots, n\}$.

If this is the case, the measure μ_ρ is given by the expectation value of the joint spectral measure,

$$\mu_\rho(S) = \omega_\rho(E^{A_1,\dots,A_n}(S)) = \sum_{\lambda \in S} E^{A_1,\dots,A_n}(\lambda). \tag{2.38}$$

Proof. b) $\Longrightarrow$ a) According to Prop. M.16, we have the joint spectral measure $E^{A_1,\dots,A_n} = \sum_{\lambda \in \sigma(A_1,\dots,A_n)} E^{A_1}(\lambda_1) \cdots E^{A_n}(\lambda_n) \, \delta_\lambda$, where $\sigma(A) \subset \mathbb{R}^n$ denotes the joint spectrum of $(A_1, \dots, A_n)$. Then

$$\mu_\rho(S) := \mathrm{Tr}\big(\rho E^{A_1,\dots,A_n}(S)\big), \qquad S \in \mathfrak{B}(\mathbb{R}^n) \tag{2.39}$$

is a Borel probability measure on $\mathbb{R}^n$, and

$$\begin{aligned} \omega_\rho(A_1^{m_1} \cdots A_n^{m_n}) &= \sum_{\lambda \in \sigma(A_1,\dots,A_n)} \lambda_1^{m_1} \cdots \lambda_n^{m_n} \, \mathrm{Tr}\big(\rho \, E^{A_1}(\lambda_1) \cdots E^{A_n}(\lambda_n)\big) \\ &= \int_{\mathbb{R}^n} s_1^{m_1} \cdots s_n^{m_n} \, d\mu_\rho(s_1, \dots, s_n). \end{aligned}$$

a) $\Longrightarrow$ b) Let $k, l \in \{1, \dots, n\}$, with $k < l$. Then

$$\mathrm{Tr}(\rho A_k A_l) = \int_{\mathbb{R}^n} \lambda_k \lambda_l \, d\mu_\rho(\lambda) = \int_{\mathbb{R}^n} \lambda_l \lambda_k \, d\mu_\rho(\lambda) = \mathrm{Tr}(\rho A_l A_k),$$

i.e. $C := [A_k, A_l]$ satisfies $0 = \mathrm{Tr}(\rho C)$ for all $\rho \in \mathscr{D}_N$. Picking $\rho = P_\psi$ with an arbitrary unit vector ψ, this condition implies $0 = \mathrm{Tr}\big(P_\psi C\big) = \langle \psi, C\psi \rangle$, and hence $C = 0$. Thus A_k commutes with A_l. □

This observation represents another drastic difference to classical mechanics: Whereas in classical mechanics, states are given by probability measures on phase space and induce probability measures for all classical observables $f \in C^\infty_{\mathbb{R}}(M)$, and joint probability measures for arbitrary families $(f_1, \dots, f_n)$ of classical observables, in quantum mechanics a state

is given by a density matrix that induces probability measures on the spectrum of each single observable, and also of joint spectra of families $(A_1, \dots, A_n)$ of *commuting* observables. However, as $M_N(\mathbb{C})$ is noncommutative, there exist also families of observables given by noncommuting matrices (called *incommensurable* in the physics literature), the joint moments of which are not given by a joint measure as above.

Note, however, that even for noncommuting $A_1, \dots, A_n$, there might exist *some* density matrices ρ such that (2.37) holds.

Selective Lüders measurements. Let $A, B \in M_N(\mathbb{C})$ be selfadjoint and not necessarily commuting (for simplicity, we restrict ourselves to two observables here), and $\rho \in \mathscr{D}_N$. Then

$$\begin{aligned} \mu_\rho^{A\to B}(S) &:= \sum_{\lambda\in S} \mu_\rho^{A\to B}(\lambda_1,\lambda_2), \\ \mu_\rho^{A\to B}(\lambda_1,\lambda_2) &:= \mathrm{Tr}\big(\rho E^A(\lambda_1)E^B(\lambda_2)E^A(\lambda_1)\big) \end{aligned} \tag{2.40}$$

is a Borel probability measure on $\mathbb{R}^2$, with support contained in $\sigma(A)\times\sigma(B)$ (this is easy to check). Its marginals are $\mu_\rho^{A\to B}(S\times\mathbb{R}) = \mu_\rho^A(S)$ and $\mu_\rho^{A\to B}(\mathbb{R}\times S) = \mu_\rho^B(S)$, i.e. fully compatible with spectral measures. In case A and B commute, we have $\mu_\rho^{A\to B}(S) = \mathrm{Tr}\big(\rho E^{(A,B)}(S)\big)$.

However, for noncommuting A, B, the measure $\mu_\rho^{A\to B}$ in general depends on the order A, B, in contrast to the integrals (2.37). It is usually interpreted as follows:

The possible values of measurements of A are the spectral values of A. There is no definite prediction about what happens in a single measurement, only statistical information – the probability of obtaining $\lambda \in \sigma(A)$ as measurement outcome in the state ρ is $\omega_\rho(E^A(\lambda))$, and this fixes the expectation value and the whole probability measure.

If the measured outcome is λ (with non-zero probability, i.e. $\rho E^A(\lambda) \neq 0$), then the state changes from ρ to the *post-measurement state*

$$\frac{1}{\mathrm{Tr}(\rho E^A(\lambda))} E^A(\lambda)\rho E^A(\lambda) \in \mathscr{D}_N,$$

which is a density matrix because $E^A(\lambda)\rho E^A(\lambda)$ is positive and has trace $\mathrm{Tr}\big(\rho E^A(\lambda)\big)$. The interpretation is that the measurement has disturbed the system and changed its state. If one measures A again, one will again obtain the outcome λ, with probability 1, because

$$\mathrm{Tr}\left(\frac{1}{\mathrm{Tr}(\rho E^A(\lambda))} E^A(\lambda)\rho E^A(\lambda)\, A^n\right) = \lambda^n, \tag{2.41}$$

$$\Delta_{\frac{1}{\mathrm{Tr}(\rho E^A(\lambda))} E^A(\lambda)\rho E^A(\lambda)}(A) = 0. \tag{2.42}$$

If one measures B after the A-measurement, the probability measure for B is $S \mapsto \mu^{A\to B}(\{\lambda\}\times S)$, explaining the significance of the measure $\mu_\rho^{A\to B}$ for conditional probabilities, and the asymmetry in $A \leftrightarrow B$ in case A and B do not commute.

We refer to Exercise 2.3 for further discussion of this, and to the book [BLM96] for a thorough analysis of measurements in quantum mechanics.

For operations on density matrices, we have so far considered convex combinations, i.e. mixing. For vector states, there is another possibility: Given two normalized vectors $\psi, \xi \in \mathbb{C}^N$, we can form a linear combination and normalize to get another normalized vector, and hence another vector state. This is called superposition.

Definition 2.22 (Superpositions). Let $\psi, \xi \in \mathbb{C}^N$, $\|\psi\| = \|\xi\| = 1$, with $\psi \neq -\xi$. A *superposition* of the vector states ω_ψ, ω_ξ is a vector state ω_η with $\eta = c\psi + c'\xi$, $\|\eta\| = 1$.

Note that the superposition of two vectors ξ, ψ gives a vector state that is different from a mixture of the vector states given by ξ and ψ! In contrast to mixtures, superpositions have no analogue in classical mechanics and lead to wave-like interference phenomena. This will be investigated in examples in Problem 2.6 and Example 2.34.

The following table collects similarities and differences between the states and observables in classical and matrix mechanics.

2.3 Dynamics and symmetries in matrix mechanics

We now move on from our kinematical discussion of states in matrix mechanics to *dynamics* (time evolution). Our guiding principle is again classical mechanics: In classical mechanics, we had considered a Hamiltonian function $H \in C^\infty_{\mathbb{R}}(M)$, which defined a Hamiltonian flow ϕ^H, and, more importantly, induced a dynamics

$$\alpha^H_t : C^\infty(M) \to C^\infty(M) \tag{2.43}$$

on the classical observable algebra $C^\infty(M)$ such that

$$\frac{d}{dt}\alpha^H_t(f) = \{\alpha^H_t(f), H\}. \tag{2.44}$$

The maps α^H_t were automorphisms of our classical observable algebra and formed a one-parameter group, namely $\alpha^H_t \circ \alpha^H_s = \alpha^H_{t+s}$.

In the following, we will investigate the analogous structures in matrix mechanics. We begin by recalling what an automorphism is.

Definition 2.23 (Automorphisms of unital *-algebras). Let $\mathcal{A}$ be a unital *-algebra. An *automorphism* of $\mathcal{A}$ (or *unital *-automorphism* for emphasis) is a bijective linear map $\alpha : \mathcal{A} \to \mathcal{A}$ satisfying for all $A, B \in \mathcal{A}$

a) $\alpha(AB) = \alpha(A)\alpha(B)$,
b) $\alpha(A^*) = \alpha(A)^*$,
c) $\alpha(1) = 1$.

The set of all automorphisms of $\mathcal{A}$ is denoted $\operatorname{Aut}\mathcal{A}$.

Note that $\operatorname{Aut}\mathcal{A}$ is a group, the *automorphism group of* $\mathcal{A}$. The automorphism group of $M_N(\mathbb{C})$ can be described quite explicitly:

Proposition 2.24 (Automorphisms of $M_N(\mathbb{C})$).
a) Let $U \in U(N)$ be a unitary matrix. Then

$$\alpha_U : M_N(\mathbb{C}) \to M_N(\mathbb{C}), \qquad \alpha_U(A) := UAU^*, \tag{2.45}$$

is a unital *-automorphism of $M_N(\mathbb{C})$, i.e. $\alpha_U \in \operatorname{Aut} M_N(\mathbb{C})$, and $U \mapsto \alpha_U$ is a

Aspect	**Classical mechanics**	**Matrix mechanics**
Observable algebra	$C^\infty(M)$ (commutative unital *-algebra)	$M_N(\mathbb{C})$ (noncommutative unital *-algebra)
Real observables	$f = \bar{f} \in C^\infty(M)$	$A = A^* \in M_N(\mathbb{C})$
States	compactly supported Borel probability measures $\mu \in \mathcal{P}_c(M)$	density matrices $\rho \in \mathscr{D}_N$
Mixtures	Convex combinations of probability measures	Convex combinations of density matrices
Superposition	—	vector state given by normalized linear combinations of other vectors
Pure states	Dirac measures μ_ξ, $\xi \in M$ (zero uncertainty in all observables)	vector states $\rho = P_\psi$, $\psi \in \mathbb{C}^N$, $\|\psi\| = 1$ (not zero uncertainty in all observables)
Expectation value of observable in a state	$\omega_\mu(f) = \int_M f(\xi)\, d\mu(\xi)$	$\omega_\rho(A) = \mathrm{Tr}(\rho A)$
Uncertainty of observable in a state	$\Delta_\mu(f)^2 = \omega_\mu(f^2) - \omega_\mu(f)^2$	$\Delta_\rho(A)^2 = \omega_\rho(A^2) - \omega_\rho(A)^2$
Prob. measure on $\mathbb{R}$ of observable in state	$f_*\mu(B) = \mu(f^{-1}(B))$	$\mu_\rho^A(B) = \mathrm{Tr}\big(\rho E^A(B)\big)$
Spectrum of observable	$\sigma(f) = \overline{f(M)}$	$\sigma(A) = \{\text{eigenvalues of } A\}$
Joint probability measure of several obs. in a state	always exists, $(f_1 \times \ldots \times f_n)_*\mu$	exists only for commuting $A_1, \ldots, A_n$. In this case, given by $\omega_\rho(E^{A_1,\ldots,A_n})$

Table 1: Similarities and differences between states and observables (kinematics) in classical mechanics and matrix mechanics.

group homomorphism, i.e.

$$\alpha_U \circ \alpha_V = \alpha_{UV}, \qquad U, V \in U(N). \tag{2.46}$$

b) Let $U, V \in U(N)$. Then $\alpha_U = \alpha_V \iff U = zV$ for some $z \in \mathbb{C}$, $|z| = 1$.
c) Let $\alpha \in \operatorname{Aut} M_N(\mathbb{C})$. Then $\alpha = \alpha_U$ for some $U \in U(N)$.

Proof. a) follows easily by checking the properties of automorphisms. For b), we note that $\alpha_U = \alpha_V$ is equivalent to $UAU^* = VAV^* \iff [(V^*U), A] = 0$ for all $A \in M_N(\mathbb{C})$, hence to V^*U being a multiple of the idenity.

c) Let $\alpha \in \operatorname{Aut} M_N(\mathbb{C})$ and consider $F_{kl} := \alpha(E_{kl})$ be the images of the matrix units. Because α is an automorphism, the F_{kl} form a system of matrix units, and in particular, the F_{kk} are mutually orthogonal rank one projections. Thus $F_{kk} = |f_k\rangle\langle f_k|$ form an orthonormal basis (f_k) of $\mathbb{C}^N$.

The unitary U defined by $Ue_k := f_k$, $k = 1, \ldots, N$, satisfies

$$\alpha_U(E_{kl}) = UE_{kl}U^* = U|e_k\rangle\langle e_l|U^* = |f_k\rangle\langle f_l| = F_{kl} = \alpha(E_{kl}).$$

As $(E_{kl})_{k,l}$ is a basis of $M_N(\mathbb{C})$, this implies $\alpha = \alpha_U$. □

We thus see that all automorphisms α of $M_N(\mathbb{C})$ are "inner", i.e. given by unitaries U according to $\alpha(A) = UAU^*$. However, U is determined by α only up to a phase (part b)). Denoting the unit circle in the complex plane $\mathbb{T}$, the content of Proposition 2.24 can therefore be summarized by saying that we have an isomorphism of groups

$$\operatorname{Aut} M_N(\mathbb{C}) \cong U(N)/\mathbb{T} =: \operatorname{PU}(N), \tag{2.47}$$

where in the quotient group $U(N)/\mathbb{T}$, unitaries that differ by a multiplicative phase in $\mathbb{T}$ are identified. The group $\operatorname{PU}(N)$ is called the *projective unitary group*. We will say more about in Section 2.4.

Automorphisms α preserve all algebraic structure of $M_N(\mathbb{C})$, i.e. they are linear, carry products into products, adjoints into adjoints, the identity into the identity. They also preserve commutators, $\alpha([A, B]) = [\alpha(A), \alpha(B)]$, and positivity: $A \geq 0 \implies \alpha(A) \geq 0$ (this is the case because $A \geq 0$ is equivalent to $A = B^*B$ for some $B \in M_N(\mathbb{C})$). Thanks to the description of automorphisms by unitaries in Proposition 2.24, $\alpha = \alpha_U$, we also see that automorphisms preserve traces, namely $\operatorname{Tr} \alpha(A) = \operatorname{Tr} A$ for all $A \in M_N(\mathbb{C})$.

A direct consequence of these observations is that any automorphism α maps density matrices to density matrices,

$$\rho \in \mathcal{D}_N \iff \alpha(\rho) \in \mathcal{D}_N, \tag{2.48}$$

and

$$\omega_\rho(\alpha(A)) = \omega_{\alpha^{-1}(\rho)}(A), \qquad A \in M_N(\mathbb{C}),\ \rho \in \mathcal{D}_N,\ \alpha \in \operatorname{Aut} M_N(\mathbb{C}). \tag{2.49}$$

We now consider one-parameter groups of automorphisms of $M_N(\mathbb{C})$ as models for dynamics in matrix mechanics, i.e. maps

$$\alpha : \mathbb{R} \to \operatorname{Aut} M_N(\mathbb{C}), \qquad \alpha_t \circ \alpha_s = \alpha_{t+s}. \tag{2.50}$$

This immediately implies $\alpha_0 = \mathrm{id}$ and $\alpha_t^{-1} = \alpha_{-t}$. It does however not imply regularity in the time parameter t. In the following, we will therefore assume that α is continuous in the sense that

$$\mathbb{R} \ni t \mapsto \alpha_t(A) \in M_N(\mathbb{C}) \tag{2.51}$$

is continuous for every $A \in M_N(\mathbb{C})$. We call α differentiable or analytic in case the maps (2.51) are differentiable or converging power series in $t \in \mathbb{C}$ for all A, i.e. $\alpha_t(A) = \sum_{n=0}^{\infty} \beta_n(A) t^n$ with $\beta_n(A) \in M_N(\mathbb{C})$ such that the series converges in norm, respectively.

In the context of the classical observable algebra $C^\infty(M)$, the Poisson bracket with the Hamiltonian gives the derivative of the dynamics (2.44). To discuss analogous structures on $M_N(\mathbb{C})$, we identify two relevant properties of Poisson brackets in the following definition (cf. Proposition 1.15):

Definition 2.25 (Derivations). Let $\mathcal{A}$ be a *-algebra. A *derivation* of $\mathcal{A}$ is a linear map[a] $\delta : \mathcal{A} \to \mathcal{A}$ such that for all $A, B \in \mathcal{A}$,

$$\begin{aligned} \delta(AB) &= \delta(A)B + A\delta(B) && \text{(Leibniz rule)}, \\ \delta(A^*) &= \delta(A)^* && \text{(compatability with adjoints)}. \end{aligned} \tag{2.52}$$

[a] It is customary to denote derivations by δ, not to be confused with Dirac measures.

In Exericse 2.1, you have already shown that for $H = H^* \in M_N(\mathbb{C})$, the map

$$\delta_H : M_N(\mathbb{C}) \to M_N(\mathbb{C}), \qquad \delta_H(A) := -i[A, H], \tag{2.53}$$

is a derivation of $M_N(\mathbb{C})$. The following proposition claims that all derivations of $M_N(\mathbb{C})$ are of this form.

Proposition 2.26 (The derivations of $M_N(\mathbb{C})$). *Let δ be a derivation of $M_N(\mathbb{C})$. Then there exists $H = H^* \in M_N(\mathbb{C})$ such that $\delta(A) = -i[A, H]$ for all $A \in M_N(\mathbb{C})$. This matrix H is uniquely determined by δ up to an additive multiple of the identity.*

The proof of this lemma uses the same kind of idea as the proof of Proposition 2.11 and Proposition 2.24: First one considers a derivation $\delta_H(A) = -i[A, H]$ of commutator form and derives a formula expressing H in terms of δ_H. Then one uses this formula for a general derivation δ as the definition of H and checks that $\delta = -i[\,\cdot\,, H]$ holds. The details of the proof will be given in the exercises (Problem 2.7).

Theorem 2.27 (Dynamics on $M_N(\mathbb{C})$).

a) Let $H = H^* \in M_N(\mathbb{C})$ and $U(t) := e^{itH}$. Then

$$\alpha_t^H(A) := U(t) A U(t)^* \tag{2.54}$$

is an analytic one-parameter group of automorphisms of $M_N(\mathbb{C})$, with

$$\frac{d}{dt}\alpha_t^H(A) = -i[\alpha_t^H(A), H], \qquad A \in M_N(\mathbb{C}),\ t \in \mathbb{R}. \tag{2.55}$$

b) Let $H = H^*, K = K^* \in M_N(\mathbb{C})$. Then $\alpha_t^H = \alpha_t^K$ for all $t \in \mathbb{R}$ if and only if $K = H + c1$ for some $c \in \mathbb{R}$.

c) *Let $(\alpha_t)_{t\in\mathbb{R}}$ be a continuous one-parameter group of automorphisms of $M_N(\mathbb{C})$. Then α is analytic and there exists $H = H^* \in M_N(\mathbb{C})$ such that $\alpha_t = \alpha_t^H$, $t \in \mathbb{R}$.*

Proof. a) In the spectral representation $H = \sum_{\lambda\in\sigma(H)} \lambda E^H(\lambda)$ of H, the exponential takes the form

$$U(t) = e^{itH} = \sum_{\lambda\in\sigma(H)} e^{it\lambda} E^H(\lambda). \tag{2.56}$$

This matrix is unitary, namely

$$U(t)^* = \sum_{\lambda\in\sigma(H)} e^{-it\lambda} E^H(\lambda) = U(t)^{-1} = U(-t), \tag{2.57}$$

and we have $U(t)U(s) = U(t+s)$ by functional calculus (exercise). This implies that α_t^H is an automorphism of $M_N(\mathbb{C})$ for every $t \in \mathbb{R}$ (Prop. 2.24), and that the one-parameter property holds. To see that α^H is analytic, we write

$$\alpha_t^H(A) = \sum_{\lambda,\lambda'\in\sigma(H)} e^{it(\lambda-\lambda')} E^H(\lambda) A E^H(\lambda') \tag{2.58}$$

$$= \sum_{n=0}^{\infty} \frac{t^n}{n!} \sum_{\lambda,\lambda'\in\sigma(H)} (\lambda-\lambda')^n E^H(\lambda) A E^H(\lambda'), \tag{2.59}$$

and note that the second sum can be estimated according to

$$\left\| \sum_{\lambda,\lambda'\in\sigma(H)} (\lambda-\lambda')^n E^H(\lambda) A E^H(\lambda') \right\| \leq \sum_{\lambda,\lambda'\in\sigma(H)} (2\|H\|)^n \, \|E^H(\lambda) A E^H(\lambda')\| \leq (2\|H\|)^n N^2 \|A\|.$$

Here we have used that any eigenvalue $\lambda \in \sigma(H)$ satisfies $|\lambda| \leq \|H\|$. As $\sum_{n=0}^{\infty} \frac{t^n}{n!}(2\|H\|)^n N^2\|A\| = e^{2t\|H\|} N^2 \|A\|$ converges for all t, this proves the analyticity.

For the derivative of α^H, we find

$$\begin{aligned}
\frac{d}{dt} e^{itH} A e^{-itH} &= i \sum_{\lambda,\lambda'\in\sigma(H)} (\lambda-\lambda') \, e^{it(\lambda-\lambda')} E^H(\lambda) A E^H(\lambda') \\
&= i \sum_{\lambda\in\sigma(H)} \lambda e^{it\lambda} E^H(\lambda) A \sum_{\lambda'\in\sigma(H)} e^{-it\lambda'} E^H(\lambda') \\
&\quad - i \sum_{\lambda\in\sigma(H)} e^{it\lambda} E^H(\lambda) A \sum_{\lambda'\in\sigma(H)} \lambda' e^{-it\lambda'} E^H(\lambda') \\
&= iHU(t)AU(-t) - iU(t)AU(-t)H \\
&= -i[U(t)AU(-t), H] \\
&= -i[\alpha_t^H(A), H].
\end{aligned}$$

b) If $K = H + c1$, $c \in \mathbb{R}$, then $\alpha^K = \alpha^H$ by Prop. 2.24 b). If $\alpha_t^H = \alpha_t^K$ for all t, their derivatives at $t = 0$ coincide,

$$\frac{d}{dt}\Big|_{t=0} \alpha_t^H(A) = -i[A,H] = -i[A,K] \, \frac{d}{dt}\Big|_{t=0} \alpha_t^K(A), \qquad A \in M_N(\mathbb{C}).$$

Hence $H - K$ commutes with all elements of $M_N(\mathbb{C})$, which implies that it is a multiple of the identity. The multiple must be real by selfadjointness.

c) Let us first assume that $\mathbb{R} \ni t \mapsto \alpha_t(A) \in M_N(\mathbb{C})$ is differentiable (for every $A \in M_N(\mathbb{C})$). Then the derivative at $t = 0$,

$$\delta(A) := \frac{d}{dt}\Big|_{t=0} \alpha_t(A), \qquad A \in M_N(\mathbb{C}),$$

defines a derivation of $M_N(\mathbb{C})$ (exercise). According to Proposition 2.26, there exists $H = H^* \in M_N(\mathbb{C})$ such that $\delta(A) = -i[A, H]$. This is the derivation of the one-parameter group α^H investigated in part a), and we claim $\alpha = \alpha^H$. To prove that, let $A \in M_N(\mathbb{C})$ and consider the matrix-valued function $f : t \mapsto \alpha^H_{-t}(\alpha_t(A))$. We have $f(0) = A$ and want to show $f(t) = A$ for all t; this then implies $\alpha^H_t = \alpha_t$. So we have to show that the derivative of f vanishes. To that end, we consider $\psi, \xi \in \mathbb{C}^N$, an orthonormal basis (b_j) of $\mathbb{C}^N$, and compute

$$\begin{aligned}
\frac{d}{dt}\langle \xi, f(t)\psi\rangle &= \frac{d}{dt}\langle e^{itH}\xi, \alpha_t(A)e^{itH}\psi\rangle \\
&= \frac{d}{dt}\sum_{j,l=1}^{N} \langle e^{itH}\xi, b_j\rangle\langle b_j, \alpha_t(A)b_l\rangle\langle b_l, e^{itH}\psi\rangle \\
&= \sum_{j,l=1}^{N} \Big(\langle iHe^{itH}\xi, b_j\rangle\langle b_j, \alpha_t(A)b_l\rangle\langle b_l, e^{itH}\psi\rangle \\
&\quad - \langle e^{itH}\xi, b_j\rangle\langle b_j, i[A,H]b_l\rangle\langle b_l, e^{itH}\psi\rangle + \langle e^{itH}\xi, b_j\rangle\langle b_j, \alpha_t(A)b_l\rangle\langle b_l, iHe^{itH}\psi\rangle \Big) \\
&= \langle iHe^{itH}\xi, \alpha_t(A)e^{itH}\psi\rangle - \langle e^{itH}\xi, i[\alpha_t(A), H]e^{itH}\psi\rangle + \langle e^{itH}\xi, \alpha_t(A), iHe^{itH}\psi\rangle \\
&= \langle e^{itH}\xi, (-iH\alpha_t(A) - i[\alpha_t(A), H] + i\alpha_t(A)H)e^{itH}\psi\rangle \\
&= 0.
\end{aligned}$$

As ξ, ψ were arbitrary, this shows that f is indeed constant.

It remains to show that the differentiability assumption on α was not necessary. We define

$$I_\varepsilon := \int_0^\varepsilon \alpha_s \, ds, \tag{2.60}$$

where the integral is carried out in the finite-dimensional linear space $\mathrm{End}\, M_N(\mathbb{C})$. By continuity of α, we see that $\varepsilon \mapsto I_\varepsilon$ is differentiable. Since $I_0 = 0$, this implies $\varepsilon^{-1} I_\varepsilon \to \mathrm{id}$ as $\varepsilon \to 0$. The identity is of course invertible, and this implies that I_ε is also invertible for small $|\varepsilon|$. We find for any $t \geq 0$

$$\alpha_t = I_\varepsilon^{-1} I_\varepsilon \alpha_t = I_\varepsilon^{-1} \int_0^\varepsilon \alpha_{t+s} ds = I_\varepsilon^{-1} \int_t^{t+\varepsilon} \alpha_s ds = I_\varepsilon^{-1} \left(I_{t+\varepsilon} - I_t\right),$$

which implies differentiability of α because I is differentiable. □

This theorem allows us to formulate more correspondences between matrix mechanics and classical mechanics: Every dynamics (one-parameter group of automorphisms) is given

by a *Hamiltonian* matrix, just as every dynamics in classical mechanics is given by a Hamiltonian function. In both cases, the Hamiltonian is only determined up to an additive constant. The infinitesimal dynamics (derivative w.r.t. t at $t = 0$) is given by the Poisson bracket with the Hamiltonian function in classical mechanics, and by $-i$ times the commutator with the Hamiltonian matrix in matrix mechanics.

We can now define the analogue of a classical Hamiltonian system.

Definition 2.28.

a) A *matrix mechanical system* consists of a matrix algebra $M_N(\mathbb{C})$ and a selfadjoint matrix $H \in M_N(\mathbb{C})$, called the *Hamiltonian*.

b) A *symmetry* of a matrix mechanical system $(M_N(\mathbb{C}), H)$ is an automorphism $\alpha \in \operatorname{Aut} M_N(\mathbb{C})$ preserving H, i.e. $\alpha(H) = H$.

c) The *dynamics* defined by the Hamiltonian H is the one-parameter group of automorphisms $\alpha_t^H : A \mapsto e^{itH} A e^{-itH} \in \operatorname{Aut} M_N(\mathbb{C})$, $t \in \mathbb{R}$.

d) In a matrix mechanical system $(M_N(\mathbb{C}), H)$, *Heisenberg's Equation* is the differential equation

$$\frac{d}{dt}\alpha_t^H(A) = -i[\alpha_t^H(A), H], \qquad A \in M_N(\mathbb{C}),\ t \in \mathbb{R}. \tag{2.61}$$

At this stage, we have completely translated dynamics into matrix mechanics. However, it should be noted that we currently have no systematic way of building interesting Hamiltonians, in contrast to the situation in classical Newtonian Mechanics where we could work with potentials of conservative force fields and kinetic energies. This will be remedied in quantum mechanics in Chapter 3.

- **(Solution of Heisenberg's Equation.)** Heisenberg's Equation (2.61) replaces Hamilton's Equation (1.35) as the dynamical equation (or equation of motion); its solution describes the time evolution of an observable $A \in M_N(\mathbb{C})$.
  The solution is explicitly given by (see Prop. 2.29 for its uniqueness properties)

  $$\alpha_t^H(A) = e^{itH} A e^{-itH}, \qquad t \in \mathbb{R}. \tag{2.62}$$

  As e^{itH} is unitary, various properties of $\alpha_t^H(A)$ are time-independent, e.g. its norm, its spectrum, selfadjointness/unitarity properties, etc. The eigenprojections of $\alpha_t^H(A)$ can however change with time, and so can its expectation values in states.
  To extract more concrete information about $t \mapsto \alpha_t^H(A)$, e.g. expectation values in certain states of interest, one often wants to compute the matrix exponential e^{itH} more explicitly. This can either be done as a power series in H or in spectral form,

  $$e^{itH} = \sum_{n=0}^{\infty} \frac{i^n t^n H^n}{n!} = \sum_{\lambda \in \sigma(H)} e^{it\lambda} E^H(\lambda). \tag{2.63}$$

  In the latter case, this requires knowledge of the spectral resolution of H, i.e. one needs to find the eigenvalues and eigenprojections of H. Examples will be discussed below.
- **(Heisenberg and Schrödinger picture.)** Exactly as in classical mechanics, we may view the dynamics as either on the side of the observables – that is, observables A change according to $\alpha_t^H(A)$ with time when a Hamiltonian defining the dynamics has been fixed, and states $\omega = \omega_\rho$ are time independent (*Heisenberg picture*), or on the

side of the states – that is, observables A are time-independent, and states $\omega = \omega_\rho$ change according to $\alpha^H_{-t}(\rho)$ with time (*Schrödinger picture*). Note that $\alpha^H_{-t}(\rho)$ is also a density matrix, and (cf. (2.49))

$$\omega_\rho(\alpha^H_t(A)) = \omega_{\alpha^H_{-t}(\rho)}(A) \tag{2.64}$$

shows that all expectation values (and hence all probability distributions) are the same in both pictures, exactly as in classical mechanics (Prop. 1.31).

- **(The Schrödinger Equation.)** In the special case that $\rho = P_\psi$ is a vector state, we have $\alpha^H_{-t}(P_\psi) = e^{-itH}|\psi\rangle\langle\psi|e^{itH} = |e^{-itH}\psi\rangle\langle e^{itH}\psi| = P_{U(-t)\psi}$, i.e. a vector $\psi_0 \in \mathbb{C}^N$ describing a vector state changes with time according to $\psi(t) := e^{-itH}\psi_0$. By differentiation, we find

$$i\frac{d}{dt}\psi(t) = H\psi(t). \tag{2.65}$$

This is a linear ordinary first order differential equation for a function taking values in $\mathbb{C}^N$. It is called the *Schrödinger Equation with Hamiltonian H*.

⚛ **(Hamiltonians and Planck's constant $\hbar$.)** As in classical mechanics, the interpretation of the Hamiltonian (matrix) is that of *energy*. Hence its entries and its eigenvalues have the physical dimension of energy, e.g. they are measured in units like Joule J or electron volt eV$\cong 1.6 \cdot 10^{-19}$J. Time, on the other hand, is measured in seconds s. Hence tH has physical units Js, and to get well-defined exponentials e^{itH}, we need to divide the exponent by a number of units Js. This is done by Planck's constant $\hbar \cong 10^{-34}$Js, i.e. all exponentials e^{itH} should be replaced by $e^{itH/\hbar}$. We have set $\hbar = 1$, which amounts to measuring energy in units s^{-1}, i.e. as a frequency.

Let us reformulate the existence and uniqueness properties of solutions of Heisenberg's and Schrödinger's Equation.

Proposition 2.29 (Initial value problems for Heisenberg and Schrödinger Equation). *Let $(H, M_N(\mathbb{C}))$ be a matrix mechanical system.*

a) For $A_0 \in M_N(\mathbb{C})$, consider the initial value problem for Heisenberg's Equation,

$$\frac{d}{dt}A(t) = -i[A(t), H], \qquad A(0) = A_0. \tag{2.66}$$

This problem has a unique solution, namely $A(t) = \alpha^H_t(A_0) = e^{itH}A_0e^{-itH}$.

b) For $\psi_0 \in \mathbb{C}^N$, consider the initial value problem for Schrödinger's Equation,

$$\frac{d}{dt}\psi(t) = -iH\psi(t), \qquad \psi(0) = \psi_0. \tag{2.67}$$

This problem has a unique solution, namely $\psi(t) = e^{-itH}\psi_0$.

Proof. a) The proof has basically already been given in Theorem 2.27 and just repeated for clarity. From Theorem 2.27, we know alredy that $A(t) = e^{itH}A_0e^{-itH}$ is a solution of the initial value problem (2.66). Suppose $t \mapsto B(t)$ is another solution, and define

$$C(t) := e^{-itH}B(t)e^{itH}.$$

Then (the derivative rules used here are proven precisely as in Theorem 2.27)

$$\begin{aligned} C'(t) &= -iHe^{-itH}B(t)e^{itH} + e^{-itH}B'(t)e^{itH} + ie^{-itH}B(t)He^{itH} \\ &= e^{-itH}\left(B'(t) - iHB(t) + iB(t)H\right)e^{itH} \\ &= e^{-itH}\left(B'(t) + i[B(t), H]\right)e^{itH} \\ &= 0. \end{aligned}$$

As $C(0) = A_0$, we conclude $C(t) = A_0$ for all $t \in \mathbb{R}$. This is equivalent to $B(t) = e^{itH}A_0e^{-itH}$.

b) By differentiation, one finds directly that $t \mapsto e^{-itH}\psi_0$ is a solution to the initial value problem (2.67). Assume that $t \mapsto \phi(t)$ is another solution, and define $\xi(t) := e^{itH}\phi(t)$. Then

$$\xi'(t) = iHe^{itH}\phi(t) + e^{itH}\phi'(t) = e^{itH}\left(iH\phi(t) + \phi'(t)\right) = 0,$$

so ξ is constant. As $\xi(0) = \psi_0$, we conclude $\xi(t) = \psi_0$ for all $t \in \mathbb{R}$, which implies $\phi(t) = e^{-itH}\psi_0$. □

The following corollary is left as an exercise (Problem 2.9).

Corollary 2.30 (Hamiltonians and unitary one-parameter groups). *There exists a bijection between selfadjoint elements $H \in M_N(\mathbb{C})$ and continuous unitary one-parameter groups $(U(t))_{t\in\mathbb{R}}$ in $M_N(\mathbb{C})$, given by*

$$H \mapsto U_H(t) := e^{itH}, \qquad (U(t))_{t\in\mathbb{R}} \mapsto H := -i\left.\frac{d}{dt}\right|_{t=0} U(t). \tag{2.68}$$

At this point we have found the quantum (matrix) analogues of states, Hamiltonians, and dynamics. In view of the commutator with H times $-i$ appearing as the derivation of the dynamics given by H, we also see that this should be considered the analogue of the classical Poisson bracket.

As in classical mechanics, it is now meaningful to consider properties that do not change with time. As we may consider the dynamics either in the Heisenberg or the Schrödinger picture, we have reason to define two dual concepts: Constants of motion (time-independent observables in the Heisenberg picture) and stationary states (time-independent states in the Schrödinger picture).

Definition 2.31 (Constants of motion and stationary states). Let $(M_N(\mathbb{C}), H)$ be a matrix mechanical system.

a) A *constant of motion* is a matrix $A \in M_N(\mathbb{C})$ such that

$$\alpha_t^H(A) = A, \qquad t \in \mathbb{R}. \tag{2.69}$$

b) A *stationary state* is a state ω_ρ such that

$$\omega_\rho \circ \alpha_t^H = \omega_\rho, \qquad t \in \mathbb{R}. \tag{2.70}$$

In stationary states, the expectation values of *all* observables are time-independent, hence the name "stationary". In non-stationary states, *some* observables have time-independent

expectation values. This is in particular true for the constants of motion, which have time-independent expectation values in *all* states.

The first parts of the following proposition are almost identical to its classical counterpart Lemma 1.19. We also include a version of Noether's Theorem (Theorem 1.22) in part d).

Proposition 2.32 (Constants of motion in matrix mechanics). *Let $(M_N(\mathbb{C}), H)$ be a matrix mechanical system.*

a) $A \in M_N(\mathbb{C})$ is a constant of motion if and only if

$$[A, H] = 0. \tag{2.71}$$

b) H is a constant of motion.

c) The constants of motion form a unital *-algebra.

d) A selfadjoint matrix $A = A^*$ is a constant of motion if and only if the unitary one-parameter group $(e^{isA})_{s\in\mathbb{R}}$ defines symmetries $\alpha_s^A(B) = e^{isA}Be^{-isA}$, $B \in M_N(\mathbb{C})$.

Proof. b) follows because H commutes with e^{-itH}, namely

$$\alpha_t^H(H) = e^{itH}He^{-itH} = e^{itH}e^{-itH}H = H.$$

a) We consider Heisenberg's Equation and rewrite it with the help of the automorphism property and b) as

$$\frac{d}{dt}\alpha_t^H(A) = -i[\alpha_t^H(A), H] = -i\alpha_t^H([A, H]). \tag{2.72}$$

In this form, we see that $t \mapsto \alpha_t^H(A)$ being constant is equivalent to $[A, H] = 0$, as claimed.

c) is only saying that in case A and B commute with H, then also all polynomials in $A, B, A^*, B^*, 1$ do so.

d) As in our discussion of dynamics, it follows that $(e^{isA})_{s\in\mathbb{R}}$ is a one-parameter group of unitaries. These unitaries induce symmetries of $(M_N(\mathbb{C}), H)$ if and only if $e^{isA}He^{-isA} = H$ for all $s \in \mathbb{R}$, i.e. $0 = [H, A]$ by part a). □

Stationary states can be characterized analogously, as we show next.

Corollary 2.33 (Characterization of stationary states). *Let $(M_N(\mathbb{C}), H)$ be a matrix mechanical system and ω_ρ a state, $\rho \in \mathcal{D}_N$.*

a) ω_ρ is stationary if and only if ρ is a constant of motion.

b) If $\rho = P_\psi$ is a rank one projection (i.e. if ω_ρ is vector state), then it is stationary if and only if ψ is an eigenvector of H.

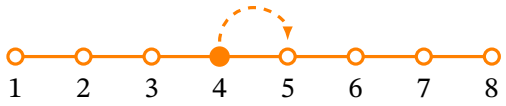
Proof. a) Since $\omega_\rho \circ \alpha_t^H = \omega_{\alpha_{-t}^H(\rho)}$, the state ω_ρ is stationary if and only if $\alpha_{-t}^H(\rho) = \rho$, which is equivalent to $[\rho, H] = 0$ by Proposition 2.32 a).

b) Let $\rho = P_\psi$ be a rank one projection. If ψ is a normalized eigenvector of H with eigenvalue λ, then $\alpha_t^H(P_\psi) = P_{e^{-it\lambda}\psi} = P_\psi$ (here we have used $P_\psi = P_{z\psi}$ for $z \in \mathbb{C}$, $|z| = 1$), so ω_ψ is stationary. If, conversely, ω_ψ is stationary, then we have $0 = [P_\psi, H]$ and in particular

$$0 = [P_\psi, H]\psi = \langle\psi, H\psi\rangle\psi - H\psi,$$

which shows that ψ is an eigenvector of H, with eigenvalue $\omega_\psi(H) = \langle\psi, H\psi\rangle$. □

Example 2.34 (A particle hopping on a one-dimensional lattice). In this example we revisit Example 2.19 of a particle on a one-dimensional finite lattice consisting of N sites. In addition to the position observable $X = \sum_{k=1}^{N} kE_{kk}$, we now add a dynamics by selecting a Hamiltonian H. The idea is to choose H in such a way that the particle can "hop" with non-zero probabilities on the lattice.



We impose periodic boundary conditions as in Problem 2.4, i.e. we identify $e_{N+1} = e_1$ and label the orthonormal basis mod N, and consider the "discrete rotation"

$$V = \sum_{k=1}^{N-1} |e_{k+1}\rangle\langle e_k|. \tag{2.73}$$

As shown in Problem 2.4, V is unitary and satisfies $Ve_j = e_{j+1}$. Its eigenvalues are

$$\sigma(V) = \{\nu_1, \dots, \nu_N\}, \qquad \nu_k := \omega^k, \qquad \omega := e^{2\pi i/N}, \tag{2.74}$$

and the vector

$$\xi_k := \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \omega^{-jk} e_j \tag{2.75}$$

is a normalized eigenvector of V with eigenvalue ν_k. Hence the spectral decomposition of V is $V = \sum_{k=0}^{N-1} \omega^k |\xi_k\rangle\langle\xi_k|$. Furthermore, it was shown that

$$V^*XV = X + 1 - NP_{e_N} \iff [X, V] = V - NVP_{e_N} = V - N|e_1\rangle\langle e_N| \tag{2.76}$$

holds.

We now choose

$$H := V + V^* = \sum_{k=1}^{N-1} (|e_k\rangle\langle e_{k+1}| + |e_{k+1}\rangle\langle e_k|) \tag{2.77}$$

as our (selfadjoint) Hamiltonian matrix and consider the time evolution $\alpha_t^H(X)$ of the position observable X.

a) We have

$$\begin{aligned} [X, H] &= [X, V + V^*] \\ &= [X, V] - [X, V]^* \\ &= V - N|e_1\rangle\langle e_N| - V^* + N|e_N\rangle\langle e_1| \\ &= V - V^* - N(|e_1\rangle\langle e_N| - |e_N\rangle\langle e_1|), \end{aligned}$$

which is different from 0. Hence X is not a constant of motion, so there exist states ω such that $\omega(\alpha_t^H(X))$ is not constant in t.

b) Since $H = V + V^*$, we see directly by the spectral mapping principle that the spectrum of H is

$$\sigma(H) = \{2\,\mathrm{Re}\,\lambda \;:\; \lambda \in \sigma(V)\} = \{\lambda_1, \dots, \lambda_N\}, \qquad \lambda_k = 2\cos\frac{2\pi k}{N}, \quad (2.78)$$

with the same eigenvectors ξ_k as V (here we have used that V is normal (even unitary)). In other words, $H = \sum_{j=1}^N \lambda_k |\xi_k\rangle\langle\xi_k|$ is the spectral decomposition of H.

c) We claim that the time evolution of X is

$$\alpha_t^H(X) = \frac{N+1}{2}\,1 + \sum_{\substack{k,l \\ k\neq l}}^{N} e^{it(\lambda_k-\lambda_l)} \frac{\omega^{k-l}}{\omega^{k-l}-1} |\xi_k\rangle\langle\xi_l|. \quad (2.79)$$

The proof is a direct calculation. We first calculate

$$\langle \xi_k, X\xi_l\rangle = \frac{1}{N}\sum_{j,n} \overline{\omega^{-jk}}\omega^{-nl}\langle e_j, Xe_l\rangle = \frac{1}{N}\sum_{j=0}^{N-1} j\omega^{(k-l)j} = \begin{cases} \frac{N+1}{2} & k = l \\ \frac{\omega^{k-l}}{\omega^{k-l}-1} & k \neq l \end{cases}.$$

Inserted into

$$\alpha_t^H(X) = e^{itH}Xe^{-itH} = \sum_{k,l} e^{it(\lambda_k-\lambda_l)} |\xi_k\rangle\langle\xi_k, X\xi_l\rangle\langle\xi_l|,$$

this yields the claimed formula.

In the following plots, you see the probability distribution for position X at times $t = 0$, $t = 2$, $t = 5$, in different vector states ψ (for $N = 40$). On the horizontal axis, you see the spectrum of X, i.e. the set $\{1, \dots, 40\}$, and at position k, the probability

$$\mathrm{Pr}_\psi(\alpha_t^H(X) = k) = \|(\alpha_t^H(E^X(k))\psi\|^2 = \|e^{itH}|e_k\rangle\langle e_k, e^{-itH}\psi\rangle\|^2 = |\langle e_k, e^{-itH}\psi\rangle|^2$$

is plotted.

(row 1) $\psi = e_{21}$, an eigenvector of X. The distribution is initially a Dirac measure at 21 and then spreads with time, similar to the behaviour found in Example 1.32 in classical mechanics.

(row 2) $\psi = \xi_6$, an eigenvector of H with eigenvalue λ_6. As explained in Corollary 2.33, the vector state given by ξ_6 is stationary, so the plots of the probability distributions look identical at different times. In more detail, we have

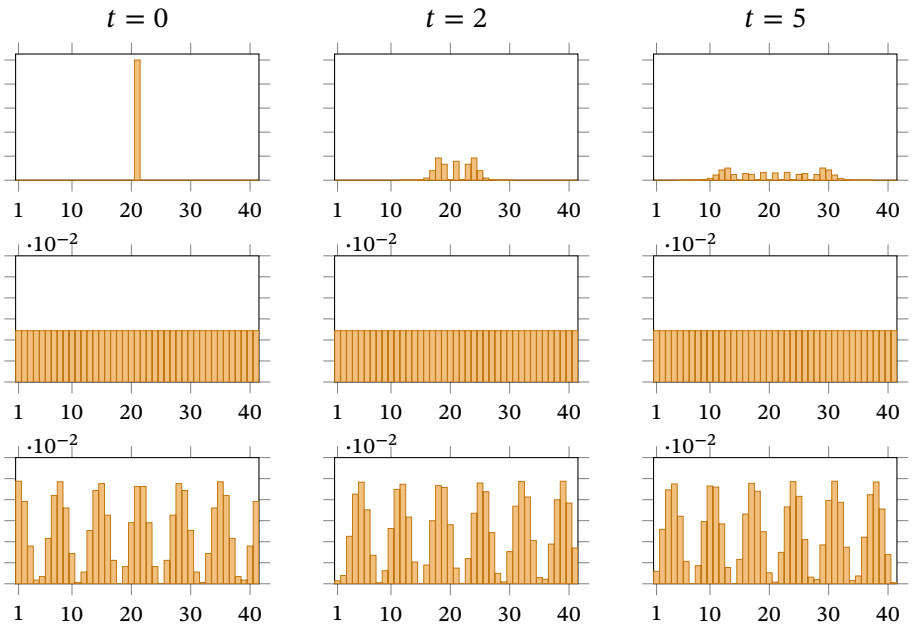
$$|\langle e_k, e^{-itH}\xi_6\rangle|^2 = \left|\left\langle e_k, \frac{1}{\sqrt{N}}\sum_{j=1}^N \omega^{-6j} e_j\right\rangle\right|^2 = \frac{1}{N},$$

i.e. a time-independent uniform distribution over $\{1, \dots, N\}$.

(row 3) $\psi = \frac{1}{\sqrt{2}}(\xi_6 + \xi_{12})$, a superposition of two energy eigenstates (eigenvectors of H) with different eigenvalues λ_6 and λ_{12}. In this case, we obtain an oscillatory behaviour, with frequency $\lambda_6 - \lambda_{12}$. To see why this is the case, observe that for this vector ψ,

$$\begin{aligned}
|\langle e_k, e^{-itH}\psi\rangle|^2 &= \frac{1}{2}|\langle e_k, e^{-it\lambda_6}\xi_6 + e^{-it\lambda_{12}}\xi_{12}\rangle|^2 \\
&= \frac{1}{2}|\langle e_k, \xi_6\rangle|^2 + \frac{1}{2}|\langle e_k, \xi_{12}\rangle|^2 + \mathrm{Re}\left(e^{-it\lambda_6}e^{it\lambda_{12}}\langle e_k, \xi_6\rangle\langle \xi_{12}, e_k\rangle\right) \\
&= \frac{1}{2}|\langle e_k, \xi_6\rangle|^2 + \frac{1}{2}|\langle e_k, \xi_{12}\rangle|^2 + \cos(t(\lambda_6 - \lambda_{12}))\,\mathrm{Re}(\langle e_k, \xi_6\rangle\langle \xi_{12}, e_k\rangle) \\
&\quad + \sin(t(\lambda_6 - \lambda_{12}))\,\mathrm{Im}(\langle e_k, \xi_6\rangle\langle \xi_{12}, e_k\rangle).
\end{aligned}$$

Such a oscillatory pattern is reminiscent of waves rather than particles. It is unexpected here from a classical perspective, but not in direct contradiction because we have no classical model to compare this example with.



On StudOn, you can find an animated version of these plots.

The following table continues Table 1 with a comparison of classical and matrix mechanics w.r.t. dynamics and symmetries.

2.4 Rotational symmetry and spin

In classical mechanics, Noether's Theorem (Theorem 1.22) gave us a correspondence between geometric symmetries, like spatial translations, and conserved quantities, like total momentum (Example 1.23). While we have no good model of continuous translational symmetry or momentum in our current finite-dimensional matrix mechanics, the example of rotational symmetry and angular momentum conservation (Example 1.24, Definition 1.25) can be investigated in matrix mechanics.

Aspect	Classical mechanics	Matrix mechanics
Dynamics	one-parameter group of automorphisms α_t^H generated by a Hamiltonian function H	one-parameter group of automorphisms α_t^H generated by a Hamiltonian matrix H
Equation of motion	Hamilton's equation $\frac{d}{dt}\alpha_t^H(f) = \{\alpha_t^H(f), H\}$	Heisenberg's equation $\frac{d}{dt}\alpha_t^H(A) = -i[\alpha_t^H(A), H]$ In Schröd. pic.: Schrödinger's eqn. $\frac{d}{dt}\psi(t) = -iH\psi(t)$
Lie bracket	Poisson bracket $\{f, g\}$	commutator $-i[A, B]$
Constants of motion	characterized by $\{f, H\} = 0$	characterized by $[A, H] = 0$
Automorphisms	symplectic diffeomorphisms (as a unital Poisson *-algebra)	PU(N), $\alpha_U(A) = UAU^*$
Dynamical system	Hamiltonian system (M, H)	matrix mechanical system $(M_N(\mathbb{C}), H)$
Symmetry of dyn. system	Auto. α with $\alpha(H) = H$	Auto. α with $\alpha(H) = H$

Table 2: Similarities and differences between classical mechanics and matrix mechanics: Dynamics.

We therefore consider the rotation group (special orthogonal group)

$$G = \mathrm{SO}(3) = \{R \in M_3(\mathbb{R}) \,:\, R^T = R^{-1}, \quad \det R = 1\}. \tag{2.80}$$

Lemma 2.35 ($SO(3)$).
a) Let $x \in \mathbb{R}^3$, $\|x\| = 1$, and $\theta \in (-\pi, \pi]$. Choose an orthonormal basis (y, z) of the orthogonal complement $\{x\}^\perp \cong \mathbb{R}^2$. Then the linear map $R_x(\theta) : \mathbb{R}^3 \to \mathbb{R}^3$ that in the orthonormal basis (x, y, z) is given by

$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \tag{2.81}$$

is an element of SO(3).
b) Every $R \in \mathrm{SO}(3)$, $R \neq 1$, is of the form described in a), with unique x and $\theta \in (0, \pi]$.

Proof. a) It is easy to check that (2.81) is orthogonal and has determinant 1. As these properties are stable under change of orthonormal basis, it is also an element of SO(3) when exprесed in the canonical basis of $\mathbb{R}^3$.

b) Let $R \in \mathrm{SO}(3)$. We begin by noting that since R is real, its characteristic polynomial is real and hence its zeros (the eigenvalues of R) occur in conjugate pairs $\lambda, \bar{\lambda}$. Thus the spectrum of R has the form $\sigma(R) = \{\lambda, \bar{\lambda}, \mu\}$. As R is in particular unitary, we have $|\lambda| = |\mu| = 1$

and

$$1 = \det R = \lambda\bar{\lambda}\mu = \mu, \tag{2.82}$$

i.e. R has eigenvalue 1. We also see that in case the multiplicity of 1 is larger than 1, then it is 3, i.e. $R = 1$. So up to the case $R = 1$, the eigenspace with eigenvalue 1 is a line, and contains two unit vectors x and $-x$. Since R is orthogonal, it preserves the orthogonal complement $\{x\}^\perp \cong \mathbb{R}^2$, and acts in this two-dimensional plane by an SO(2)-matrix, say

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SO}(2). \tag{2.83}$$

Then $\det A = ad - bc = 1$, and

$$A^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix} = A^{-1} = \frac{1}{\det A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \implies A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad a^2 + b^2 = 1. \tag{2.84}$$

There exists unique $\theta \in (-\pi, \pi]$ such that $a = \cos\theta, b = \sin\theta$. There is a slight redundancy in this description: If we choose $-x$ instead of x as our basis vector, and rotate by the negative angle $-\theta$ instead of θ, we obtain the same transformation. Hence restricting the angle to $(0, \pi]$ but allowing all axis (x and $-x$) yields a bijective description of SO(3). □

The eigenvalue 1 eigenvector x plays the role of the axis of rotation, and θ the angle by which R rotates around x. Note that in this parameter description, we may think of SO(3) as the ball $B = \{y \in \mathbb{R}^3 \ : \ \|y\| \leq \pi\}$, where $y \in B$ corresponds to rotation axis $x := y/\|y\|$ (unless $y = 0$, which corresponds to $R = 1$), and $\|y\| = \theta$ to the rotation angle. However, since rotations around a fixed axis are 2π-periodic, we have to identify the antipodal points y and $-y$ for $\|y\| = \pi$. This gives a more complicated looking topological space $B/\!\sim \cong \mathrm{SO}(3)$, which in particular is not simply connected.

The rotations $R_x(\theta)$, $\theta \in \mathbb{R}$, around a fixed axis x, form continuous ((2π)-periodic) unitary one-parameter groups and are hence of the form $R_x(\theta) = e^{i\theta L_x}$ for some $L_x = L_x^* \in M_3(\mathbb{C})$ (Corollary 2.30). In particular, if we take the rotation axis to be one of the canonical basis vectors e_k, then we have

$$\begin{aligned} R_1(\theta) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}, \qquad R_2(\theta) = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}, \\ R_3(\theta) &= \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{2.85}$$

Example 2.36 (Spin 1). As explained above, we have $R_k(\theta) = e^{i\theta L_k}$ for some $L_k = L_k^* \in M_3(\mathbb{C})$. To determine them, we take the derivative at angle $\theta = 0$ of the matrices (2.85), and obtain

$$L_1 = -i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad L_2 = -i\begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad L_3 = -i\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.86}$$

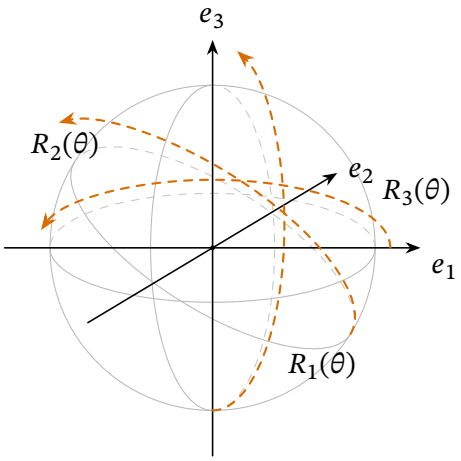


Figure 3: One-parameter groups of rotations around the coordinate axes.

A direct calculation shows that each L_j has spectrum $\{-1, 0, +1\}$, and the equations

$$L_3 = -i[L_1, L_2], \qquad L_1 = -i[L_2, L_3], \qquad L_2 = -i[L_3, L_1] \tag{2.87}$$

and

$$L_1^2 + L_2^2 + L_3^2 = 2 \cdot \mathbb{1} \tag{2.88}$$

(2 times the identity matrix) hold.

We note that the equations (2.87) are precisely analogous to the classical Poisson brackets of the components of angular momentum (Exercise 1.5). In classical mechanics, angular momentum was a vector observable. In that analogy, $L_1^2 + L_2^2 + L_3^2$ corresponds to the length (squared) of the angular momentum vector, the "total angular momentum". In the example above, total angular momentum is $\sqrt{2}$ times the identity. In general, we make the following definition.

Definition 2.37 (Angular momentum in matrix mechanics). A system of *angular momentum operators*[a] in $M_N(\mathbb{C})$ is a triple (L_1, L_2, L_3) of selfadjoint matrices in $M_N(\mathbb{C})$ such that

$$L_3 = -i[L_1, L_2], \qquad L_1 = -i[L_2, L_3], \qquad L_2 = -i[L_3, L_1], \tag{2.89}$$

holds. The *total angular momentum* of this system is defined as

$$L := (L_1^2 + L_2^2 + L_3^2)^{1/2}. \tag{2.90}$$

[a] For those who know Lie algebras: This is a representation of the Lie algebra of $\mathrm{SO}(3)$.

Of course the matrices from Example 2.36 form a system of angular momentum operators. Another example is the following one.

Example 2.38 (Pauli matrices and spin $\frac{1}{2}$). In $M_2(\mathbb{C})$, consider the *Pauli matrices*

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.91}$$

often also called $\sigma_x, \sigma_y, \sigma_z$ instead of $\sigma_1, \sigma_2, \sigma_3$.

a) For each $j \in \{1,2,3\}$, we have

$$\sigma_j = \sigma_j^*, \qquad \operatorname{Tr}\sigma_j = 0, \qquad \sigma_j^2 = 1, \qquad \sigma(\sigma_j) = \{1,-1\}. \tag{2.92}$$

b) The real span of the Pauli matrices is

$$\operatorname{span}_{\mathbb{R}}\{\sigma_1,\sigma_2,\sigma_3\} = \{X \in M_2(\mathbb{C}) \,:\, X = X^*,\ \operatorname{Tr} X = 0\}, =: i\mathfrak{su}(2). \tag{2.93}$$

As a shorthand, we write $S(x) := x \cdot \sigma := x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3$ for $x \in \mathbb{R}^3$ and note that $S : \mathbb{R}^3 \to i\mathfrak{su}(3)$ is a linear bijection with

$$\det(S(x)) = -\|x\|^2. \tag{2.94}$$

c) The Pauli matrices satisfy the commutation relations

$$-i[\sigma_1,\sigma_2] = 2\sigma_3 \tag{2.95}$$

and the corresponding cyclic rearrangements of this. This does not quite match (2.89) because of the factor 2. But the rescaled matrices

$$L_j := \frac{1}{2}\sigma_j, \qquad j = 1,2,3, \tag{2.96}$$

do form a system of angular momentum operators in $M_2(\mathbb{C})$. As $\sigma_j^2 = 1$, the total angular momentum of this system is

$$L^2 = \frac{3}{4}\cdot \mathbb{1}. \tag{2.97}$$

We now consider general features of angular momentum operators and first derive some commutation relations.

Lemma 2.39 (Properties of angular momentum operators). *Let (L_1,L_2,L_3) be a system of angular momentum operators.*

a) $[L^2, L_j] = 0$ *for each* $j \in \{1,2,3\}$.

b) Define $L_\pm := L_1 \pm iL_2$. *Then* $L_\pm^* = L_\mp$ *and*

$$[L_3, L_\pm] = \pm L_\pm, \tag{2.98}$$
$$[L_+, L_-] = 2L_3, \tag{2.99}$$
$$L_-L_+ = L^2 - L_3(L_3+1), \qquad L_+L_- = L^2 - L_3(L_3-1). \tag{2.100}$$

Proof. a) We show the argument for $j = 1$:

$$\begin{aligned}[L^2, L_1] &= [L_1^2, L_1] + [L_2^2, L_1] + [L_3^2, L_1] \\ &= 0 + L_2[L_2, L_1] + [L_2, L_1]L_2 + L_3[L_3, L_1] + [L_3, L_1]L_3 \\ &= -iL_2L_3 - iL_3L_2 + iL_3L_2 + iL_2L_3 = 0.\end{aligned}$$

b) $L_\pm^* = L_\mp$ is clear by selfadjointness of L_1 and L_2. The other equations follow from

$$\begin{aligned}[L_3, L_\pm] &= [L_3, L_1] \pm i[L_3, L_2] = iL_2 \pm i(-i)L_1 = \pm(L_1 \pm iL_2) = \pm L_\pm, \\ [L_+, L_-] &= [L_1 + iL_2, L_1 - iL_2] = -i[L_1, L_2] + i[L_2, L_1] = 2L_3, \\ L_\mp L_\pm &= (L_1 \mp iL_2)(L_1 \pm iL_2) = L_1^2 + L_2^2 \mp iL_2L_1 \pm iL_1L_2 = L^2 - L_3^2 \mp L_3.\end{aligned}$$

□

The fact that the relations in b) are not symmetric in L_1, L_2, L_3 can be explained as follows. According to a), L^2 commutes with all L_j. Picking one j, say $j = 3$, we may diagonalize L^2 and L_3 simultaneously. But L_1 and L_2 do note commute with L_3, hence non-trivial commutation relations with these operators follow.

In the examples considered so far, L^2 was a multiple of the identity. (This is a consequence of an irreducible representation.) We now investigate what we can say about such a situation in general.

Theorem 2.40 (Angular momentum operators in finite dimension). *Let (L_1, L_2, L_3) be a system of angular momentum operators in $M_N(\mathbb{C})$ such that L^2 is a multiple of the identity, say $L^2 = l(l+1)$ for some $l \geq 0$.*

a) $l \in \frac{1}{2}\mathbb{N}_0$ and $\sigma(L_3) = \{-l, -l+1, \dots, l-1, l\}$.

b) If (L_1, L_2, L_3) is irreducible in the sense that there exists no proper subspace $V \subset \mathbb{C}^N$ which is invariant under the three operators L_1, L_2, L_3 simultaneously[a], then the dimension is $N = 2l + 1$.

[a] Irreducibility implies that L^2 is a multiple of the identity: Since L^2 commutes with L_1, L_2, L_3, its spectral projections $E^{L^2}(\lambda)$ also commute with these operators. Hence the range of $E^{L^2}(\lambda)$ is an invariant subspace. As this must be trivial, it follows that L^2 has only one eigenvalue, namely $L^2 = \lambda \mathbb{1}$.

Proof. Let ψ be a normalized eigenvector of L_3, with eigenvalue $m \in \mathbb{R}$. Using the commutation relations from the previous lemma, we find

$$\begin{aligned}&L_-L_+\psi = (L^2 - L_3(L_3+1))\psi = (l(l+1) - m(m+1)) \cdot \psi = (l-m)(l+m+1) \cdot \psi, \\ &L_+L_-\psi = (L^2 - L_3(L_3-1))\psi = (l+m)(l-m+1) \cdot \psi \\ \implies &\|L_\pm\psi\|^2 = \langle \psi, L_\mp L_\pm \psi \rangle = (l \mp m)(l \pm m + 1).\end{aligned}$$

As a norm is non-negative, we see that we get $-l \leq m \leq l$. In the same way, one sees

$$L_\pm\psi = 0 \iff m = \pm l.$$

Now it is important to observe that

$$L_3L_\pm\psi = [L_3, L_\pm]\psi + L_\pm L_3\psi = (m \pm 1)L_\pm\psi.$$

Thus either $L_\pm\psi$ is an eigenvector of L_3, with eigenvalue $m \pm 1$, or $L_\pm\psi = 0$, which is equivalent to $m = \pm l$. If we apply $L_\pm$ repeatedly, we obtain eigenvalues $m+1, m+2, m+3, \dots$ and $m-1, m-2, m-3, \dots$ unless $L_\pm^k\psi = 0$ for some k. Since L_3 has only finitely many eigenvalues, we conclude that $L_+^{k+1}\psi = 0$ and $L_-^{k'+1} = 0$ for some $k, k' \in \mathbb{N}_0$. This implies $\pm l \in \sigma(L_3)$ and $m + k = l$, $m - k' = -l$, namely $m = \frac{1}{2}(k' - k) \in \frac{1}{2}\mathbb{Z}$. As l differs from m by an integer, we have shown $l \in \frac{1}{2}\mathbb{N}_0$, and the raising/lowering procedure gives $\sigma(L_3) = \{-l, \dots, l\}$.

To show the second part, observe that the eigenvectors of L_3 with eigenvalues $m \in \{-l, \dots, l\}$ are linearly independent. Their span $V \subset \mathbb{C}^N$ is invariant under $L_\pm = L_1 \pm iL_2$ by construction, hence under L_1, L_2, L_3. Irreducibility implies $V = \mathbb{C}^N$, hence $N = \dim V = 2l + 1$. □

The examples discussed so far correspond to the cases $l = 1$ (dimension 3) and $l = \frac{1}{2}$ (dimension 2). Even simpler is the situation with $l = 0$ (dimension 1), which is realized in the trivial matrix algebra $M_1(\mathbb{C}) = \mathbb{C}$.

We state here without proof that for any $l \in \frac{1}{2}\mathbb{N}_0$, there exists an irreducible system of angular momentum operators that is unique up to unitary equivalence. A system of angular momentum operators with a given value $l \in \frac{1}{2}\mathbb{N}_0$ (or the corresponding (projective) representation of SO(3), see below) is called *spin l representation*.

Angular momentum and spin. In our later investigations of quantum particles moving in $\mathbb{R}^3$, we will see vector states of varying angular momenta $l \in \mathbb{N}_0$ (integer valued). This angular momentum is analogous to the one found in classical mechanics and describes the behaviour of the state under rotations in space.

There is, however, a second way in which angular momentum operators occur in quantum mechanics: Each elementary particle carries, besides its mass, a *spin s*, which can be thought of as an inner degree of freedom. As a rough analogy, one may picture the particle as spinning around like a top. But there is no classical explanation of spin and hence no convincing illustration of it. Nonetheless, without the assumption of spin the predictions of the theory would deviate from experimental data, so that spin is now a universally accepted concept.

This "internal angular momentum" (spin) can take integer and half-integer values and depends on the type of particle considered. There are many particle types of spin $\frac{1}{2}$, like electrons, protons, neutrons, neutrinos, quarks. The Higgs boson has spin 0. Photons and gluons have spin 1 (one actually needs to speak about helicity rather than spin because these are massless particles, but we do not into this further), the massive W and Z bosons mediating the weak nuclear force have spin 1, a hypothetical graviton particle is thought to have spin 2 (if it exists), and composite particles can have higher spin values.

Because of the frequent appearance of spin 1/2 (all known matter particles have spin 1/2), this is a particularly important case. The mathematical description of spin 1/2 particles involves the Pauli matrices from Example 2.38. As an analogy, one may think of a spin 1/2 particle as spinning around the e_3-axis in two possible directions: spin up, i.e. eigenvalue $+1/2$ of L_3 respectively eigenvalue $+1$ of σ_3, and spin down, i.e. eigenvalue $-1/2$ of L_3 respectively eigenvalue -1 of σ_3. This also explains the physics

notation

$$\begin{pmatrix}1\\0\end{pmatrix} =: |\uparrow\rangle \qquad \text{spin up,} \tag{2.101}$$

$$\begin{pmatrix}0\\1\end{pmatrix} =: |\downarrow\rangle \qquad \text{spin down,} \tag{2.102}$$

for the eigenvectors of σ_3.

Particles with half-integer spin $s \in \mathbb{N}_0 + \frac{1}{2}$ are called *fermions* and particles with integer spin $s \in \mathbb{N}_0$ are called *bosons*. Bosons and fermions behave in a fundamentally different way that is however only visible when considering multi-particle systems (next section).

In general, angular momentum describing motion in $\mathbb{R}^3$ and spin combine into more complex angular momentum phenomena. We will see this later.

After this discussion of angular momentum and spin, we come back to the rotation group SO(3) and have another look at Example 2.38 with the Pauli matrices. In this example, the relation to SO(3) is not obvious. But just as in Example 2.36, we may look at the unitary one-parameter groups generated by the angular momentum operators, $e^{i\theta\sigma_j/2}$, $\theta \in \mathbb{R}$. As σ_j is selfadjoint with eigenvalues ± 1, the matrices $e^{i\theta\sigma_j/2}$ are unitary with eigenvalues $e^{\pm i\theta/2}$ and in particular with determinant 1. Hence they are elements of the group

$$\mathrm{SU}(2) = \{U \in M_2(\mathbb{C}) \,:\, U^* = U^{-1}, \quad \det U = 1\}. \tag{2.103}$$

It is easy to see that every element of SU(2) has the form

$$U = \begin{pmatrix}\alpha & \beta\\ -\overline{\beta} & \overline{\alpha}\end{pmatrix}, \qquad |\alpha|^2 + |\beta|^2 = 1. \tag{2.104}$$

Clearly α and β are uniquely determined by U, and any $\alpha, \beta \in \mathbb{C}$ with $|\alpha|^2 + |\beta|^2 = 1$ fixes a unique $U \in \mathrm{SU}(2)$. As a topological space, we may therefore identify $\mathrm{SU}(2) \cong S^3$, the 3-sphere in $\mathbb{R}^4$ (here we read α, β in terms of their real and imaginary parts). This description shows in particular that SU(2) is connected and simply connected.[9]

For the further developments, it is very helpful to understand the relation between SU(2) and SO(3).

Proposition 2.41 (SU(2) and SO(3)). *We consider the linear bijection $S : \mathbb{R}^3 \to i\mathfrak{su}(2)$, $x \mapsto x \cdot \sigma =: S(x)$ (Example 2.38), and the map ρ defined by*

$$\rho : \mathrm{SU}(2) \to \mathrm{SO}(3), \qquad S(\rho(U)x) := USU^*, \quad x \in \mathbb{R}^3. \tag{2.105}$$

a) ρ is a well-defined continuous group homomorphism with $\ker\rho = \{+1, -1\}$, i.e. $\rho(U) = \rho(V) \iff U = \pm V$.

b) ρ is surjective.

[9] Our discussion of topological groups is quite rudimentary here – if you want to learn more about this, you should attend a lecture on Lie groups.

Proof. a) For $x \in \mathbb{R}^3$, the matrix $S(x)$ is selfadjoint and traceless (this means it has trace zero). These properties are preserved under conjugation with a unitary $U \in \mathrm{SU}(2)$. As S is a bijection, we see that (2.105) defines a map $\rho(U) : \mathbb{R}^3 \to \mathbb{R}^3$ which is linear because S is linear.

We next show that $\rho(U) \in \mathrm{SO}(3)$ for $U \in \mathrm{SU}(2)$. Recalling $\det S(x) = -\|x\|^2$ (example 2.38), we find

$$-\|x\|^2 = \det S(x) = \det(US(x)U^*) = \det(S(\rho(U)x)) = -\|\rho(U)x\|^2,$$

and conclude that $\rho(U)$ is an isometry, i.e. $\rho(U) \in O(3)$. The map ρ is continuous[10], and $\rho(1) = 1$ has determinant 1. As $\mathrm{SU}(2)$ is connected and $R \in O(3)$ can only have determinant $+1$ or -1, it follows that $\rho(U)$ has determinant $+1$, i.e. ρ maps $\mathrm{SU}(2)$ into $\mathrm{SO}(3)$.

For the homomorphism property, note that for any $x \in \mathbb{R}^3$, $U, V \in \mathrm{SU}(2)$,

$$S(\rho(U)\rho(V)x) = US(\rho(V)x)U^* = UVS(x)(UV)^* = S(\rho(UV)x),$$

namely $\rho(UV) = \rho(U)\rho(V)$.

It remains to establish $\ker\rho = \pm 1$. Indeed, let $\rho(U) = 1$. Then $US(x)U^* = S(x)$ for all $x \in \mathbb{R}^3$, i.e. for all matrices of trace zero. As U also commutes with all multiples of the identity matrix, this implies that U commutes with all matrices and must therefore be trivial, $U = c1$. Now $1 = \det U = c^2$ implies $c = \pm 1$.

b) is shown in the exercises (Problem 2.10). □

The map $\rho : \mathrm{SU}(2) \to \mathrm{SO}(3)$ is called *covering homomorphism*. It covers the non-simply connected group $\mathrm{SO}(3)$ in a $2:1$ fashion by the simply connected group $\mathrm{SU}(2)$. While the angular momentum operators of $\mathrm{SO}(3)$ and $\mathrm{SU}(2)$ satisfy the same relations, meaning that these groups look the same locally, they differ in their global structure.

After this digression on $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$ we come back to rotational symmetry in matrix mechanics. Given a matrix algebra $M_N(\mathbb{C})$ (for now, we do not fix a Hamiltonian H), we would like to describe how rotations act by automorphisms of $M_N(\mathbb{C})$, i.e. we want to study group homomorphisms $\mathrm{SO}(3) \to \mathrm{Aut}\, M_N(\mathbb{C})$ – or, as we shall see, group homomorphims $\mathrm{SU}(2) \to \mathrm{Aut}\, M_N(\mathbb{C})$. We therefore write G to denote either $\mathrm{SO}(3)$ or $\mathrm{SU}(2)$ and consider homomorphims $G \to \mathrm{Aut}\, M_N(\mathbb{C}) \cong \mathrm{PU}(N)$. This is equivalent to studying maps

$$U : G \to U(N), \qquad U(g_1)U(g_2) = z(g_1, g_2)U(g_1g_2), \tag{2.106}$$

where $z(g_1, g_2) \in \mathbb{T}$ is a complex number of absolute value 1. Such a map U is called a *projective unitary representation* of G, and a *unitary representation*[11] in case $z(g_1, g_2) = 1$ for all $g_1, g_2 \in G$. As a group homomorphism, a representation is easier to work with than a projective representation and we may wonder whether the study of projective representations can be reduced to the study of honest representations.

Since we are working in the *projective* unitary group, we can freely change the matrices $U(g)$ by multiplicative phases $c(g) \in \mathbb{T}$ (continuous functions $c : G \to \mathbb{T}$), i.e. switch from

[10] This is for example checked with the explicit form $\rho(U)_{jk} = \frac{1}{2}\mathrm{Tr}(\sigma_j U \sigma_k U^*)$ (exercise).

[11] Sometimes we also say "honest" (unitary) representation to emphasize that the phases $z(g_1, g_2)$ are all trivial in this case.

$U(g)$ to $c(g)U(g)$, without changing the action $A \mapsto U(g)AU(g)^{-1}$ on observables. This *might* make it possible to eliminate the factors $z(g_1, g_2)$, because in

$$U(g_1)U(g_2) = \frac{c(g_1g_2)}{c(g_1)c(g_2)} z(g_1, g_2) \cdot U(g_1g_2), \tag{2.107}$$

the factor $\frac{c(g_1g_2)}{c(g_1)c(g_2)} z(g_1, g_2)$ might be 1 for a clever choice of c. In this case, one says that the projective representation U can be lifted to the honest representation cU. Whether this is possible or not depends on simple connectedness, as we sketch now (a full argument is best done with the help of covering theory from topology, and not given here).

Any unitary matrix U has determinant $\det U \in \mathbb{T}$. We now would like to choose $c : G \to \mathbb{T}$ such that $c(g)^N = (\det U(g))^{-1}$ for all $g \in G$, which then implies $\det(c(g)U(g)) = 1$. The problem is to choose c in a continuous manner. Modifying $U(e)$ by a factor if necessary, we may assume $U(e) = 1$, and choose $c(e) = 1$. Now locally (for g varying in a small neighbourhood of e), the function $g \mapsto \det U(g)$ does not cover the full circle $\mathbb{T}$ and hence we find a continuous branch of the logarithm on its range; this gives a continuous $c(g) = \exp -\frac{1}{N} \log \det U(g)$ locally. Continuing this function along a continuous curve γ in G, passing through an arbitrary element g_0 and back to e in a continuous closed loop, we have continuous $g \mapsto c(\gamma(g))$, possibly up to a discontinuity at the end point: We might not arrive back at the value $c(e) = 1$ that we started with. But, if G is simply connected, we can contract γ continuously to a point, which rules out this possibility. Hence a continuous $c : G \to \mathbb{T}$ with $c(g)^N = (\det U(g))^{-1}$ exists.

Once this is done, the modified projective representation $V(g) := c(g)U(g)$ satisfies $\det V(g) = 1$ for all $g \in G$, and taking determinants in $V(g_1)V(g_2) = z(g_1, g_2)V(g_1g_2)$ implies $z(g_1, g_2)^N = 1$ for all $g_1, g_2 \in G$, i.e. z takes values in the discrete set of the N-th roots of unity. If z is continuous (we do not explain here why a continuous choice of z is possible as a consequence of G being simply connected), it must be constant. Looking at $g_1 = g_2 = e$ then gives $z(g_1, g_2) = 1$.

Theorem 2.42. *Any continuous projective unitary representation of* $\mathrm{SU}(2)$ *lifts to an honest representation.*

See [Var06, Chapter 7] for a proof. The argument sketched above suggests that simple connectedness of G is crucial. Indeed, in the non-simply connected group $\mathrm{SO}(3)$, there exist projective representations that can not be made into actual representations by multiplying with phase factors.

To see this explicitly, we now consider continuous unitary representations of $\mathrm{SU}(2)$ or $\mathrm{SO}(3)$ (and not just projective ones). These define angular momentum operators as in Example 2.36 and Example 2.38:

Lemma 2.43.

a) Let $U : \mathrm{SO}(3) \to \mathrm{U}(N)$ *be a continuous unitary representation. Then the matrices*

$$L_j^U := -i \left.\frac{d}{d\theta}\right|_{\theta=0} U(R_{e_j}(\theta)) \in M_N(\mathbb{C}), \qquad j = 1, 2, 3, \tag{2.108}$$

form a system of angular momentum operators, with

$$\sigma(L_3^U) \subset \mathbb{Z}. \tag{2.109}$$

b) Let $U : \mathrm{SU}(2) \to \mathrm{U}(N)$ be a continuous unitary representation. Then the matrices

$$L_j^U := -i \left.\frac{d}{d\theta}\right|_{\theta=0} U(e^{\frac{i}{2}\theta S(x)}) \in M_N(\mathbb{C}), \qquad j = 1,2,3, \tag{2.110}$$

form a system of angular momentum operators, with

$$\sigma(L_3^U) \subset \frac{1}{2}\mathbb{Z}. \tag{2.111}$$

Proof. We only give a brief sketch of some parts of the proof. The details are carried out in Problem 2.11.

The proof that the L_j^U defined by the representation U form angular momentum operators is similar to the one used in Examples 2.36 and 2.38 but less direct because the representing matrices are not explicitly specified. In the case of SO(3), one uses the fact that

$$TR_x(\theta)T^{-1} = R_{T(x)}(\theta), \qquad T \in \mathrm{SO}(3). \tag{2.112}$$

In the case of SU(2), the matrices $e^{\frac{i}{2}\theta(x\cdot\sigma)} = e^{\frac{i}{2}\theta S(x)}$ can be used instead of the rotation $R_x(\theta)$ with axis x and angle θ.

For the claims about the spectra of L_3, note that a rotation around the e_3-axis with angle $\theta = 2\pi$ is the identity in SO(3), so

$$1 = U(R_{e_3}(2\pi)) = e^{2\pi i L_3^U}.$$

By the spectral mapping principle, this implies $e^{2\pi i\lambda} = 1$ for all $\lambda \in \sigma(L_3^U)$, i.e. $\lambda \in \mathbb{Z}$. In SU(2), the matrices $e^{\frac{i}{2}\theta\sigma_3}$ play the role of rotations around the e_3-axis. Because of the factor $\frac{1}{2}$, we need to set $\theta = 4\pi$ to reach the identity, namely

$$1 = U(e^{2\pi i\sigma_3}) = e^{4\pi i L_3^U},$$

which only implies $\sigma(L_3^U) \subset \frac{1}{2}\mathbb{Z}$. □

We conclude that the spin l angular momentum operators from Theorem 2.40 can be of the form described in Lemma 2.43 a) (with an SO(3) representation) only if l is an integer. We also state without proof that for any system of angular momentum operators in $M_N(\mathbb{C})$, there exists a continuous unitary representation $U : \mathrm{SU}(2) \to \mathrm{U}(N)$ such that $L_j = L_j^U$ (Lemma 2.43) [Hal03, Thm. 3.7]. For an irreducible representation of integer spin $l \in \mathbb{N}_0$, this representation factors through SO(3), i.e. is given by an honest representation of SO(3).

Summary

Rotational symmetry in matrix mechanics amounts to having SO(3) act by automorphisms of the observable algebra $M_N(\mathbb{C})$, which is equivalent to studying projective representations of SO(3). Any projective representation U of SO(3) can be regarded as a projective representation $\tilde{U} := U \circ \rho$ of the simply connected covering group SU(2):

$$\begin{array}{ccc} \mathrm{SU}(2) & \xrightarrow{\rho} & \mathrm{SO}(3) \\ & \searrow^{\tilde{U}} & \downarrow U \\ & & \mathrm{PU}(N) \end{array},$$

and any projective representation of SU(2) lifts to an honest representation of SU(2) (Thm. 2.42). By passing to generators, SU(2)-representations can be described by angular momentum operators. In the case of an irreducible representation, the angular momentum operators are also irreducible and are labelled by their total angular momentum $l \in \frac{1}{2}\mathbb{N}_0$. From here, it is easy to see whether U lifts to an honest representation of SO(3): This is the case if and only if $l \in \mathbb{N}_0$ is integer. For half-integer spin $l \in \mathbb{N}_0 + \frac{1}{2}$, a rotation by (2π) is represented by $-1 \in M_N(\mathbb{C})$, and only a rotation by (4π) is represented by the identity matrix 1. This is in particular the case for the important spin $\frac{1}{2}$ representation which describes the spin of many elementary particles.

2.5 Composite systems and entanglement in matrix mechanics

We continue our exploration of similarities and differences between classical and quantum mechanics by considering composite systems and correlations, as described in Section 1.5 for the classical setting. We have already observed that the definitions of covariances and correlations directly carry over to matrix mechanics. Also regarding subsystems and composite systems, little changes, in particular when taking into account that product systems had a connection with tensor products (Lemma 1.38). The following definition is therefore directly analogous to Definition 1.36. For the time being, we will be concerned with subsystems and bipartite systems and ignore part d), which is related to Hamiltonians and dynamics.

Definition 2.44 (Subsystems and composite systems in matrix mechanics). Let $M_N(\mathbb{C})$ be a full matrix algebra.

a) A *subsystem* is a unital *-subalgebra $\mathcal{A} \subset M_N(\mathbb{C})$.
b) Given two subsystems $\mathcal{A}, \mathcal{B} \subset M_N(\mathbb{C})$, the *bipartite system generated by* $\mathcal{A}$ *and* $\mathcal{B}$ is the smallest subsystem $(\mathcal{A}, \mathcal{B})$ of $M_N(\mathbb{C})$ that contains $\mathcal{A}$ and $\mathcal{B}$.
c) The *product* (or *composite system*) of $M_{N_1}(\mathbb{C})$ and $M_{N_2}(\mathbb{C})$ is $M_{N_1}(\mathbb{C}) \otimes M_{N_2}(\mathbb{C})$.
d) The *uncoupled product* (or *interaction-free product*) of two matrix mechanical systems $(M_{N_1}(\mathbb{C}), H_1)$ and $(M_{N_2}(\mathbb{C}), H_2)$ is the matrix mechanical system

$$(M_{N_1}(\mathbb{C}) \otimes M_{N_2}(\mathbb{C}), H_1 \otimes 1_{N_2} + 1_{N_1} \otimes H_2). \tag{2.113}$$

For some reminders about algebraic tensor products, see Appendix M.2.

Example 2.45 (Subsystems of $M_N(\mathbb{C})$). There are a number of subsystems (unital *-subalgebras $\mathcal{A}$) of $M_N(\mathbb{C})$ that one sees immediately:

a) The system itself, $\mathcal{A} = M_N(\mathbb{C})$, and the trivial subsystem $\mathcal{A} = \mathbb{C}1$.
b) Let $N_1, \ldots, N_k \in \mathbb{N}$ such that $N_1 + \ldots + N_k = N$ and consider matrices of block form,

$$A = A_1 \oplus A_2 \oplus \ldots \oplus A_k = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}, \tag{2.114}$$

where $A_j \in M_{N_j}(\mathbb{C})$ is an $(N_j \times N_j)$-matrix. It is clear that matrices of the form (2.114) form a unital *-subalgebra[a], the algebra $\mathcal{A} = M_{N_1}(\mathbb{C}) \oplus \ldots \oplus M_{N_k}(\mathbb{C})$.

For $k > 1$, this is not isomorphic to a full matrix algebra because it has central elements, namely all elements of the form $c_1 1_{N_1} \oplus \ldots \oplus c_k 1_{N_k}$, with 1_{N_j} the identity in $M_{N_j}(\mathbb{C})$, and $c_j \in \mathbb{C}$.
One can show that any unital *-subalgebra $\mathcal{A} \subset M_N(\mathbb{C})$ is essentially of this form, namely there exists $\alpha \in \operatorname{Aut} M_N(\mathbb{C})$ such that $\mathcal{A} = \alpha(M_{N_1}(\mathbb{C}) \oplus \ldots \oplus M_{N_k}(\mathbb{C}))$.

c) A special subalgebra contained in the previous general example is the case $N_1 = \ldots = N_k = 1$, i.e. the algebra of diagonal $(N \times N)$-matrices. It is commutative and behaves in many ways like a classical mechanics system.

[a] It is an unsatisfactory feature of our definitions that a subsystem of a matrix mechanical system can have a more complicated structure than a matrix mechanical system. In order to avoid this, we should have defined a matrix mechanical system not based on a full matrix algebra $M_N(\mathbb{C})$, but on a general finite-dimensional unital *-algebra. This introduces some changes in the formalism because such algebras can have central elements, and has been avoided here to keep the presentation simple.

In the following, we will be mostly interested in bipartite systems consisting of two matrix algebras

$$\mathcal{A} := M_{N_1}(\mathbb{C}), \qquad \mathcal{B} := M_{N_2}(\mathbb{C}), \tag{2.115}$$

where $N_1, N_2 \in \mathbb{N}$ are arbitrary. In the interest of concise terminology, we call such a bipartite system a *full* bipartite system (tensor product of two full matrix algebras). The algebra they generate is the tensor product matrix algebra

$$\mathcal{A} \otimes B = M_{N_1}(\mathbb{C}) \otimes M_{N_2}(\mathbb{C}) \cong \operatorname{End}(\mathbb{C}^{N_1} \otimes \mathbb{C}^{N_2}) \cong M_{N_1 N_2}(\mathbb{C}). \tag{2.116}$$

To prove that, it suffices to note $M_{N_1}(\mathbb{C}) \otimes M_{N_2}(\mathbb{C}) \subset \operatorname{End}(\mathbb{C}^{N_1} \otimes \mathbb{C}^{N_2})$ (Appendix M.2), and equality holds because $M_{N_1}(\mathbb{C}) \otimes M_{N_2}(\mathbb{C})$ contains the $(N_1 N_2)^2$ many linearly independent matrix units $E_{k,l} \otimes E_{i,j}$, $k, l \in \{1, \ldots, N_1\}$, $i, j \in \{1, \ldots, N_2\}$.

The two matrix algebras $\mathcal{A}, \mathcal{B}$ can be regarded as subsystems (subalgebras) of $\mathcal{A} \otimes \mathcal{B}$ by identifying

$$\mathcal{A} \ni A \cong A \otimes 1 \in \mathcal{A} \otimes \mathcal{B}, \tag{2.117}$$

$$\mathcal{B} \ni B \cong 1 \otimes B \in \mathcal{A} \otimes \mathcal{B}, \tag{2.118}$$

where $1 = 1_{\mathcal{B}}$ denotes the identity in $\mathcal{B}$ in the first line, and $1 = 1_{\mathcal{A}}$ denotes the identity in $\mathcal{A}$ in the second line.

We begin our analysis of full bipartite systems with an elementary remark about tensor product operators.

Lemma 2.46 (Tensor product matrices). *Let $A \in \mathcal{A}$, $B \in \mathcal{B}$ be normal. Then $(A \otimes 1)(1 \otimes B) = A \otimes B = (1 \otimes B)(A \otimes 1)$. In particular, $A \otimes 1$ and $1 \otimes B$ commute and hence have a joint spectral resolution, which is given by*

$$E^{(A\otimes 1, 1\otimes B)}(\alpha, \beta) = E^A(\alpha) \otimes E^B(\beta), \qquad \alpha \in \sigma(A), \beta \in \sigma(B). \tag{2.119}$$

The joint spectrum is $\sigma(A \otimes 1, 1 \otimes B) = \sigma(A) \times \sigma(B)$.

Proof. The equation $(A\otimes 1)(1\otimes B) = A\otimes B = (1\otimes B)(A\otimes 1)$ is a consequence of general tensor product facts (Appendix M.2). For $\alpha \in \sigma(A), \beta \in \sigma(B)$, the operator $E^A(\alpha)\otimes E^B(\beta)$ is an orthogonal projection,

$$\begin{aligned}(E^A(\alpha)\otimes E^B(\beta))^2 &= E^A(\alpha)^2 \otimes E^B(\beta)^2 \\ &= E^A(\alpha)\otimes E^B(\beta) = E^A(\alpha)^* \otimes E^B(\beta)^* = (E^A(\alpha)\otimes E^B(\beta))^*.\end{aligned}$$

Furthermore,

$$\sum_{\alpha\in\sigma(A),\beta\in\sigma(B)} E^A(\alpha)\otimes E^B(\beta) = \sum_{\alpha\in\sigma(A)} E^A(\alpha)\otimes \sum_{\beta\in\sigma(B)} E^B(\beta) = 1_{\mathcal{A}}\otimes 1_{\mathcal{B}} = 1,$$

and $E^A(\alpha)\otimes E^B(\beta)$ is orthogonal to $E^A(\alpha')\otimes E^B(\beta')$ for $(\alpha,\beta) \neq (\alpha',\beta')$. Therefore $(\alpha,\beta)\mapsto E^A(\alpha)\otimes E^B(\beta)$ is a projection-valued measure on $\mathbb{R}^2$. Its support is $\sigma(A)\times\sigma(B)$, and

$$\sum_{\alpha\in\sigma(A),\beta\in\sigma(B)} \alpha\beta E^A(\alpha)\otimes E^B(\beta) = \sum_{\alpha\in\sigma(A)} \alpha E^A(\alpha)\otimes \sum_{\beta\in\sigma(B)} \beta E^B(\beta) = A\otimes B,$$

which proves the claim. □

As for vectors, we emphasize that a general element of $\mathcal{A}\otimes\mathcal{B}$ is not of the form $A\otimes B$, but only a linear combination of such simple tensors.

Note that since $\mathcal{A}\otimes\mathcal{B} \cong M_{N_1N_2}(\mathbb{C})$ is also a full matrix algebra, its states are described by density matrices in $\mathcal{A}\otimes\mathcal{B}$, i.e. positive elements of unit trace. The following definition and lemma describe how states of a full bipartite system restrict to its two factors.

Definition 2.47 (Reduced density matrices). Let $\rho \in \mathcal{D}(\mathbb{C}^{N_1}\otimes\mathbb{C}^{N_2})$ be a density matrix in $M_{N_1}\otimes M_{N_2}$. The *reduced density matrices* of ρ are

$$\rho_1 := (\mathrm{id}_{M_{N_1}}\otimes \mathrm{Tr}_2)(\rho), \qquad \rho_2 := (\mathrm{Tr}_1\otimes \mathrm{id}_{M_{N_2}})(\rho), \tag{2.120}$$

these are elements $\rho_j \in \mathcal{D}_j$, $j = 1, 2$.

To compute the reduced density matrices of ρ, write it as a linear combination of simple tensors, i.e. $\rho = \sum_{j,l} c_{jl} X_j \otimes Y_l$, and take the trace in just one of the two tensor factors[12],

$$\rho_1 = \sum_{j,l} c_{jl}\,\mathrm{Tr}_2(Y_l)X_j, \qquad \rho_2 = \sum_{j,l} c_{jl}\,\mathrm{Tr}_1(X_j)Y_l. \tag{2.121}$$

Here the index 1 or 2 on the trace functional indicates that the trace runs over $\mathbb{C}^{N_1}$ or $\mathbb{C}^{N_2}$. As this is clear from the argument of the trace (the X_j are elements of $\mathcal{A} = M_{N_1}(\mathbb{C})$, and the Y_l are elements of $\mathcal{B} = M_{N_2}(\mathbb{C})$), we will mostly drop these indices in the following.

The claim made in Definition 2.47, namely the claim that the ρ_j are density matrices, can be checked directly. We will see it a bit later indirectly (comment after Lemma 2.49).

[12] For this reason, reduced density matrices are also called *partial traces*.

A decomposition $\rho = \sum_{j,l} c_{jl} X_j \otimes Y_l$ is not unique. One can also compute the reduced density matrices via their matrix elements. Indeed, for any $v, w \in \mathbb{C}^{N_1}$ and any orthonormal basis $(b_k)_{k=1,\dots,N_2}$ of $\mathbb{C}^{N_2}$, we have

$$\begin{aligned}\langle v, \rho_1 w\rangle &= \sum_{j,l} c_{jl} \langle v, X_j w\rangle \operatorname{Tr}(Y_l) \\ &= \sum_k \sum_{j,l} c_{jl} \langle v, X_j w\rangle \langle b_k, Y_l b_k\rangle \\ &= \sum_k \sum_{j,l} c_{jl} \langle v \otimes b_k, (X_j \otimes Y_l) w \otimes b_k\rangle \\ &= \sum_k \langle v \otimes b_k, \rho(w \otimes b_k)\rangle,\end{aligned}$$

and analogously for the matrix elements of ρ_2. This method is illustrated in the following example.

Example 2.48. Let $N_1 = N_2 = 2$ (i.e. $\mathcal{A} = \mathcal{B} = M_2(\mathbb{C})$) and consider the superposition vector

$$\psi = \frac{1}{\sqrt{2}} (e_1 \otimes e_1 + e_2 \otimes e_2) \in \mathbb{C}^2 \otimes \mathbb{C}^2. \tag{2.122}$$

We have $\|\psi\| = 1$ because the two tensor product vectors have norm one and are orthogonal to each other w.r.t. the canonical scalar product on the tensor product.

We want to compute the reduced density matrices of the density matrix (projection) $\rho := P_\psi \in M_2 \otimes M_2$. By definition of ρ_1, for $v, w \in \mathbb{C}^2$, we have

$$\begin{aligned}\langle v, \rho_1 w\rangle &= \langle v, (\mathrm{id}_{M_2} \otimes \operatorname{Tr})(\rho) w\rangle \\ &= \langle v \otimes e_1, \rho(w \otimes e_1)\rangle + \langle v \otimes e_2, \rho(w \otimes e_2)\rangle \\ &= \langle v \otimes e_1, \psi\rangle\langle\psi, w \otimes e_1\rangle + \langle v \otimes e_2, \psi\rangle\langle\psi, w \otimes e_2\rangle \\ &= \frac{1}{2} (\langle v, e_1\rangle\langle e_1, w\rangle + \langle v, e_2\rangle\langle e_2, w\rangle) \\ &= \left\langle v, \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} w \right\rangle,\end{aligned}$$

from which we read off $\rho_1 = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$. Similarly one computes $\rho_2 = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$ in this example.

Lemma 2.49 (Reduced density matrices). *Let ω_ρ, $\rho \in \mathcal{D}(\mathcal{A} \otimes \mathcal{B})$, be a state on a full bipartite system $\mathcal{A} \otimes \mathcal{B}$. Then, $A \in \mathcal{A}, B \in \mathcal{B}$,*

$$\omega_\rho(A \otimes 1) = \operatorname{Tr}_1(\rho_1 A), \qquad \omega_\rho(1 \otimes B) = \operatorname{Tr}_2(\rho_2 B), \tag{2.123}$$

with ρ_1, ρ_2 the reduced density matrices of ρ.

Proof. We give the proof for the first subsystem. Write $\rho = \sum_{jl} c_{jl} X_j \otimes Y_l$ as a linear combination of simple tensors, and let $A \in \mathcal{A}$ be arbitrary. Then

$$\begin{aligned}\omega_\rho(A \otimes 1) &= \sum_{j,l} c_{jl} \operatorname{Tr}\big(AX_j \otimes Y_l\big) \\ &= \sum_{j,l} c_{jl} \operatorname{Tr}\big(AX_j\big) \operatorname{Tr}(Y_l) \\ &= \operatorname{Tr}\left(A \sum_{j,l} c_{jl} \operatorname{Tr}(Y_l) X_j\right) = \operatorname{Tr}(\rho_1 A).\end{aligned}$$

□

As $A \mapsto \omega_\rho(A \otimes 1)$ is clearly a normalized positive linear functional (state) on $\mathcal{A} = M_{N_1}(\mathbb{C})$, and states are in bijection with density matrices, it follows that ρ_1 is indeed a density matrix, as claimed in Def. 2.47. The same holds for ρ_2.

The role of ρ_1 and ρ_2 is to describe the restriction of the state ω_ρ of the bipartite system to its two defining subsystems. Note that, as shown in Example 2.48, a *pure* state of the bipartite system may restrict to *mixed* states of its two defining subsystems, in contrast to the situation in classical mechanics.

We now consider correlations of states of bipartite systems. As in classical mechanics, we call a state ω of a bipartite system $\mathcal{A} \otimes \mathcal{B}$ *uncorrelated*, or *product state*, if

$$\omega(A \otimes B) = \omega(A \otimes 1)\omega(1 \otimes B), \qquad A \in \mathcal{A}, B \in \mathcal{B}, \tag{2.124}$$

and *correlated* otherwise. The state ω being a product state is equivalent to all covariances $\operatorname{Cov}_\omega(A, B) = 0$ vanishing when A is taken from the first and B from the second subsystem.

Product states can be characterized easily.

Proposition 2.50 (Product states on matrix bipartite systems). *A state $\omega = \omega_\rho$ of the full bipartite system $\mathcal{A} \otimes \mathcal{B}$ is a product state if and only if its density matrix ρ satisfies*

$$\rho = \rho_1 \otimes \rho_2, \tag{2.125}$$

where $\rho_j \in \mathcal{D}_{N_j}$ are the reduced density matrices of ρ. Given any $\rho_j \in \mathcal{D}_{N_j}$, the density matrix $\rho := \rho_1 \otimes \rho_2$ defines a product state on $\mathcal{A} \otimes \mathcal{B}$.

Proof. Consider two arbitrary density matrices $\rho_j \in \mathcal{D}_{N_j}$ and their tensor product $\rho := \rho_1 \otimes \rho_2$. Then

$$\rho = \rho_1^{1/2}\rho_1^{1/2} \otimes \rho_2^{1/2}\rho_2^{1/2} = (\rho_1^{1/2} \otimes \rho_2^{1/2})^*(\rho_1^{1/2} \otimes \rho_2^{1/2}) \geq 0,$$

so ρ is positive. Furthermore (cf. Appendix M.2)

$$\operatorname{Tr}\rho = \operatorname{Tr}\rho_1 \cdot \operatorname{Tr}\rho_2 = 1 \cdot 1 = 1, \tag{2.126}$$

i.e. ρ has unit trace. Hence ρ is a density matrix and defines a state of the bipartite system. Given any $A \in \mathcal{A}, B \in \mathcal{B}$,

$$\begin{aligned}\omega_\rho(A \otimes B) &= \operatorname{Tr}(\rho(A \otimes B)) \\ &= \operatorname{Tr}(\rho_1 A \otimes \rho_2 B) = \operatorname{Tr}(\rho_1 A) \operatorname{Tr}(\rho_2 B).\end{aligned}$$

As

$$\mathrm{Tr}(\rho_1 A) = \mathrm{Tr}(\rho(A \otimes 1)) = \omega_\rho(A \otimes 1),$$

and similarly $\mathrm{Tr}(\rho_2 B) = \omega_\rho(1 \otimes B)$, the claimed equation

$$\omega_\rho(A \otimes B) = \omega_\rho(A \otimes 1)\omega_\rho(1 \otimes B)$$

follows. This shows that the tensor product of two density matrices defines a product state.

Consider now conversely a product state ω_ρ on $\mathcal{A} \otimes \mathcal{B}$. Then, $A \in \mathcal{A}, B \in \mathcal{B}$,

$$\begin{aligned}\omega_\rho(A \otimes B) &= \omega_\rho(A \otimes 1)\omega_\rho(1 \otimes B) = \mathrm{Tr}(\rho_1 A)\,\mathrm{Tr}(\rho_2 B) \\ &= \mathrm{Tr}((\rho_1 \otimes \rho_2)(A \otimes B)) = \omega_{\rho_1 \otimes \rho_2}(A \otimes B),\end{aligned}$$

where we have used Lemma 2.49. Since every element of $M_{N_1} \otimes M_{N_2}$ is a linear combination of simple tensors $A \otimes B$, it follows that $\omega_\rho = \omega_{\rho_1 \otimes \rho_2}$, which is equivalent to $\rho = \rho_1 \otimes \rho_2$. □

Example 2.51. The pure state considered in Example 2.48 is not a product state: Its reduced density matrices are $\rho_1 = \rho_2 = \frac{1}{2}\mathbb{1}$, but $\frac{1}{2}\mathbb{1} \otimes \frac{1}{2}\mathbb{1} = \frac{1}{4}\mathbb{1}$ is different from P_ψ (2.122). For example, P_ψ is a rank one projection, but $\frac{1}{4}\mathbb{1}$ is not (has eigenvalue $\frac{1}{4}$).

Note, however, that a density matrix that does not look like a simple tensor might still be one. For example, given density matrices ρ_1, ρ_2,

$$\frac{1}{16}\rho_1 \otimes \rho_1 + \frac{3}{16}\rho_1 \otimes \rho_2 + \frac{3}{16}\rho_2 \otimes \rho_1 + \frac{9}{16}\rho_2 \otimes \rho_2$$

might not seem to be of tensor product form, but coincides with the tensor square $(\frac{1}{4}\rho_1 + \frac{3}{4}\rho_2) \otimes (\frac{1}{4}\rho_1 + \frac{3}{4}\rho_2)$ and therefore defines a product state.

The following definitions are the same as in classical mechanics but a bit simpler because due to the finite dimensionality of the observable algebras considered here, we do not need to consider limits of convex combinations.

Definition 2.52 (Classically correlated and entangled states). A correlated state ω on a bipartite system $\mathcal{A} \otimes \mathcal{B}$ is called *classically correlated* or *separable* if it is a convex combination of product states, and *entangled* otherwise.

In the case of pure states, entanglement can be characterized directly: Pure states on full bipartite systems are either product states or entangled (in contrast to mixed states).

Lemma 2.53 (Entanglement of pure states on full bipartite systems). *Let ω be a pure state on a full bipartite system $\mathcal{A} \otimes \mathcal{B}$, given by a unit vector $\psi \in \mathbb{C}^{N_1} \otimes \mathbb{C}^{N_2}$ via $\omega = \omega_{P_\psi}$. Then ω is a product state if and only if ψ is a simple tensor, i.e. $\psi = \psi_1 \otimes \psi_2$ for unit vectors $\psi_j \in \mathbb{C}^{N_j}$. If ω is not a product state then it is entangled.*

Proof. We know already that ω is a product state if and only if its density matrix P_ψ is a simple tensor, $P_\psi = \rho_1 \otimes \rho_2$ for density matrices ρ_1, ρ_2 on $\mathcal{A}$ and $\mathcal{B}$, respectively. This is equivalent to

$$1 = \mathrm{Tr}\big(P_\psi^2\big) = \mathrm{Tr}\big(\rho_1^2 \otimes \rho_2^2\big) = \mathrm{Tr}\big(\rho_1^2\big)\,\mathrm{Tr}\big(\rho_2^2\big).$$

As $\mathrm{Tr}\left(\rho_j^2\right) \leq 1$, this implies $\mathrm{Tr}(\rho_1^2) = \mathrm{Tr}\left(\rho_2^2\right) = 1$, i.e. ρ_1 and ρ_2 are rank one projections. Hence pure product states are necessarily tensor products of pure states.

Now suppose that $\omega = \omega_{P_\psi}$ is a pure separable or uncorrelated state, i.e. a convex combination of product states, namely

$$P_\psi = \sum_j \lambda_j \rho_{1,j} \otimes \rho_{2,j}$$

for coefficients $\lambda_j \in (0,1)$ with $\sum_j \lambda_j = 1$, and density matrices $\rho_{1,j}, \rho_{2,j}$. As any $\rho_{1,j} \otimes \rho_{2,j}$ defines a state, purity implies that the convex combination must be trivial, namely $P_\psi = \rho_1 \otimes \rho_2$ defines a product state. □

The observation made in this lemma that pure states on full bipartite systems are either product states or entangled is not true for mixed states. For pure states, this lemma however clarifies the structure of entangled states and in particular, we see that in matrix mechanics, many entangled states exist. For instance, the pure state from Example 2.48 is entangled: In Example 2.51 it was shown that it is not a product state, hence it is entangled by Lemma 2.53.

Another way of verifying that a given state is entangled is by using the entanglement entropy defined below. This quantitiy is based on the *information function*

$$f : [0,\infty) \to \mathbb{R}, \qquad f(\lambda) := -\lambda \log \lambda, \tag{2.127}$$

where $f(0) := 0$ is understood. By functional calculus, we may define $f(\rho)$ for a density matrix ρ, and the *von Neumann entropy*

$$S(\rho) := \mathrm{Tr}(f(\rho)) = -\mathrm{Tr}(\rho \log \rho). \tag{2.128}$$

Definition 2.54 (Entanglement entropy). The *entanglement entropy* of a pure state of a full bipartite system is

$$E : S_{\mathrm{pure}}(\mathcal{A}) \to \mathbb{R}, \quad E(P) := S(P_1) = -\mathrm{Tr}(P_1 \log P_1), \tag{2.129}$$

where P_1 is the first reduced density matrix of the rank one projection P.

Equivalently, $E(P) = S(P_2) = S(P_1)$ can be defined via the second reduced density matrix of P. In Problem 2.12, you are asked to prove this, together with some other important properties of the entanglement entropy, which are summarized here.

Proposition 2.55 (Properties of entanglement entropy). *The entanglement entropy of the full bipartite system $M_{N_1}(\mathbb{C}) \otimes M_{N_2}(\mathbb{C})$ has the following properties: For any rank one projection P, one has*

a) $0 \leq E(P) \leq \log \min\{N_1, N_2\}$,
b) $E(P) = 0$ *if and only if P defines a product state,*
c) $E((U_1 \otimes U_2)P(U_1 \otimes U_2)^*) = E(P)$ (*invariance under local operations*)

The idea of the entanglement entropy is that it not only gives a definite way of testing whether a given pure state is entangled or not (compute $E(P)$ and check whether it is zero), but also a measure of "how much" entangled it is. States for which the entanglement entropy matches the upper bound $\log \min\{N_1, N_2\}$ exist, they are called *maximally entangled states*.

The state P_ψ in Example 2.48 has reduced density matrix $(P_\psi)_1 = \frac{1}{2}\mathbb{1}$, and hence entanglement entropy

$$E(P_\psi) = -\operatorname{Tr}\left(\frac{1}{2}\mathbb{1}\log\left(\frac{1}{2}\mathbb{1}\right)\right) = -\log\left(\frac{1}{2}\right) = \log 2. \tag{2.130}$$

This shows that not only P_ψ is entangled, but even maximally entangled.

For mixed states, the entanglement entropy defined here is not a good entanglement measure, and more complicated methods have to be used [Hor+09].

⚛ **(Experiments, data analysis, and joint probabilities.)** At this point, entanglement is only surprising insofar that it is a kind of correlation that we did not see in classical mechanics. To understand better what it means, we have to discuss hidden variable models.

For the time being, we do not think about theory, but only experiments. We consider a bipartite system formed of two parties (experimental physicists), usually called "Alice" and "Bob" that work in separate locations/laboratories. We think of a state of this system, which experimentally speaking means a preparation of the system in a certain way. For example, we might consider a machine in Erlangen which produces pairs of black and white marbles (always one white and one black marble), packs them into boxes and ships one box to Alice (in her subsystem in Alicante) and one to Bob (in his subsystem in Boston). Then Alice and Bob can make measurements on their respective subsystems, for instance they may look into the box they have received and observe the color of the marble in it. Unsurprisingly, when Alice finds a white marble she knows that Bob will have received a black one, i.e. there will be correlations. Another example would be a machine producing a pair of spin $\frac{1}{2}$ particles from a spin 0 state – in this case, a spin conservation law would demand that when the measurement of Alice shows spin up then Bob must find spin down (if the system is not disturbed).

In general, we have a state ω describing the preparation of the system, and the physicists Alice and Bob may choose which observables $A_1, A_2, \ldots$ (Alice) and $B_1, B_2, \ldots$ (Bob) they measure. To perform statistics on the measurement data, the experiment will be repeated often and Alice/Bob take note of which observables they measured in run $n \in \{1, \ldots, N\}$, and what the outcome was in each case (e.g. Alice writes down "in run 1 I measured the observable A and the outcome was $\alpha \in \mathbb{R}$").

After the experiment is finished, Alice and Bob communicate, compare their experimental data, and compute statistical averages. The interesting quantities to consider are the joint probabilities $\Pr_\omega((A, B) = (\alpha, \beta)) \in [0, 1]$, i.e. the probability that in the state ω, Alice's observable A has the value α and Bob's observable B has the value β. In an experiment with N iterations, this is the fraction (number of runs with $A = \alpha$ and $B = \beta)/N$.

Alice and Bob have now finished their experimental work and begin to think about theoretical explanations for the results they found. For example, they could wonder whether classical mechanics is in agreement with the experiment. To that end, they need to identify the correct mathematical description of their setup, i.e. a phase space M, a compactly supported probability measure μ on M (to describe the state $\omega = \omega_\mu$), and for each observable A, B, a smooth function on phase space (they acquire this information by asking a theoretical/mathematical physicist with experience in classical

mechanics). Then (see Def. 1.28 and the discussion below it)

$$\begin{aligned}\Pr_{\omega_\mu}((A,B)=(\alpha,\beta)) &= \mu(\{\xi\in\mathbb{R}^{2d} \;:\; A(\xi)=\alpha,\, B(\xi)=\beta\}) \\ &= \int_M \chi_{\{\alpha\}}(A(\xi))\chi_{\{\beta\}}(B(\xi))\,d\mu(\xi)\end{aligned}$$

is the theoretical prediction of classical mechanics for the joint probability that A has value α and B has value β.

In case the predictions of classical mechanics do not match with experiment, the task is to find a better theory. Instead of developing a full-fledged theory, one could also make only some basic assumptions about it – as we shall see, already this can lead to interesting statements.

Definition 2.56 (Local hidden variable models). Joint probabilities of a state ω as discussed above admit a *local hidden variable model* if there exists a measure space X with probability measure ν, and for every pair of observables A (for Alice) and B (for Bob) integrable functions $f_A : \sigma(A)\times X\to[0,1]$, $f_B : \sigma(B)\times X\to[0,1]$ such that

$$\Pr_\omega((A,B)=(\alpha,\beta)) = \int_X f_{A,\omega}(\alpha,x)\,f_{B,\omega}(\beta,x)\,d\nu(x). \tag{2.131}$$

Here $\sigma(A),\sigma(B)\subset\mathbb{R}$ denote the sets of possible values of A and B, respectively.

It should be noted that this is a very general setting. No detailed assumptions about the underlying theory, i.e. the way in which the joint probabilities $\Pr_\omega((A,B)=(\alpha,\beta))$ are computed, are made. The space X and the measure ν is not fixed. The idea is that x is a "hidden variable", i.e. a quantity that we have not yet discovered or cannot observe, and the observed probabilistic features arise as a consequence of our ignorance of $x\in X$, hence the average (integration against ν) over the hidden variable. We interpret $f_{A,\omega}(\alpha,x)$ as the probability of outcome α for the observable A in the state ω given the hidden variable x, and likewise for B and β.

In particular, we could consider the more specific deterministic scenario in which the functions $f_{A,\omega}$ and $f_{B,\omega}$ take only values in $\{0,1\}$. In that case, they would only function to select for which hidden variables x the observable A has value α and for which not. This is basically only saying that we assume that all observables have *some* values $\alpha,\beta,..$ (this assumption is called *realism*), and due to our lack of complete knowledge of the system, we see probabilistic features.

While this seems tautological to the classically trained mind, there is one assumption that *does* enter the local hidden variable model, namely that the probabilities $f_{A,\omega}(\alpha,x)$ of A do not explicitly depend on B or β. This is justified in a setup in which Alice does not know which observables Bob is measuring, and the measurements on the two subsystems do not influence each other in any way. If the laboratories of Alice and Bob are sufficiently far apart, and the measurements take place in a short period of time, this will always be the case in a theory complying with a maximal speed of communication (speed of light). Hence models of the form (2.131) are called *local* hidden variable models.

Example 2.57 (Separable states admit local hidden variable models). Of course, in classical mechanics, any two observables and any state admit local hidden variable models. Here the phase space point $x = \xi$ can be regarded as the hidden variable.

In matrix mechanics, let us consider a full bipartite system $\mathcal{A} \otimes \mathcal{B}$ and a separable state $\rho = \sum_{j=1}^{n} \lambda_j \rho_{1,j} \otimes \rho_{2,j}$ (convex combination of product states). Then we may consider $X := \{1, \dots, n\}$ with the probability measure $\nu(x) := \lambda_x \in [0,1]$, as well as $f_{A,\rho}(\alpha, x) := \mathrm{Tr}\big(\rho_{1,x} E^A(\alpha)\big)$, $f_{B,\rho}(\beta, x) := \mathrm{Tr}\big(\rho_{2,x} E^B(\beta)\big)$. This gives

$$\begin{aligned}\int_X f_{A,\rho}(\alpha, x) f_{B,\rho}(\beta, x)\, d\nu(x) &= \sum_{x=1}^{n} \lambda_x \,\mathrm{Tr}\big(\rho_{1,x} E^A(\alpha)\big)\,\mathrm{Tr}\big(\rho_{2,x} E^B(\beta)\big) \\ &= \mathrm{Tr}\left(\sum_{x=1}^{n} \lambda_x (\rho_{1,x} \otimes \rho_{2,x})(E^A(\alpha) \otimes E^B(\beta))\right) \\ &= \mathrm{Tr}\big(\rho E^{A\otimes 1, 1\otimes B}(\alpha, \beta)\big),\end{aligned}$$

which is precisely the expectation value of the joint spectral projection of the two commuting observables $A \otimes 1, 1 \otimes B$ corresponding to the joint spectral value $(\alpha, \beta) \in \sigma(A \otimes 1, 1 \otimes B)$, and hence coincides with the matrix mechanical prediction for the joint probability $\mathrm{Pr}_\rho((A, B) = (\alpha, \beta))$ (Proposition 2.17).

We conclude that separable states on full bipartite systems admit local hidden variable models, in agreement with the idea that the classical correlations in such states are reflections of our partial ignorance of the preparation of the state.

We refer to [WW01] for a more detailed review of hidden variables and entanglement.

So far we have seen that local hidden variable models are very natural from a classical point of view, and always realized in classical physics as well as in separable states of matrix mechanics. For entangled states, we are not sure at this point of our discussion. In general, it is also not clear how to decide in a given theory whether local hidden variable models exist. In this context, it is important to learn that the existence of local hidden variable models leads to certain inequalities – so-called Bell Inequalities and their generalizations – that can be checked in theory and measured in experiment.

We restrict ourselves to a simple version in which the considered observables can only take the two values ± 1. Such observables are called *binary*. The following definition again takes place on the level of the experimental/statistical data collected by Alice and Bob and will be compared with the matrix mechanical setting afterwards.

Given the joint probabilities $\mathrm{Pr}_\omega((A, B) = (\alpha, \beta))$ (with $\alpha, \beta \in \{\pm 1\}$), we introduce the averaged quantities

$$\mathbb{E}_\omega(A, B) := \sum_{\alpha,\beta=\pm 1} \alpha\beta \,\mathrm{Pr}_\omega((A, B) = (\alpha, \beta)). \tag{2.132}$$

In classical mechanics, we have[13]

$$\begin{aligned}\mathbb{E}_\omega(A,B) &= \sum_{\alpha,\beta=\pm1} \alpha\beta \int_M \chi_{\{\alpha\}}(A(\xi))\chi_{\{\beta\}}(B(\xi))\,d\mu(\xi)\\ &= \int_M \big(\chi_{\{1\}}(A(\xi)) - \chi_{\{-1\}}(A(\xi))\big)\big(\chi_{\{1\}}(B(\xi)) - \chi_{\{-1\}}(B(\xi))\big)\,d\mu(\xi)\\ &= \int_M A(\xi)B(\xi)\,d\mu(\xi)\\ &= \omega_\mu(AB),\end{aligned}$$

the expectation value of the product of both observables. In a full bipartite system in matrix mechanics with two observables $A\otimes 1, 1\otimes B$, we have

$$\begin{aligned}\mathbb{E}_\omega(A\otimes 1, 1\otimes B) &= \sum_{\alpha,\beta=\pm1} \alpha\beta\,\omega(E^{A\otimes 1,1\otimes B}((\alpha,\beta)))\\ &= \sum_{\alpha,\beta=\pm1} \alpha\beta\,\omega(E^A(\alpha)\otimes E^B(\beta))\\ &= \omega(A\otimes B),\end{aligned}$$

the expectation value of the product of both observables. In general, $\mathbb{E}_\omega(A,B)$ plays the role of expectation value of the product observable AB also when we have no observable algebra at our disposal.

Definition 2.58. Consider a bipartite system with binary observables A_1, A_2 for Alice and B_1, B_2 for Bob. In a given state ω, the *Bell correlation* (in CHSH form[a]) is

$$\mathbb{b}_\omega(A_1,A_2,B_1,B_2) := \mathbb{E}_\omega(A_1,B_1) + \mathbb{E}_\omega(A_1,B_2) + \mathbb{E}_\omega(A_2,B_1) - \mathbb{E}_\omega(A_2,B_2).$$

[a] CHSH stands for the names Clauser-Horne-Shimony-Holt.

Theorem 2.59 (Local hidden variable models imply Bell Inequalities). *Let A_1, A_2, B_1, B_2 be binary observables of a bipartite system, and let ω be a state on this system. If ω admits a local hidden variable model, then the* Bell Inequality *holds, namely*

$$|\,\mathbb{b}_\omega(A_1,A_2,B_1,B_2)| \le 2. \tag{2.133}$$

Proof. Using the assumption of a local hidden variable model, there exists a measure space X with probability measure ν and integrable functions $f_{A_j,\omega} : \{\pm1\}\times X \to [0,1], f_{B_j,\omega} : \{\pm1\}\times X \to [0,1]$, $j = 1,2$. We write $f^{\pm,x}_{A_j,\omega} := f_{A_j,\omega}(\pm1,x)$, $f^{\pm,x}_{B_k,\omega} := f_{B_k,\omega}(\pm1,x)$ as a shorthand, and note that for fixed hidden variable $x\in X$, the expectation value of A_j is

$$\mathbb{E}^x_\omega(A_j) := \sum_{\alpha=\pm1} \alpha f^{\alpha,x}_{A_j,\omega} = f^{+,x}_{A_j,\omega} - f^{-,x}_{A_j,\omega} \in [-1,1],$$

[13] Note that the idealization to observables taking only the two values ±1 is extremely restrictive for smooth functions $A, B : M \to \mathbb{R}$. But one can extend this part of classical mechanics easily to functions in $L^\infty(M,\mu)$, where μ is the measure describing the state under consideration.

and similarly $\mathbb{E}^x_\omega(B_j) \in [-1,1]$. Now the hidden variable model gives

$$\mathbb{E}_\omega(A_j, B_k) = \int_X \sum_{\alpha,\beta=\pm 1} \alpha\beta f^{\alpha,x}_{A_j,\omega} f^{\beta,x}_{B_k,\omega}\, d\nu(x) = \int_X \mathbb{E}^x_\omega(A_j)\mathbb{E}^x_\omega(B_k)\, d\nu(x),$$

and the Bell correlation can be estimated by triangle inequality according to

$$|\,\mathbb{b}_\omega(A_1, A_2, B_1, B_2)| \le \int_X \Big|\mathbb{E}^x_\omega(A_1)(\mathbb{E}^x_\omega(B_1) + \mathbb{E}^x_\omega(B_2)) + \mathbb{E}^x_\omega(A_2)(\mathbb{E}^x_\omega(B_1) - \mathbb{E}^x_\omega(B_2))\Big| d\nu(x).$$

Since ν is a probability measure, the claimed Bell Inequality (2.133) follows once we have shown that the integrand is bounded by 2 for every $x \in X$. But this is elementary: Writing $a_j := E^x_\omega(A_j)$, $b_k := E^x_\omega(B_k) \in [-1,1]$, we claim $|a_1(b_1+b_2) + a_2(b_1-b_2)| \le 2$. Indeed,

$$\begin{aligned} |a_2(b_1-b_2)| &= |a_2b_1(1 \pm a_1b_2) - a_2b_2(1 \pm a_1b_1)| \\ &\le |1 \pm a_1b_2| + |1 \pm a_1b_1| \\ &= 2 \pm a_1(b_1+b_2) \\ \implies |a_1(b_1+b_2) + a_2(b_1-b_2)| &\le |a_1(b_1+b_2)| + |a_2(b_1-b_2)| \\ &\le 2 \pm a_1(b_1+b_2) + |a_1(b_1+b_2)| \\ &= 2. \end{aligned}$$

In the last line, we have chosen the sign $\pm$ such that $\pm a_1(b_1+b_2) + |a_1(b_1+b_2)| = 0$. □

From our earlier discussion it is clear that the Bell Inequality holds in classical mechanics, and for all separable states in matrix mechanics. We next show that it does *not* hold for all states in matrix mechanics, and begin with a lemma about a particular state vector.

Lemma 2.60 (The singlet state). *Consider the vector*[a]

$$\psi = \frac{1}{\sqrt{2}}(e_2 \otimes e_1 - e_1 \otimes e_2) \in \mathbb{C}^2 \otimes \mathbb{C}^2.$$

Then the projection onto $\mathbb{C}\psi$ can be expressed in terms of the Pauli matrices as

$$P_\psi = \frac{1}{4}\left(1 \otimes 1 - \sum_{j=1}^3 \sigma_j \otimes \sigma_j\right). \tag{2.134}$$

[a] This vector has various interesting properties: for example, ψ is invariant under $\mathrm{SU}(2)$ in the sense $(U \otimes U)\psi = \psi$ for all $U \in \mathrm{SU}(2)$.

Proof. We define $F := \frac{1}{2}\left(1 \otimes 1 + \sum_{j=1}^3 \sigma_j \otimes \sigma_j\right)$ and claim that F is the *tensor flip*, i.e. $F(v \otimes w) = w \otimes v$ for all $v, w \in \mathbb{C}^2$. This claim can be checked in the tensor product basis

$(e_j \otimes e_k)_{j,k=1,2}$:

$$
\begin{aligned}
F(e_1 \otimes e_1) &= \frac{1}{2}(e_1 \otimes e_1 + e_2 \otimes e_2 - e_2 \otimes e_2 + e_1 \otimes e_1) = e_1 \otimes e_1, \\
F(e_1 \otimes e_2) &= \frac{1}{2}(e_1 \otimes e_2 + e_2 \otimes e_1 + e_2 \otimes e_1 - e_1 \otimes e_2) = e_2 \otimes e_1, \\
F(e_2 \otimes e_1) &= \frac{1}{2}(e_2 \otimes e_1 + e_1 \otimes e_2 + e_1 \otimes e_2 - e_2 \otimes e_1) = e_1 \otimes e_2, \\
F(e_2 \otimes e_2) &= \frac{1}{2}(e_2 \otimes e_2 + e_1 \otimes e_1 - e_1 \otimes e_1 + e_2 \otimes e_2) = e_2 \otimes e_2.
\end{aligned}
$$

The claimed formula (2.134) is equivalent to $P_\psi = \frac{1}{2}(1 - F)$. As $F\psi = -\psi$, we have $\frac{1}{2}(1 - F)\psi = \psi = P_\psi\psi$. The three vectors $e_1 \otimes e_1, e_2 \otimes e_2, \frac{1}{\sqrt{2}}(e_1 \otimes e_2 + e_2 \otimes e_1)$ complete ψ to an orthonormal basis of $\mathbb{C}^2 \otimes \mathbb{C}^2$ and are invariant under F (symmetric). Hence $\frac{1}{2}(1 - F)$ maps them to zero, in agreement with the orthogonal projection P_ψ. □

Example 2.61 (Violation of Bell Inequalities in matrix mechanics). On the full bipartite system $M_2(\mathbb{C}) \otimes M_2(\mathbb{C})$, we consider the pure entangled state ω_ψ, with the vector

$$
\psi = \frac{1}{\sqrt{2}}(e_2 \otimes e_1 - e_1 \otimes e_2)
$$

from Lemma 2.60. The interpretation of the state ψ is that it describes a pair of coupled spin $\frac{1}{2}$ particles.

To test the Bell Inequality, we have to choose binary observables for Alice and Bob, i.e. $A \in M_2(\mathbb{C})$ selfadjoint with eigenvalues ± 1. In other words, $A = A^*$ must have $\mathrm{Tr}(A) = 0$ and $\det(A) = -1$. Recall from Example 2.38 that such matrices are always of the form $A = x\sigma = x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3$ with a unit vector $x \in \mathbb{R}^3$. This observable corresponds to a measurement of the spin along the axis $x \in \mathbb{R}^3$.

Using Lemma 2.60 for expressing P_ψ in terms of Pauli matrices, recalling that all Pauli matrices have trace zero, and that $\mathrm{Tr}(\sigma_j\sigma_k) = 2\delta_{jk}$ (this was shown in Problem 2.10), expectation values of tensor products of these matrices are

$$
\begin{aligned}
\omega_\psi((x\sigma) \otimes (y\sigma)) &= \mathrm{Tr}\big(P_\psi((x\sigma) \otimes (y\sigma))\big) \\
&= \frac{1}{4}\,\mathrm{Tr}\left((x\sigma) \otimes (y\sigma) - \sum_{j=1}^{3} \sigma_j(x\sigma) \otimes \sigma_j(y\sigma)\right) \\
&= -\frac{1}{4}\sum_{j,k,l=1}^{3} x_k y_l\,\mathrm{Tr}\big(\sigma_j\sigma_k\big)\,\mathrm{Tr}\big(\sigma_j\sigma_l\big) \\
&= -\langle x, y\rangle.
\end{aligned}
$$

Similarly, one finds $\omega_\psi(1 \otimes (x\sigma)) = \omega_\psi((x\sigma) \otimes 1) = 0$. In particular, spin measurements in the same direction x are perfectly anticorrelated for Alice and Bob,

$$
\mathrm{Cov}_{\omega_\psi}((x\sigma) \otimes 1, 1 \otimes (x\sigma)) = \mathrm{Corr}_{\omega_\psi}((x\sigma) \otimes 1, 1 \otimes (x\sigma)) = -1
$$

(the uncertainties are 1). This does however not rule out interesting correlations when Alice and Bob choose different directions for their spin observables. Consider $A_1 = (x\sigma) \otimes 1$, $A_2 = (x'\sigma) \otimes 1$, $B_1 = 1 \otimes (y\sigma)$, $B_2 = 1 \otimes (y'\sigma)$, with unit vectors $x, x', y, y' \in \mathbb{R}^3$. Then the Bell correlation is

$$|\,\mathbb{b}_{\omega_\psi}(A_1, A_2, B_1, B_2)| = |\langle x, y\rangle + \langle x, y'\rangle + \langle x', y\rangle - \langle x', y'\rangle|\,.$$

To make this correlation maximal, let $x = e_1, x' = e_2$ be the canonical basis of a two-dimensional subspace $\mathbb{R}^2 \subset \mathbb{R}^3$, and $y = \frac{1}{\sqrt{2}}(e_1 + e_2)$, $y' = \frac{1}{\sqrt{2}}(e_1 - e_2)$ an orthonormal basis of the same subspace that is rotated by $\pi/4$ w.r.t. the canonical one. Then

$$\mathbb{b}_{\omega_\psi}(A_1, A_2, B_1, B_2) = \left|\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} - \left(-\frac{1}{\sqrt{2}}\right)\right| = 2\sqrt{2} > 2.$$

This example shows that in matrix mechanics, there exist situations in which Bell's Inequality is violated. We also remark that this is the maximal possible violation of Bell's Inequality:

Lemma 2.62 (Tsirelson's bound). *Let $A_1, A_2, B_1, B_2 \in M_N(\mathbb{C})$ be selfadjoint with $A_1^2 = A_2^2 = B_1^2 = B_2^2 = 1$ and $[A_j, B_k] = 0$, $j, k = 1, 2$. Then, for any state $\omega \in S(M_N(\mathbb{C}))$,*

$$\mathbb{b}_\omega(A_1, A_2, B_1, B_2) = |\omega(A_1(B_1 + B_2) + A_2(B_1 - B_2))| \leq 2\sqrt{2}.$$

Proof. Let $C := A_1(B_1 + B_2) + A_2(B_1 - B_2) = C^*$. Then, by explicit calculation, $C^2 = 4 - [A_1, A_2][B_1, B_2]$, and therefore $C^2 \leq 4 + \|[A_1, A_2]\|\|[B_1, B_2]\| \leq 8$, hence $C \leq \sqrt{8} = 2\sqrt{2}$. Evaluating in ω proves the claim. □

From our discussion of the entangled singlet state, it is clear that the four binary observables A_1, A_2, B_1, B_2 have to be chosen carefully in order to violate the Bell Inequality, i.e. for many choices we will have $|\,\mathbb{b}_\omega(A_1, A_2, B_1, B_2)| \leq 2$. Furthermore, there exist mixed entangled states that satisfy Bell Inequality for all choices of binary A_1, A_2, B_1, B_2. Examples are the so-called Werner states, see Problem 2.14.

⚛ **Experimental tests of Bell's Inequality.** We have seen that Bell Inequalities are satisfied in theories admitting local hidden variable descriptions, but violated in matrix mechanics. As Bell correlations can be directly computed from measurement data from experiments, one can run an experimental test of the idea of hidden variable descriptions. The first experimental test of Bell's Inequality was carried out by the team of Aspect in 1981 (Nobel Prize of Physics 2022) in an optical experiment with entangled polarized photons (these are not spin $\frac{1}{2}$ particles, but their polarization behaves in the same way). As in Example 2.61, the observables in the two subsystems measure polarizations along different axes that are set up in such a way that the Bell correlation becomes maximal.

The experimental value of the Bell correlation was $\mathbb{b}_\omega \approx 2.73$, close to the matrix mechanics prediction $2\sqrt{2} \approx 2.82$ and certainly larger than 2. Many later experiments confirmed this. So the experimental findings rule out local hidden variable

models and give strong support for quantum mechanics.

When accepting the local nature of physics (no faster than light communication), this leads us to conclude that the observables of matrix mechanics do not have fixed values before the measurement process, and can not be explained by hidden variable theories. Only after measuring them, they acquire definite values, as described by a "collapse of the wavefunction" or more carefully by measurement theory as indicated in Exercise 2.3.

Although this state of affairs is a radical break from classical physics and in conflict with what we would intuitively expect based on our everyday (classical) experience, it is the perspective that is widely accepted today.

Nowadays, entanglement is known not only as a fundamental property of quantum physics, but also as a computational resource of central importance in quantum information. Time permitting, we will come back to this point later in this course.

2.6 Exercises and Problems

Exercise 2.1 (The commutator on $M_N(\mathbb{C})$). Verify the properties of the commutator claimed in Lemma 2.4.

Exercise 2.2 (Numerical range and uncertainty). Let $A = A^* \in M_N(\mathbb{C})$. Denote by $\lambda^A_{\max/\min} \in \mathbb{R}$ the maximal/minimal eigenvalue of A, by $\operatorname{diam}\sigma(A) := \frac{1}{2}(\lambda^A_{\max} - \lambda^A_{\min})$ the diameter of the spectrum of A, and by $\operatorname{Conv}\sigma(A) := [\lambda^A_{\min}, \lambda^A_{\max}]$ the convex hull of the spectrum of A. Show the following:

a) The possible expectation values of A in arbitrary states are

$$\{\omega_\rho(A) \,:\, \rho \in \mathscr{D}_N\} = \{\omega_\psi(A) \,:\, \psi \in \mathbb{C}^N,\ \|\psi\| = 1\} = \operatorname{Conv}\sigma(A). \tag{2.135}$$

b) The uncertainties satisfy

$$\sup_{\rho\in\mathscr{D}_A} \Delta_\rho(A) = \operatorname{diam}(\sigma(A)) \leq \|A\|. \tag{2.136}$$

Exercise 2.3 (State update rule). Let $A \in M_N(\mathbb{C})$ be normal and $\rho \in \mathscr{D}_N$. The non-selective state update rule of ρ in A is defined as

$$U^A_{\text{n.s.}}(\rho) := \sum_{\lambda\in\sigma(A)} E^A(\lambda)\rho E^A(\lambda). \tag{2.137}$$

Similarly, given $S \subset \sigma(A)$ and assuming $E^A(S)\rho \neq 0$, the selective state update rule of ρ in $S \subset \sigma(A)$ is defined as

$$U^{A,S}_{\text{s}}(\rho) := \frac{1}{\operatorname{Tr}(\rho E^A(S))} E^A(S)\rho E^A(S). \tag{2.138}$$

a) Show that $U^{A}_{\text{n.s.}}(\rho), U^{A,S}_{\text{s}}(\rho) \in \mathscr{D}_N$.
b) Show that if ρ is a rank one projection and $E^A(\lambda)\rho \neq 0$, then $U^{A,S}_{\text{s}}(\rho)$ is a rank one projection.
c) Let ρ_ψ be a rank one projection. Show that $U^{A}_{\text{n.s.}}(\rho_\psi)$ is a rank one projection if and only if ψ is an eigenvector of A.

Think of the following model: Fix $S \subset \sigma(A)$ and $S' \subset \sigma(B)$. If a measurement of A in the state ω_ρ yields a value $\lambda \in S$, update the state ω_ρ to $\omega_{U^{A,S}_{\text{s}}(\rho)}$. One may then perform a new measurement in B.

d) Show that in general the order of updates plays a role, i.e. there exist normal $A, B \in M_N(\mathbb{C})$ s.t.

$$U^{A,S}_{\text{s}} \circ U^{B,S'}_{\text{s}} \neq U^{B,S'}_{\text{s}} \circ U^{A,S}_{\text{s}}.$$

e) If A, B are normal and commute, show that

$$\Pr_{U^{A,S}_{\text{s}}(\rho)}(B \in S') = \frac{1}{\mu^A_\rho(S)} \mu^{A,B}_\rho(S \times S'), \tag{2.139}$$

where μ^A_ρ and $\mu^{A,B}_\rho$ are the Borel probability measures associated to A and (A, B).

Exercise 2.4 (Schrödinger equation). On $\mathbb{C}^2$ consider the inital value problem for Schrödinger's Equation for

$$H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \psi_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{2.140}$$

a) Show that e^{itH} defined via spectral calculus and via the exponential series coincide.
b) Show that the unique solution of Schrödinger's equation for H, ψ_0 is given by

$$\psi(t) = \begin{pmatrix} \cos t \\ -i \sin t \end{pmatrix}. \tag{2.141}$$

Exercise 2.5 (Commuting observables). Let $H = H^* \in M_N(\mathbb{C})$ and $L = L^* \in M_N(\mathbb{C})$ commute.
a) Show that H leaves $E^L(\lambda)\mathbb{C}^N$ invariant for all $\lambda \in \sigma(L)$.
b) Show that $E^H(\mu)$ and $E^L(\lambda)$ commute for all $\mu \in \sigma(H), \lambda \in \sigma(L)$. What does this imply for the diagonalisation of H?
c) Discuss the difference between $\sigma(H, L)$, $\sigma(H) \times \sigma(L)$ and $\sigma(H)$.

Exercise 2.6 (Particle hopping on one-dimensional lattice). Consider X and V as in Problem 5.1 and denote by $(e_j)_{j=1,\dots,N}$ the eigenbasis of X. Recall that V is unitary, satisfies $Ve_j = e_{j+1}$ and its eigenvalues are

$$\sigma(V) = \{\nu_1, \dots, \nu_N\}, \qquad \nu_k := \omega^k, \qquad \omega := e^{2\pi i/N}, \tag{2.142}$$

and the vector

$$\xi_k := \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \omega^{-jk} e_j \tag{2.143}$$

is a normalized eigenvector of V with eigenvalue ν_k. Consider now

$$H := V + V^* = \sum_{k=1}^{N-1} (|e_k\rangle\langle e_{k+1}| + |e_{k+1}\rangle\langle e_k|) \in M_N(\mathbb{C}). \tag{2.144}$$

and consider the time evolution $\alpha_t^H(X)$ of the position observable X.

a) Show that X is a not a constant of motion.

b) Show that the spectral decomposition of H is given by

$$H = \sum_{j=1}^{N} \lambda_k |\xi_k\rangle\langle\xi_k|, \quad \sigma(H) = \{\lambda_1, \dots, \lambda_N\}, \quad \lambda_k = 2\cos\frac{2\pi k}{N}. \tag{2.145}$$

c) Show that

$$\alpha_t^H(X) = \sum_{k,l} e^{it(\lambda_k - \lambda_l)} |\xi_k\rangle\langle\xi_k, X\xi_l\rangle\langle\xi_l|. \tag{2.146}$$

d) Show that $\Pr_\psi(\alpha_t^H(X) = k) = |\langle e_k, e^{-itH}\psi\rangle|^2$ and discuss the result for

- $\psi = e_k$ for some k,
- $\psi = \xi_k$,
- $\psi = \frac{1}{\sqrt{2}}(e_r + e_l)$ for arbitrary r, l,
- $\psi = \frac{1}{\sqrt{2}}(\xi_r + \xi_l)$ for arbitrary r, l.

Exercise 2.7 (Dynamics of uncoupled product). Consider the uncoupled product $(\mathcal{A}, H)$ of two matrix mechanical systems $(M_{N_1}(\mathbb{C}), H_1)$ and $(M_{N_2}(\mathbb{C}), H_2)$. Show that for the dynamics it holds

$$\alpha_t^H = \alpha_t^{H_1} \otimes \alpha_t^{H_2}, \quad t \in \mathbb{R}. \tag{2.147}$$

Exercise 2.8 (Reduced density matrices). Let $M_{N_1}(\mathbb{C}) \otimes M_{N_2}(\mathbb{C})$ be a full bipartite system and denote for $\rho \in \mathcal{D}_{N_1 N_2}$ the reduced density matrices by $\rho_1 \in \mathcal{D}_{N_1}$ and $\rho_2 \in \mathcal{D}_{N_2}$.

a) Show that for $\rho_1 \in \mathcal{D}_{N_1}$ and $\rho_2 \in \mathcal{D}_{N_2}$, there exists a $\rho \in \mathcal{D}_{N_1 N_2}$ s.t. ρ_1 and ρ_2 are the reduced density matrices of ρ.

b) Show that the map

$$\mathcal{D}_{N_1 N_2} \to \mathcal{D}_{N_1} \times \mathcal{D}_{N_2}, \quad \rho \mapsto (\rho_1, \rho_2) \tag{2.148}$$

is not injective for $N_1, N_2 \geq 2$.

c) What happens in case N_1 or N_2 equals 1?

Exercise 2.9 (Schmidt decomposition). Use the singular value decomposition of an $N_1 \times N_2$-matrix to show that any $\psi \in \mathbb{C}^{N_1} \otimes \mathbb{C}^{N_2}$ can be written as

$$\psi = \sum_{i=1,\cdots,M} \lambda_i u_i \otimes v_i, \tag{2.149}$$

where $0 \leq M \leq \min\{N_1, N_2\}$, $(u_i)_{i=1,\cdots,M} \subset \mathbb{C}^{N_1}$ and $(v_i)_{i=1,\cdots,M} \subset \mathbb{C}^{N_2}$ are orthonormal sets and $\lambda_i > 0$ for all $i = 1, \cdots, M$.

Problem 2.1 (Matrix units). A system of matrix units in $M_N(\mathbb{C})$ is a family $E_{kl} \in M_N(\mathbb{C}), k, l \in \{1, \ldots, N\}$ such that for all $i, j, k, l \in \{1, \ldots, N\}$,

$$E_{kl}E_{ij} = \delta_{li}E_{kj}, \qquad (E_{kl})^* = E_{lk}, \qquad \sum_{k=1}^{N} E_{kk} = 1. \tag{2.150}$$

Show that every orthonormal basis (b_k) of $\mathbb{C}^N$ defines a system of matrix units via $E_{kl} := |b_k\rangle\langle b_l|$, and every system of matrix units is of this form. Show that a system of matrix units forms a basis of $M_N(\mathbb{C})$.

Problem 2.2 (The trace functional on $M_N(\mathbb{C})$). The matrix trace is the linear functional

$$\mathrm{Tr} : M_N(\mathbb{C}) \to \mathbb{C}, \qquad \mathrm{Tr}(A) := \sum_{k=1}^{N} A_{kk}. \tag{2.151}$$

Show that the matrix trace has the following properties:

a) Tr : $M_N(\mathbb{C})$ is linear and positive (i.e. $\mathrm{Tr}(A) \geq 0$ for $A \geq 0$),

b) Tr : $M_N(\mathbb{C})$ is faithful: If a positive matrix $A \geq 0$ satisfies $\mathrm{Tr}(A) = 0$, then $A = 0$.

c) Tr is invariant under cyclic permutations of its argument,

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA), \qquad A, B \in M_N(\mathbb{C}). \tag{2.152}$$

d) If $\tau : M_N(\mathbb{C}) \to \mathbb{C}$ is linear and satisfies $\tau(AB) = \tau(BA)$ for all $A, B \in M_N(\mathbb{C})$, then τ is a multiple of Tr.

Problem 2.3 (No uncorrelated states on $M_N(\mathbb{C})$). Let $\mathcal{H}$ be a finite dimensional Hilbert space with $N := \dim \mathcal{H} > 1$. In this exercise you show that $B(\mathcal{H})$ does not have any product (or completely uncorrelated) states. Let ω be a state on $\mathcal{B}(\mathcal{H})$ and assume

$$\mathrm{Cov}_\omega(A, B) = 0 \quad \text{for all } A, B \in B(\mathcal{H}).$$

Proceed in the following steps:

a) Show that $B(\mathcal{H}) \simeq M_N(\mathbb{C})$ as $*$-algebras.

b) Show that $\ker\omega = [M_N(\mathbb{C}), M_N(\mathbb{C})] := \mathrm{Span}\{AB - BA \ : \ A, B \in M_N(\mathbb{C})\}$.
c) Show that $[M_N(\mathbb{C}), M_N(\mathbb{C})] = \ker \mathrm{Tr}$ the set of all traceless matrices and conclude that $\omega = c\,\mathrm{Tr}$ for some $c \in \mathbb{C}^\times$.
d) Conclude that no state ω with $\mathrm{Cov}_\omega(A, B) = 0$ for all $A, B \in B(\mathcal{H})$ exists.

Problem 2.4 (Rotation on a discrete circle). Consider a particle on a cyclic one-dimensional finite lattice consisting of N sites. On the associated space $\mathbb{C}^N$ one considers the position observable

$$X := \mathrm{diag}(1, \dots, N) = \sum_{k=1}^{N} kE_{kk}, \tag{2.153}$$

which is a diagonal matrix (see Example 2.19) in the canonical orthonormal basis $(e_j)_{j=1,\cdot,N}$. In the identification $e_{N+1} \equiv e_1$, consider the matrix

$$V := \sum_{j=1}^{N} |e_{j+1}\rangle\langle e_j|. \tag{2.154}$$

a) Show that V is unitary and $V^N = 1$.
b) Show that

$$V = \sum_{j=0}^{N-1} \omega^j |\xi_j\rangle\langle\xi_j|,$$

where $\xi_j := \frac{1}{\sqrt{N}} \sum_{l=1}^{N} \omega^{-jl} e_l$ and $\omega = e^{\frac{1}{N} 2\pi i}$.
c) Show that $V^* X V = X + 1 - N|e_N\rangle\langle e_N|$.
d) Show that there exists no invertible $U \in M_N(\mathbb{C})$ s.t. $UXU^{-1} = X + 1$.

Problem 2.5 (Polynomial Functional Calculus). Denote by $\mathrm{Poly}(\mathbb{C})$ the space of polynomials in one variable with complex coefficients. Let $A \in M_N(\mathbb{C})$ and $p \in \mathrm{Poly}(\mathbb{C})$ and define

$$p(A) := \sum_{k=0}^{m} c_k A^k \in M_N(\mathbb{C}), \quad \text{where} \quad p(\lambda) = \sum_{k=0}^{m} c_k \lambda^k. \tag{2.155}$$

a) Show the spectral mapping principal $\sigma(p(A)) = p(\sigma(A))$.
b) Show that if A is normal, then $\|p(A)\| = \sup_{\lambda\in\sigma(A)} |p(\lambda)|$.
c) Show that if A is normal, then the map

$$\phi^A : \mathrm{Poly}(\mathbb{C}) \to M_N(\mathbb{C}), \quad \phi^A(p) := p(A) \tag{2.156}$$

is an isometric homomorphism of normed unital algebras, where $\mathrm{Poly}(\mathbb{C})$ is endowed with $\|p\| := \sup_{\lambda\in\sigma(A)} |p(\lambda)|$.

Let now A be normal and E^A its spectral measure and define for $f : \mathbb{C} \to \mathbb{C}$

$$\Phi^A(f) := \sum_{\lambda \in \sigma(A)} f(\lambda) E^A(\lambda). \tag{2.157}$$

d) Show that for any function $f : \mathbb{C} \to \mathbb{C}$ there exists $p \in \mathrm{Poly}(\mathbb{C})$ s.t. $p(\lambda) = f(\lambda)$ for all $\lambda \in \sigma(A)$. Given $S \subset \sigma(A)$, give an explicit expression for the polynomial that agrees with the characteristic function χ_S of S on $\sigma(A)$.

e) Show that for any $p \in \mathrm{Poly}(\mathbb{C})$ the definitions coincide, i.e.

$$\Phi^A(p) = p(A). \tag{2.158}$$

f) Show that $\Phi^A(f^*) = \Phi^A(f)^* = \Phi^{A^*}(f^\#)$, where $f^*(z) = \overline{f(z)}$ and $f^\#(z) = \overline{f(\overline{z})}$.

g) Given a state $\omega_\rho \in S(M_N(\mathbb{C}))$, show that (see equation (2.27))

$$\omega_\rho(\chi_S(A)) = \mu_\rho^A(S). \tag{2.159}$$

Problem 2.6 (Superposition states). Let $\psi = e_1, \xi = e_2 \in \mathbb{C}^2$ and consider the vector

$$\eta := \frac{1}{\sqrt{2}}(e_1 + e^{i\alpha} e_2), \tag{2.160}$$

where $\alpha \in \mathbb{R}$ is a parameter.

a) Show that η is a superposition vector.

b) Show that for $A \in M_2(\mathbb{C})$, $\langle \eta, A\eta \rangle = \omega_{P_\eta}(A)$ for $P_\eta = \frac{1}{2}\begin{pmatrix} 1 & e^{-i\alpha} \\ e^{i\alpha} & 1 \end{pmatrix}$.

c) Consider the density matrix $\rho = \frac{1}{2}P_{e_1} + \frac{1}{2}P_{e_2}$ and show that $\rho \neq P_\eta$.

Problem 2.7 (Derivations on $M_N(\mathbb{C})$). Let δ be a derivation of $M_N(\mathbb{C})$ and consider a system of matrix units $(E_{kl})_{kl}$ in $M_N(\mathbb{C})$.

a) Let δ_X and δ_Y be inner derivations satisfying $\delta_X = \delta_Y$, show that $X - Y \in \mathbb{C} \cdot 1$.

b) Define $X := -i \sum_{k=1}^N \delta(E_{k1}) E_{1k}$ and show that

$$\delta_X(E_{kl}) = \delta(E_{kl}), \quad k, l = 1, \cdots, N. \tag{2.161}$$

Conclude that every derivation is inner.

c) Consider an analytic one-parameter group of automorphisms $(\alpha_t)_{t \in \mathbb{R}}$. Show that

$$\delta := \frac{d}{dt}|_{t=0} \alpha_t \tag{2.162}$$

is a derivation of $M_N(\mathbb{C})$ and that $\delta \circ \alpha_t = \alpha_t \circ \delta$.

Problem 2.8 (Spectrum and periodicity). Let $H = H^* \in M_N(\mathbb{C})$ with either a) $\sigma(H) \subset \mathbb{Z}$ or b) $\sigma(H) \subset \frac{1}{2} + \mathbb{Z}$. Denote $U_H(t) = e^{itH}$ and $\alpha_t^H : A \mapsto e^{itH} A e^{-itH} \in \mathrm{Aut}\, M_N(\mathbb{C})$, $t \in \mathbb{R}$.

a) Show that U_H is periodic with period 2π for a) and 4π for b).

b) Show that α^H is periodic with period 2π in both cases a) and b).

Problem 2.9 (Unitary one-parameter groups). Show that on $\mathbb{C}^N$, the map between selfadjoint matrices and continuous unitary one-parameter groups, given by

$$H \mapsto (e^{itH})_{t\in\mathbb{R}}, \tag{2.163}$$

is a bijection.

Problem 2.10 (SU(2) and SO(3)). In this exercise we investigate the groups SU(2) and SO(3).

a) Show that $R_x(\theta)$, $\|x\| = 1, \theta \in \mathbb{R}$ as defined in Lemma 2.35 satisfies

$$R_x(\theta)p = \cos(\theta)p + (1 - \cos(\theta))(x \cdot p)x + \sin(\theta)(x \times p), \quad p \in \mathbb{R}^3, \tag{2.164}$$

where $\times$ denote the cross product on $\mathbb{R}^3$. *Hint:* Decompose p into $p_{||}$ and $p_\perp$ w.r.t. x.

b) Consider $L_1, L_2, L_3 \in M_3(\mathbb{C})$ as in Example 2.36. For $x \in \mathbb{R}^3$ define

$$x \cdot L := x_1 L_1 + x_2 L_2 + x_3 L_3 \in M_3(\mathbb{C}). \tag{2.165}$$

Show that $e^{i\theta x\cdot L}$, $\theta \in [0, \pi]$ is the rotation around the x-axis with angle θ.

c) Consider $\sigma_1, \sigma_2, \sigma_3 \in M_2(\mathbb{C})$ as in Example 2.38. Show that any $U \in \mathrm{SU}(2)$ has the form

$$U = e^{\frac{-i}{2}\theta x\cdot\sigma} \tag{2.166}$$

for some $x \in \mathbb{R}^3$, $\|x\| = 1$ and $\theta \in \mathbb{R}$. How unique are x, θ?

d) Show that the group homomorphism $\rho : \mathrm{SU}(2) \to \mathrm{SO}(3)$ of Prop. 2.41 satisfies

$$\rho(e^{\frac{-i}{2}\theta x\cdot\sigma}) = e^{i\theta x\cdot L}. \tag{2.167}$$

Conclude that ρ is surjective.
Hint: Use the identity $(x \cdot \sigma)(y \cdot \sigma) = (x \cdot y)1 + i(x \times y) \cdot \sigma$, where $\times$ denotes the cross product on $\mathbb{R}^3$.

e) Show that in coordinates

$$\rho(U)_{jk} = \frac{1}{2}\operatorname{Tr}\left(\sigma_j U \sigma_k U^*\right) \tag{2.168}$$

and conclude that ρ is continuous.

Problem 2.11 (SO(3), SU(2) and angular momentum operators). Let $U : \mathrm{SO}(3) \to U_N(\mathbb{C}) \subset M_N(\mathbb{C})$ be a unitary representation, i.e. a continuous group homomorphism into $U_N(\mathbb{C})$.

a) Consider the maps

$$U_i(\theta) : \mathbb{R} \to M_N(\mathbb{C}), \quad \theta \mapsto U(R_{e_i}(\theta)), \quad i \in \{1,2,3\}. \tag{2.169}$$

Show that U_i is a continuous unitary one parameter group.

b) Denote the generator of U_i by L_i^U. Show that (L_1^U, L_2^U, L_3^U) is a system of angular momentum operators.

c) A finite-dimensional representation on $\mathbb{C}^N$ of a group is called irreducible, if the only invariant subspaces are 0 and $\mathbb{C}^N$. Show that the unitary representation U is irreducible if and only if (L_1^U, L_2^U, L_3^U) is irreducible.

d) Show that a), b) and c) hold true for a unitary representation U of SU(2) and

$$U_i(\theta) : \mathbb{R} \to M_N(\mathbb{C}), \quad \theta \mapsto U(e^{\frac{-i}{2}\theta x\cdot\sigma}), \quad i \in \{1,2,3\}. \tag{2.170}$$

e) Show that for a representation $U : \mathrm{SU}(2) \to U_N(\mathbb{C})$ the map

$$V : \mathrm{SO}(3) \to U_N(\mathbb{C}), \quad V(\rho(A)) := U(A) \tag{2.171}$$

is well-defined if and only if $\{\pm 1\} \subset \ker(U)$, where ρ is given in Proposition 2.41. Show that in this case V is a unitary representation.

f) Show that for a representation $U : \mathrm{SU}(2) \to U_N(\mathbb{C})$, also in case $\{\pm 1\} \not\subseteq \ker(U)$, the map

$$\alpha : \mathrm{SO}(3) \to \mathrm{Aut}(M_N(\mathbb{C})), \quad \alpha(\rho(A))(X) := U(A)XU(A^{-1}) \tag{2.172}$$

is well-defined and a continuous representation of SO(3).

Problem 2.12 (Entanglement and entropy). Consider the function

$$f : (0,\infty) \to \mathbb{R}, \quad x \mapsto -x\ln(x). \tag{2.173}$$

a) Show that the unique continuous extension of f to $[0,\infty)$ vanishes at 0. This extension will also be denoted by f in the following.

For $\omega = \omega_\rho$ a state on a full matrix algebra, the von Neumann entropy $S(\rho)$ is defined as

$$S(\rho) = \operatorname{Tr}(f(\rho)) \in \mathbb{R}. \tag{2.174}$$

b) Show that $S(\rho) = 0$ if and only if ω_ρ is a pure state.

Consider now the full bipartite system $\mathcal{A} = M_{N_1}(\mathbb{C}) \otimes M_{N_2}(\mathbb{C})$. We define the entanglement entropy of pure states on $\mathcal{A}$ by

$$E : S_{\text{pure}}(\mathcal{A}) \to \mathbb{R}, \quad E(P) := S(P_1), \tag{2.175}$$

where P_1 is the 1-reduced density matrix of P. Show the following properties:

c) $E(P) = S(P_2)$ (independence of choice of subsystem),

d) $E(P) = 0$ if and only if P is not entangled (entanglement measure),

e) $E(P) \in [0, \infty)$ (positivity),

f) $\sup_P E(P) = \log(\min\{N_1, N_2\})$ (upper bound),

g) $E((U_1 \otimes U_2)P(U_1 \otimes U_2)^*) = E(P)$ (invariance under local operations).

Hint: For c) and f) use the Schmidt decomposition. For f) use additionally that $\ln : (0, \infty) \to \mathbb{R}$ is a concave function, i.e. for $\sum_{i=1}^M \alpha_i = 1$ and $x_i \in (0, \infty)$

$$\sum_{i=1}^M \alpha_i \ln(x_i) \leq \ln\left(\sum_{i=1}^M \alpha_i x_i\right).$$

Give an example of a pure state on $\mathcal{A}$ that attains the upper bound in e).

Why is E not a good entanglement measure for mixed states on $\mathcal{A}$?

Problem 2.13 (Entangled states). Consider in $\mathbb{C}^2 \otimes \mathbb{C}^2$

$$\psi_1 = \frac{1}{\sqrt{2}}(e_1 \otimes e_1 + e_2 \otimes e_2), \quad \psi_2 = \frac{1}{\sqrt{2}}(e_1 \otimes e_1 - e_2 \otimes e_2), \tag{2.176}$$

$$\psi_3 = \frac{1}{\sqrt{2}}(e_1 \otimes e_2 - e_2 \otimes e_1), \quad \psi_4 = (e_1 \otimes e_2). \tag{2.177}$$

Compute the reduced density matrices of the following density matrices and decide, which of them define entangled states.

$$P_{\psi_i}, \quad \text{for} \quad i = 1, 2, 3, 4 \tag{2.178}$$

and

$$D = \frac{1}{4}\Big(\Big(\sum_{i,j=1,2} |e_i \otimes e_j\rangle\langle e_i \otimes e_j|\Big) + i|e_2 \otimes e_2\rangle\langle e_1 \otimes e_2| - i|e_1 \otimes e_2\rangle\langle e_2 \otimes e_2|\Big). \tag{2.179}$$

Problem 2.14 (Werner states). Consider the singlet vector state given by

$$\psi = \frac{1}{\sqrt{2}}(e_2 \otimes e_1 - e_1 \otimes e_2) \in \mathbb{C}^2 \otimes \mathbb{C}^2$$

and for a probability parameter $0 \leq p \leq 1$, the density matrix

$$\rho_p := p P_\psi + \frac{1-p}{4} 1,$$

and the state $\omega_p := \omega_{\rho_p} \in S(M_2 \otimes M_2)$. Show the following equivalences:

a) ω_p is a product state $\iff p = 0$,

b) ω_p is separable $\iff 0 < p \leq \frac{1}{3}$,

c) ω_p is entangled $\iff p > \frac{1}{3}$,

d) ω_p satisfies the Bell Inequality

$$\sup_{A_1,A_2,B_1,B_2 \text{ binary}} |\mathbb{b}_{\omega_p}(A_1 \otimes 1, A_2 \otimes 1, 1 \otimes B_1, 1 \otimes B_2)| \leq 2$$

if and only if $p > \frac{1}{\sqrt{2}}$.

Conclude that Bell correlation is no good entanglement measure.

Hints: For b), first show that $\omega_{1/3}$ is separable. To do so, consider the density matrices $\tau_x := \frac{1}{2}(1 + x \cdot \sigma) \in \mathcal{D}_2$, the tensor products $\tau_x \otimes \tau_{-x}$, and averages over $x \in \{\pm e_1, \pm e_2, \pm e_3\}$.

For c), first prove that the tensor flip F satisfies $\omega(F) \geq 0$ in every separable state ω, and then use this criterion on ω_p (this is called an "entanglement witness").

For d), follow the analysis presented in the lectures for the singlet state ψ.

3 General quantum mechanical systems (quantum mechanics)

Having learned many concepts and results of quantum mechanics in its finite-dimensional matrix mechanics form, we now move on to consider general quantum mechanical systems. Conceptually, not much changes, but on a mathematical level, we have to work with infinite-dimensional Hilbert spaces and operators on them rather than only $\mathbb{C}^N$ and matrices in $M_N(\mathbb{C})$. Background knowledge in functional analysis, in particular about Hilbert space operators and spectral theory, is required from now on.

See Appendix M.4 for some reminders, and [RS72; Sch12] for detailed text books. We will also need the Fourier transform, recalled briefly in Appendix M.3.

3.1 From canonical commutation relations to Schrödinger representations

A main problem of matrix mechanics is that it does not give us counterparts of the fundamental position and momentum observables of classical mechanics. Recall that given a phase space $M = U \times \mathbb{R}^d \subset \mathbb{R}^{2d}$, these are the maps (1.39), $k \in \{1, \dots, d\}$,

$$X_k^{\mathrm{cl}} : M \to \mathbb{R}, \qquad (x, p) \mapsto x_k, \tag{3.1}$$

$$P_k^{\mathrm{cl}} : M \to \mathbb{R}, \qquad (x, p) \mapsto p_k, \tag{3.2}$$

where we have included the superscript "cl" for "classical" for emphasis. We have seen that their Poisson brackets are given by (1.44), $k, l \in \{1, \dots, d\}$,

$$\{X_k^{\mathrm{cl}}, X_l^{\mathrm{cl}}\} = 0, \qquad \{P_k^{\mathrm{cl}}, P_l^{\mathrm{cl}}\} = 0, \qquad \{X_k^{\mathrm{cl}}, P_l^{\mathrm{cl}}\} = \delta_{k,l} \cdot 1. \tag{3.3}$$

In matrix mechanics, we have learned that the quantum analogue of the classical Poisson bracket is the commutator times $(-i)$. It is therefore natural to search for selfadjoint elements $X_1, \dots, X_d, P_1, \dots, P_d$ of a unital *-algebra $\mathcal{A}$ satisfying

$$[X_k, X_l] = 0, \qquad [P_k, P_l] = 0, \qquad [X_k, P_l] = i\delta_{k,l} \cdot 1, \tag{3.4}$$

as a replacement for the classical relations. We remark that in this case, the X_k, P_l also satisfy Heisenberg's uncertainty relations (2.1) (with $\hbar = 1$): For any state $\omega \in S(\mathcal{A})$,

$$\Delta_\omega(X_k) \cdot \Delta_\omega(P_l) \geq \frac{\delta_{kl}}{2}. \tag{3.5}$$

The proof of this follows immediately from the general uncertainty relation (Proposition 2.9) and the normalization of ω.

We are therefore well motivated to analyze the relations (3.4). For a basis-independent approach, it is best to consider the *symplectic space* $(\mathbb{R}^{2d}, \sigma)$, with the symplectic form (bilinear antisymmetric non-degenerate map)

$$\sigma : \mathbb{R}^{2d} \times \mathbb{R}^{2d} \to \mathbb{R}, \tag{3.6}$$

$$\sigma(\xi, \xi') := \langle \xi, \Sigma \xi' \rangle_{\mathbb{R}^{2d}} = \left\langle \begin{pmatrix} x \\ p \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x' \\ p' \end{pmatrix} \right\rangle_{\mathbb{R}^{2d}} = \langle x, p' \rangle_{\mathbb{R}^d} - \langle x', p \rangle_{\mathbb{R}^d}, \tag{3.7}$$

where we have used the canonical symplectic matrix Σ (1.33) and written $\xi = (x, p)$, $\xi' = (x', p')$ for the position/momentum parts of the phase space vectors ξ, ξ'.

Definition 3.1 (CCR algebra).

a) The *CCR algebra* (algebra of *canonical commutation relations*) $\mathrm{CCR}(\mathbb{R}^{2d}, \sigma)$ over $(\mathbb{R}^{2d}, \sigma)$ is the unital *-algebra generated[a] by elements $\phi(\xi)$, $\xi \in \mathbb{R}^{2d}$, such that

$$\begin{aligned} [\phi(\xi), \phi(\xi')] &= i\sigma(\xi, \xi') \cdot \mathbb{1}, \\ \phi(\xi)^* &= \phi(\xi), \end{aligned} \tag{3.8}$$

holds for all $\xi, \xi' \in \mathbb{R}^{2d}$.

b) A *representation* of the CCR algebra in another unital *-algebra $\mathcal{A}$ is a homomorphism of unital *-algebras $\mathrm{CCR}(\mathbb{R}^{2d}, \sigma) \to \mathcal{A}$.

[a] That is, $\mathrm{CCR}(\mathbb{R}^{2d}, \sigma)$ consists of all polynomials in the $\phi(\xi)$. See an algebra lecture for a more detailed explanation of "generated algebra" (one starts from a free algebra and divides by the ideal generated by the relations).

In the CCR algebra, the elements $\phi(\xi)$ are abstract generating elements, and in a representation, they are mapped to elements of $\mathcal{A}$ satisfying the same relations (3.8). Consider a representation mapping $\phi(\xi) \in \mathrm{CCR}(\mathbb{R}^{2d}, \sigma)$ to $\phi_{\mathcal{A}}(\xi) \in \mathcal{A}$, and set $X_k := \phi_{\mathcal{A}}((e_k, 0))$ and $P_l := \phi_{\mathcal{A}}((0, e_l))$. Then

$$[X_k, P_l] = i\sigma((e_k, 0), (0, e_l)) = i\delta_{kl}\mathbb{1},$$

in agreement with (3.4).

Lemma 3.2. *Let $N \in \mathbb{N}$. There exist no representations of $\mathrm{CCR}(\mathbb{R}^{2d}, \sigma)$ in $M_N(\mathbb{C})$.*

Proof. Assume a representation $\phi_{\mathcal{A}} : \mathrm{CCR}(\mathbb{R}^{2d}, \sigma) \to M_N(\mathbb{C})$ exists. Then $X_k = \phi_{\mathcal{A}}((e_k, 0))$, $P_k = \phi_{\mathcal{A}}((0, e_k))$ satisfy $[X_k, P_k] = i\mathbb{1}_{\mathcal{A}}$, which evaluates in the trace of $M_N(\mathbb{C})$ to

$$0 = \mathrm{Tr}(X_k P_k - P_k X_k) = \mathrm{Tr}([X_k, P_k]) = \mathrm{Tr}(i) = iN, \tag{3.9}$$

a contradiction. □

This lemma can be seen as an explanation why we did not find good replacements for position and momentum in matrix mechanics: The finite-dimensional matrix algebras are "too small" to allow for representations of the canonical commutation relations.

We next want to show that on infinite-dimensional spaces, representations of $\mathrm{CCR}(\mathbb{R}^{2d}, \sigma)$ do exist. To that end, we first recall some function spaces known from the analysis courses: We write $\mathrm{Pol}(\mathbb{R}^d)$ for the space of all polynomial functions $\mathbb{R}^d \to \mathbb{C}$ (polynomials in d variables), $C_c^\infty(\mathbb{R}^d)$ for the space of all smooth compactly supported functions, and $L^2(\mathbb{R}^d)$ for the space of all square integrable (w.r.t. Lebesgue measure ℓ on $\mathbb{R}^d$) functions $\psi : \mathbb{R}^d \to \mathbb{C}$. We will also need the slightly less familiar Schwartz space, defined as follows.

Definition 3.3 (Schwartz space). A *Schwartz function* on $\mathbb{R}^d$ is a smooth function $\psi : \mathbb{R}^d \to \mathbb{C}$ such that for every two multiindices[a] $\alpha, \beta \in \mathbb{N}_0^d$,

$$\sup_{x \in \mathbb{R}^d} |x^\alpha (\partial^\beta \psi)(x)| < \infty. \tag{3.10}$$

The *Schwartz space* is the linear space $\mathcal{S}(\mathbb{R}^d)$ of all Schwartz functions on $\mathbb{R}^d$.

[a] A multiindex is an element $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$. Its length is $|\alpha| := \alpha_1 + \ldots + \alpha_d$. For

$x \in \mathbb{R}^d$, one sets $x^\alpha := x_1^{\alpha_1} \cdots x_d^{\alpha_d}$, and $\partial^\alpha := \partial_{x_1}^{\alpha_1} \cdots \partial_{x_d}^{\alpha_d}$ for partial derivatives.

By definition of Schwartz space, we have

$$f \in \mathrm{Pol}(\mathbb{R}^d),\ \alpha \in \mathbb{N}_0^d,\ \psi \in \mathcal{S}(\mathbb{R}^d) \implies f\psi \in \mathcal{S}(\mathbb{R}^d), \quad \partial^\alpha \psi \in \mathcal{S}(\mathbb{R}^d). \tag{3.11}$$

Furthermore,

$$C_c^\infty(\mathbb{R}^d) \subsetneq \mathcal{S}(\mathbb{R}^d) \subsetneq L^2(\mathbb{R}^d), \tag{3.12}$$

and we recall that $C_c^\infty(\mathbb{R}^d)$ (and in particular $\mathcal{S}(\mathbb{R}^d)$) is a dense subspace of $L^2(\mathbb{R}^d)$. Here we of course consider the Hilbert space $L^2(\mathbb{R}^d)$ with the norm $\|\cdot\|_2$ given by its scalar product

$$\langle \psi, \varphi \rangle_{L^2(\mathbb{R}^d)} := \int_{\mathbb{R}^d} \overline{\psi(x)} \varphi(x)\, dx, \tag{3.13}$$

namely $\|\psi\|_2^2 = \langle \psi, \psi \rangle_{L^2(\mathbb{R}^d)}$. We will mostly drop the subscript on the scalar product when it is clear from context that the L^2-scalar product is meant.

Definition 3.4 (Position and momentum operators on $\mathcal{S}(\mathbb{R}^d)$).

a) A linear differential operator with polynomial coefficients is a map of the form

$$D : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d), \qquad (D\psi)(x) = \sum_\alpha M_{f_\alpha} \partial^\alpha \psi, \tag{3.14}$$

where the sum is a finite sum over multiindices $\alpha \in \mathbb{N}_0^d$, the $f_\alpha \in \mathrm{Pol}(\mathbb{R}^d)$ are polynomials in $x_1, \ldots, x_d$, and $M_{f_\alpha}\psi := f_\alpha \cdot \psi$. The set of all these maps is denoted $\mathrm{DiffPol}(\mathbb{R}^d)$.

b) The (*Schrödinger*) *position/momentum operators* are the maps $X_k, P_l \in \mathrm{DiffPol}(\mathbb{R}^d)$ given by, $k, l \in \{1, \ldots, d\}$,

$$\begin{aligned} X_k &: \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d), \quad (X_k\psi)(x) := x_k \psi(x), \\ P_l &: \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d), \quad (P_k\psi)(x) := -i(\partial_{x_l}\psi)(x), \end{aligned} \tag{3.15}$$

Every $D \in \mathrm{DiffPol}(\mathbb{R}^d)$ is a well-defined linear map $\mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ because of (3.11).

Proposition 3.5 (The Schrödinger Representation).

a) $\mathrm{DiffPol}(\mathbb{R}^d)$ *is a unital* **-algebra with the* **-operation given by*

$$\left(\sum_\alpha M_{f_\alpha} \partial^\alpha \right)^* := \sum_\alpha (-1)^{|\alpha|} \partial^\alpha \circ M_{\overline{f_\alpha}}, \qquad f_\alpha \in \mathrm{Pol}(\mathbb{R}^d). \tag{3.16}$$

b) *For* $\psi, \varphi \in \mathcal{S}(\mathbb{R}^d)$*, one has*

$$\langle \psi, D\varphi \rangle = \langle D^*\psi, \varphi \rangle \tag{3.17}$$

(*the scalar product is the* L^2*-scalar product*).

c) *The map* $\mathrm{CCR}(\mathbb{R}^{2d}, \sigma) \to \mathrm{DiffPol}(\mathbb{R}^d)$ *given by*

$$\phi((e_k, 0)) \mapsto X_k, \qquad \phi((0, e_l)) \mapsto P_l \tag{3.18}$$

defines a representation of $\mathrm{CCR}(\mathbb{R}^{2d}, \sigma)$ *in* $\mathrm{DiffPol}(\mathbb{R}^d)$.

Proof. a) and b) will be proven in the exercises. For c), we have to verify the commutation relations (3.4).

As multiplication operators, the X_k commute among themselves, i.e. $[X_k, X_l] = 0$ for all k, l. The same holds true for the momentum operators, i.e. $[P_k, P_l] = 0$ for all k, l, because partial derivatives of smooth functions commute. For the mixed commutator, we have by the Leibniz rule, $\psi \in \mathcal{S}(\mathbb{R}^d)$, $x \in \mathbb{R}^d$,

$$([X_k, P_l]\psi)(x) = x_k \cdot (-i(\partial_l\psi)(x)) + i\partial_l(x_k\psi(x)) = i\delta_{kl}\psi(x),$$

which implies

$$[X_k, P_l] = i\delta_{kl}\mathbb{1}_{\mathrm{DiffPol}(\mathbb{R}^d)} \tag{3.19}$$

because ψ and x were arbitrary.

It remains to show $X_k^* = X_k$ and $P_k^* = P_k$. With the linear polynomial $f_k(x) = x_k$, we have $X_k = M_{f_k}$, hence $X_k^* = M_{\overline{f_k}} = X_k$ according to the *-operation from a). For the momentum operators, we have $P_k = -i\partial_k$, corresponding to the multiindex $\alpha = e_k$ of length $|\alpha| = 1$. Hence $P_k^* = (-1)^1\overline{(-i)}\partial_k = i\partial_k = P_k$. □

The representation constructed in part c) of this proposition is called the *Schrödinger representation of* $\mathrm{CCR}(\mathbb{R}^{2d}, \sigma)$. We will see later that it is not just some representation of the CCR algebra $\mathrm{CCR}(\mathbb{R}^{2d}, \sigma)$, but a special one.

Every $D \in \mathrm{DiffPol}(\mathbb{R}^d)$ can be written as a polynomial in the $2d$ operators $X_1, \dots, X_d$, $P_1, \dots, P_d$. As these do not commute, the order matters, e.g. $P_1^2X_1 \neq P_1X_1P_1$. In case only position operators $X_1, \dots, X_d$ are involved, or only momentum operators $P_1, \dots, P_d$, the order does not matter because these operators commute. In particular, we may unambiguously speak about a polynomial in the X's, say $f(X) \in \mathrm{DiffPol}(\mathbb{R}^d)$ (and similarly $f(P) \in \mathrm{DiffPol}(\mathbb{R}^d)$) for given $f \in \mathrm{Pol}(\mathbb{R}^d)$.

In contrast to the commutative algebra $C^\infty(\mathbb{R}^{2d})$ or the finite-dimensional algebra $M_N(\mathbb{C})$, the state space $S(\mathrm{DiffPol}(\mathbb{R}^d))$ is not easy to describe completely. Nonetheless, it is easy to give examples of states. The following proposition partially relies on the Fourier transform (see Appendix M.3).

Proposition 3.6 (Vector states on $\mathrm{DiffPol}(\mathbb{R}^d)$**).** *Let* $\psi \in \mathcal{S}(\mathbb{R}^d)$ *be* L^2*-normalized, i.e.* $\|\psi\|_2 = 1$.

a) *The functional*

$$\omega_\psi : \mathrm{DiffPol}(\mathbb{R}^d) \to \mathbb{C}, \qquad \omega_\psi(D) := \langle\psi, D\psi\rangle \tag{3.20}$$

is a state of $\mathrm{DiffPol}(\mathbb{R}^d)$.

b) *Let* ℓ *denote Lebesgue measure on* $\mathbb{R}^d$. *Then the measure* $|\psi|^2\ell$ *with density* $|\psi|^2$ *(i.e.* $d|\psi|^2\ell(x) = |\psi(x)|^2\,dx$*) is a probability measure on* $\mathbb{R}^d$, *and*

$$\omega_\psi(f(X)) = \int_{\mathbb{R}^d} f(x)\,|\psi(x)|^2\,dx, \qquad f \in \mathrm{Pol}(\mathbb{R}^d). \tag{3.21}$$

c) *Let* ℓ *denote Lebesgue measure on* $\mathbb{R}^d$. *Then the measure* $|\tilde\psi|^2\ell$ *(here* $\tilde\psi$ *denotes the Fourier transform of* ψ*) with density* $|\tilde\psi|^2$ *(i.e.* $d|\tilde\psi|^2\ell(p) = |\tilde\psi(p)|^2\,dp$*) is a*

probability measure on $\mathbb{R}^d$, and

$$\omega_\psi(f(P)) = \int_{\mathbb{R}^d} f(p)\,|\tilde\psi(p)|^2\,dp, \qquad f \in \mathrm{Pol}(\mathbb{R}^d). \tag{3.22}$$

Proof. a) As $D\,\mathcal{S}(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$, the scalar product $\langle\psi, D\psi\rangle$ is well-defined for $\psi \in \mathcal{S}(\mathbb{R}^d)$. It is clearly linear in D and satisfies $\omega_\psi(\mathbb{1}) = \langle\psi,\psi\rangle = \|\psi\|^2 = 1$ by normalization of ψ. For positivity, we use Prop. 3.5 b) and get $\omega_\psi(D^*D) = \langle D\psi, D\psi\rangle = \|D\psi\|^2 \geq 0$. Hence $\omega_\psi \in S(\mathrm{DiffPol}(\mathbb{R}^d))$.

b) The function $|\psi|^2$ lies in $L^1(\mathbb{R}^d)$, with $\||\psi|^2\|_1 = \|\psi\|_2^2 = 1$, hence $|\psi|^2\ell$ is a probability measure on $\mathbb{R}^d$. For any polynomial $f \in \mathrm{Pol}(\mathbb{R}^d)$, we have $f(X)\psi = f\cdot\psi$, and therefore

$$\omega_\psi(f(X)) = \langle\psi, f\cdot\psi\rangle = \int_{\mathbb{R}^d} \overline{\psi(x)} f(x)\psi(x)\,dx = \int_{\mathbb{R}^d} f(x)|\psi(x)|^2\,dx.$$

c) With ψ, also the Fourier transform $\tilde\psi$ lies in $\mathcal{S}(\mathbb{R}^d)$ (Theorem M.23 a)), and we have $\|\tilde\psi\|_2 = \|\psi\|_2 = 1$ (Proposition M.24). Thus by the same argument as in b), we see that $|\tilde\psi|^2\ell$ is a probability measure on $\mathbb{R}^d$, and

$$\omega_\psi(f(P)) = \langle\psi, f(P)\psi\rangle = \langle\mathcal{F}\psi, \mathcal{F}f(P)\psi\rangle,$$

where $\mathcal{F}$ denotes the Fourier transform, and we have used the Plancherel Theorem (Prop. M.24) again. To simplify $\mathcal{F}f(P)\psi$, we first consider the linear polynomial $f(P) = P_k$. According to Theorem M.23 c), we have $\mathcal{F}P_k\psi = X_k\mathcal{F}\psi$. For a multinomial $f(P) = P_{k_1}\cdots P_{k_n}$, we get

$$\mathcal{F}P_{k_1}\cdots P_{k_n}\psi = X_{k_1}\mathcal{F}P_{k_2}\cdots P_{k_n}\psi = X_{k_1}\cdots X_{k_n}\mathcal{F}\psi = f\cdot\mathcal{F}\psi.$$

By linearity, this implies for general polynomials $\mathcal{F}f(P)\psi = f\cdot\mathcal{F}\psi$, and therefore

$$\omega_\psi(f(P)) = \langle\mathcal{F}\psi, f(X)\mathcal{F}\psi\rangle = \int_{\mathbb{R}^d} f(p)|\tilde\psi(p)|^2\,dp$$

as in b). □

As for any unital *-algebra, the state space of $\mathrm{DiffPol}(\mathbb{R}^d)$ is a convex set, i.e. convex combinations of vector states are also states (mixed ones). This gives us many more examples of states, namely

$$D \mapsto \sum_{n=1}^N \lambda_n\langle\psi_n, D\psi_n\rangle, \tag{3.23}$$

where $\psi_n \in \mathcal{S}(\mathbb{R}^d)$ have unit L^2-norms, and the $\lambda_n \geq 0$ sum to 1. As in matrix mechanics, we can also form superpositions (Def. 2.22) of vector states.

As explained above, the integrals (3.21) and (3.22) exist for all polynomials $f \in \mathrm{Pol}(\mathbb{R}^d)$ because Schwartz functions decay faster than any inverse polynomial. We therefore have two probability measures, $|\psi|^2\ell$ and $|\tilde\psi|^2\ell$, on $\mathbb{R}^d$ for which all mixed moments (cf. (1.86)) are finite. We state here without proof that in case ψ decays exponentially in the sense that there exists $a > 0$ such that $\int_{\mathbb{R}^d} e^{a\|x\|}|\psi(x)|^2\,dx < \infty$ (this is not true for all $\psi \in \mathcal{S}(\mathbb{R}^d)$, but

of course in particular satisfied for $\psi \in C_c^\infty(\mathbb{R}^d)$), it follows that these moments uniquely determine the measure $|\psi|^2\ell$ among all Borel probability measures on $\mathbb{R}^d$, and analogously for $|\tilde{\psi}|^2\ell$.

When considering only functions (polynomials) of the position operators (or only the momentum operators), we are therefore in a situation quite analogous to classical mechanics and interpret these measures as the probability measures for position and momentum (cf. Def. 1.27 and Def. 1.28).

Definition 3.7 (Position and momentum probability measures). For $\psi \in \mathcal{S}(\mathbb{R}^d)$, $\|\psi\| = 1$, the measure $|\psi|^2\ell$ is the *probability measure for position*, and the measure $|\tilde{\psi}|^2\ell$ is the *probability measure for momentum*.

In particular, given a Borel set $B \subset \mathbb{R}^d$,

$$\Pr_\psi(X \in B) = \int_B |\psi(x)|^2 \, dx \tag{3.24}$$

is the probability that the position of the particle described by ψ lies in B, and

$$\Pr_\psi(P \in B) = \int_B |\tilde{\psi}(p)|^2 \, dp \tag{3.25}$$

is the probability that the momentum of the particle described by ψ lies in B.

As in classical mechanics, we can derive expectation values, uncertainties, and correlations from the state ω_ψ, and analogously for convex combinations of such states.

Note, however, one crucial difference to classical mechanics: In classical mechanics, states are given by probability measures on *phase space*. In particular, we may consider measures $\nu_X \times \nu_P$ of product form, i.e. prescribe the probability distributions of position and momentum independently. In quantum mechanics, however, the probability measures for position and momentum, $|\psi|^2\ell$ and $|\tilde{\psi}|^2\ell$, are not independent because they both derive from the same function ψ.

Example 3.8. Working in dimension $d = 1$ for simplicity, the expectation value of $D := XP$ in ω_ψ is

$$\omega_\psi(XP) = \langle \psi, XP\psi \rangle = -i \int_{\mathbb{R}} \overline{\psi(x)} x \psi'(x) \, dx.$$

Using the Fourier transform, this can also be written as

$$\begin{aligned} \omega_\psi(XP) = \langle \tilde{\psi}, \mathcal{F}XP\psi \rangle = -\langle \tilde{\psi}, PX\tilde{\psi} \rangle = i \int_{\mathbb{R}} \overline{\tilde{\psi}(p)} \partial_p (p\psi(p)) \, dp \\ = i + i \int_{\mathbb{R}} \overline{\tilde{\psi}(p)} p \partial_p \tilde{\psi}(p) \, dp, \end{aligned}$$

but the expression does not simplify to integration against a measure on phase space, as in classical mechanics. Certainly it is typically *not* true that $\omega_\psi(XP) = \int_{\mathbb{R}^2} xp \, d\nu_\psi(x, p)$

and $\omega_\psi(PX) = \int_{\mathbb{R}^2} px\, d\nu_\psi(x,p)$ for some measure ν_ψ, because the integrals coincide (the numbers x and p commute), but the operators X and P do not commute.

⚛ **Position and momentum space.** Proposition 3.6 used the fact that Schwartz space $\mathcal{S}(\mathbb{R}^d)$ is invariant under Fourier transform, and led to the two probability measures $|\psi|^2\ell$ and $|\tilde{\psi}|^2\ell$. As the probability for the position lying in a Borel set $B \subset \mathbb{R}^d$ is given by an integral over $x \mapsto |\psi(x)|^2$, we should think of the argument of ψ as a position variable, hence the symbol "x". The probability for the *momentum* lying in a Borel set $B \subset \mathbb{R}^d$ is however given by an integral over $p \mapsto |\tilde{\psi}(p)|^2$, so we should think of the argument of $\tilde{\psi}$ as a momentum variable, hence the symbol "p".

It does of course not matter which symbols we use for the variables, but this convention helps to keep track of the significance of the variables under consideration. In physics, one speaks of working "in position space" or "in momentum space", respectively.

At this stage we have found quantum counterparts X and P of our most important classical position and momentum X^{cl} and P^{cl}. Such a correspondence $X^{\mathrm{cl}} \rightsquigarrow X$, $P^{\mathrm{cl}} \rightsquigarrow P$ is called *quantization*. We would of course like to go further and understand quantum counterparts of more, or maybe even all, classical observables. Restricting ourselfes to polynomial observables for now, we would like to set up a quantization map $\mathrm{Pol}(\mathbb{R}^{2d}) \to \mathrm{DiffPol}(\mathbb{R}^d)$ that maps classical observables to quantum observables. We start to define such a map by mapping

$$P_k^{\mathrm{cl}} \mapsto P_k, \qquad X_k^{\mathrm{cl}} \mapsto X_k, \qquad 1_{\mathbb{R}^{2d}} \mapsto \mathrm{id}_{\mathcal{S}(\mathbb{R}^d)} = \mathbb{1}_{\mathrm{DiffPol}(\mathbb{R}^d)}. \tag{3.26}$$

It is natural to extend this correspondence by mapping polynomials in the X_k^{cl}'s (or the P_k^{cl}'s) to the corresponding polynomials in the X_k's (or the P_k's), but things get more complicated for products involving position *and* momentum factors. For example, we should not map the polynomial $f(x,p) = x_1 p_1$ (this is $X_1^{\mathrm{cl}} P_1^{\mathrm{cl}}$) to $X_1 P_1$ because

$$(X_1P_1)^* = P_1X_1 = [P_1, X_1] + X_1P_1 = X_1P_1 - i\mathbb{1} \tag{3.27}$$

is not selfadjoint. As observables must be selfadjoint in order to give real expectation values and non-negative uncertainties, this is unacceptable. A possible way out would be to map $X_1^{\mathrm{cl}} P_1^{\mathrm{cl}} = \frac{1}{2}(X_1^{\mathrm{cl}} P_1^{\mathrm{cl}} + P_1^{\mathrm{cl}} X_1^{\mathrm{cl}})$ to the selfadjoint element $\frac{1}{2}(X_1P_1 + P_1X_1)$, the symmetrization of X_1P_1. This is done systematically for all polynomials by the Weyl quantization map.

Definition 3.9 (Weyl quantization). The *Weyl quantization* is the linear map

$$\begin{aligned} W &: \mathrm{Pol}(\mathbb{R}^{2d}) \to \mathrm{DiffPol}(\mathbb{R}^d), \\ (W(f)\psi)(x) &:= (2\pi)^{-d} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{i\langle x-y,p\rangle} f\left(\tfrac{x+y}{2}, p\right)\psi(y)\, dy\, dp. \end{aligned} \tag{3.28}$$

We first note that the integral (3.28) is well-defined: Any polynomial f is a linear combination of monomials $f_{\alpha\beta}(x,p) = x^\alpha p^\beta = x_1^{\alpha_1}\cdots x_d^{\alpha_d} p_1^{\beta_1}\cdots p_d^{\beta_d}$ (multiindex notation).

The inner integral over y takes the Fourier transform of the Schwartz function ψ times a polynomial in y, which is then multiplied by p^β and inverse Fourier transformed. By the binomial theorem[14] and Fourier transform, we obtain explicitly

$$\begin{aligned}(W(f_{\alpha\beta})\psi)(x) &= (2\pi)^{-d}2^{-|\alpha|}\sum_{\mu\le\alpha}\binom{\alpha}{\mu}x^{\alpha-\mu}\int_{\mathbb{R}^d}e^{i\langle x,p\rangle}p^\beta\int_{\mathbb{R}^d}e^{-i\langle y,p\rangle}y^\mu\psi(y)\,dy\,dp\\ &= (2\pi)^{-d/2}2^{-|\alpha|}\sum_{\mu\le\alpha}\binom{\alpha}{\mu}x^{\alpha-\mu}\int_{\mathbb{R}^d}e^{i\langle x,p\rangle}p^\beta((i\partial_p)^\mu\tilde{\psi})(p)\\ &= 2^{-|\alpha|}\sum_{\mu\le\alpha}\binom{\alpha}{\mu}x^{\alpha-\mu}((-i\partial_x)^\beta x^\mu\psi)(x),\end{aligned}$$

namely

$$W(f_{\alpha\beta}) = 2^{-|\alpha|}\sum_{\mu\le\alpha}\binom{\alpha}{\mu}X^{\alpha-\mu}P^\beta X^\mu. \tag{3.29}$$

Here we have extended our multiindex notation to commuting operators, i.e. $X^\mu = X_1^{\mu_1}\cdots X_d^{\mu_d}$, etc.

As (3.29) lies in $\mathrm{DiffPol}(\mathbb{R}^d)$ and is selfadjoint, we conclude that W maps indeed into $\mathrm{DiffPol}(\mathbb{R}^d)$, and maps real polynomials to selfadjoint elements of $\mathrm{DiffPol}(\mathbb{R}^d)$. For example, in dimension $d = 1$,

$$\begin{aligned}W(X^{\mathrm{cl}}P^{\mathrm{cl}}) &:= W(f_{11}) = \frac{1}{2}(XP+PX),\\ W((X^{\mathrm{cl}})^2P^{\mathrm{cl}}) &:= W(f_{21}) = \frac{1}{3}(X^2P+XPX+PX^2).\end{aligned}$$

Example 3.10 (Some observables in $\mathrm{DiffPol}(\mathbb{R}^d)$). We use the Weyl quantization map to define some polynomial observables of interest. These observables get their names from their classical counterparts.

a) In Section 1, we repeatedly encountered the free Hamiltonian function $H_0^{\mathrm{cl}}(x,p) = \frac{1}{2m}(p_1^2+\ldots+p_d^2)$ (kinetic energy), where $m>0$ was the mass of the particle (cf. Def. 1.8). Under Weyl quantization, this is mapped to the *free (quantum) Hamiltonian* (*(quantum) operator of kinetic energy*), namely

$$H_0 := W(H_0^{\mathrm{cl}}) = \frac{1}{2m}P^2 = \frac{1}{2m}(P_1^2+\ldots+P_d^2) = -\frac{1}{2m}\Delta, \tag{3.30}$$

where $\Delta = \partial_1^2+\ldots+\partial_d^2$ is the *Laplace operator*.

b) In classical mechanics, a simple Hamiltonian with non-zero potential was the harmonic oscillator $H_{\mathrm{osc}}^{\mathrm{cl}}(x,p) = \frac{1}{2m}(p_1^2+\ldots+p_d^2)+\frac{k}{2}(x_1^2+\ldots+x_d^2)$, where again $m>0$ is the mass, and $k>0$ the strength of harmonic force, cf. Problem 1.2 for the one-dimensional case. Under Weyl quantization, this is mapped

[14] For $x,y\in\mathbb{R}^d$ and $\alpha\in\mathbb{N}_0^d$, we have $(x+y)^\alpha = \sum_{\mu\le\alpha}\binom{\alpha}{\mu}x^{\alpha-\mu}y^\mu$, with $\binom{\alpha}{\mu} = \prod_{j=1}^d\binom{\alpha_j}{\mu_j}$.

to the (*quantum*) *Hamiltonian of the harmonic oscillator*, namely

$$H_{\text{osc}} := W(H_{\text{osc}}^{\text{cl}}) = \frac{1}{2m}P^2 + \frac{k}{2}X^2, \tag{3.31}$$

with $X^2 := X_1^2 + \ldots + X_d^2$.

c) In classical mechanics in $d = 3$ dimensions, we studied the angular momentum operators (Def. 1.25). Under Weyl quantization, these are mapped to the (*quantum*) *angular momentum operators* $L_j := W(L_j^{\text{cl}})$, $j = 1, 2, 3$, which are explicitly given by

$$L_1 := X_2P_3 - X_3P_2, \quad L_2 := X_3P_1 - X_1P_3, \quad L_3 := X_1P_2 - X_2P_1. \tag{3.32}$$

Note that in the products appearing here, the X and P factors commute, so the symmetrization drops out.

All of these operators will be studied later.

Weyl quantization $W : \text{Pol}(\mathbb{R}^{2d}) \to \text{DiffPol}(\mathbb{R}^d)$ is a linear map. One can show that W is bijective and satisfies (Problem 3.2)

$$W(\mathbb{1}_{\text{Pol}(\mathbb{R}^{2d})}) = \mathbb{1}_{\text{DiffPol}(\mathbb{R}^d)}, \tag{3.33}$$

$$W(\overline{f}) = W(f)^*, \qquad f \in \text{Pol}(\mathbb{R}^{2d}). \tag{3.34}$$

Weyl quantization is *not* an algebra homomorphism. For example,

$$W(X^{\text{cl}}P^{\text{cl}}) = \frac{1}{2}(XP + PX) \neq XP = W(X^{\text{cl}})W(P^{\text{cl}}).$$

It does match Poisson brackets with commutators for *some* observables, for example,

$$W(\{X^{\text{cl}}, P^{\text{cl}}\}) = W(\mathbb{1}) = \mathbb{1} = -i[X, P] = -i[W(X^{\text{cl}}), W(P^{\text{cl}})],$$

but does *not* satisfy $W(\{f, g\}) = -i[W(f), W(g)]$ in general. Such a Lie algebra homomorphism property is indeed impossible in a very general sense, as will be analyzed in Problem 3.1.

Many other quantization schemes exist, some of which also preserve the *-operation as in (3.34) (for example Wick quantization). There is no quantization scheme that is fully "canonical". The nonetheless often used term *canonical quantization* refers to quantization of the canonical position/momentum observables of classical mechanics in terms of quantum-mechanical operators. We refer to [Wal07, Chapter 5] for a detailed discussion of quantization schemes.

By inspection of the formula (3.28), it is however clear that Weyl quantization can be generalized to more classical observables than just polynomials: One only has to make sure that the integrals (3.28) exist and define functions in $\mathcal{S}(\mathbb{R}^d)$, for instance phase space functions like $g(x)p^\alpha$, with $g \in C^\infty(\mathbb{R}^d)$ a smooth function such that $\psi \in \mathcal{S}(\mathbb{R}^d) \implies g\psi \in \mathcal{S}(\mathbb{R}^d)$ (such functions are called Schwartz multipliers, they have to satisfy polynomial bounds on all derivatives) pose no problem here.

The free Schrödinger equation

We pause the development of the theoretical framework for a moment to look at an example. Since we now have some quantum mechanical Hamiltonians at our disposal, we consider the simplest possibility, the free Hamiltonian

$$H_0 = \frac{1}{2m}P^2 = -\frac{1}{2m}\Delta, \tag{3.35}$$

where $m > 0$ is the mass parameter. More generally, we might consider Hamitlonians of the form $H = \frac{1}{2m}P^2 + V(X)$ with a potential function V of Schwartz multiplier form, in close analogy to classical mechanics (Def. 1.8), but for now we stick to H_0.

Since we have concentrated on vector states on DiffPol($\mathbb{R}^d$) for now, it is meaningful to consider Schrödinger's equation as in (2.65) rather than Heisenberg's Equation, namely

$$i\frac{d}{dt}\psi(t) = H_0\psi(t), \tag{3.36}$$

where $\psi(0) \in \mathcal{S}(\mathbb{R}^d)$, $\|\psi(0)\| = 1$, is given (initial condition). Because of the free Hamiltonian, this is called the *free Schrödinger Equation*. We are looking for a time-dependent function $(t, x) \mapsto \psi(t, x)$ satisfying the above equation. It is however not clear how we should interpret the derivative $\frac{d}{dt}$: Either we read it as a partial derivative ∂_t, or we treat $t \mapsto \psi(t) \in \mathcal{S}(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$ as a function taking values in a function space. We will show below that in this situation, both versions work.

To get an idea for the solution of (3.36), we first proceed somewhat formally and give proofs afterwards. We start from the free Schrödinger equation in the form

$$i\frac{\partial}{\partial t}\psi(t, x) = -\frac{1}{2m}\Delta_x\psi(t, x), \tag{3.37}$$

where $\Delta_x = \partial_{x_1}^2 + \ldots + \partial_{x_d}^2$ is the Laplace operator in the x-variables. We now carry out a Fourier transform in the x-variables and consider $\hat{\psi}(t, p) = (2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{-i\langle p,x\rangle}\psi(t, x)\,dx$.

The Laplace operator on the right hand side is transformed to multiplication by $\frac{\|p\|^2}{2m}$. On the left hand side, we *expect* the t-derivative to commute with the Fourier transform (no proof yet) because the Fourier integral is linear and independent of t. We therefore expect the free Schrödinger Equation to be equivalent to

$$\frac{\partial}{\partial t}\hat{\psi}(t, p) = -\frac{i\|p\|^2}{2m}\hat{\psi}(t, p), \tag{3.38}$$

which is easily seen to have the unique solution

$$\hat{\psi}(t, p) = e^{-\frac{i\|p\|^2 t}{2m}}\hat{\psi}(0, p). \tag{3.39}$$

We then expect that by inverse Fourier transform,

$$\psi(t, x) = (2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{i\langle p,x\rangle}e^{-\frac{i\|p\|^2 t}{2m}}\hat{\psi}(0, p)\,dp. \tag{3.40}$$

Proposition 3.11 (Solution of the free Schrödinger Equation). *Let $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$ and define for $t \in \mathbb{R}$*

$$\psi_t(x) := \psi(t,x) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{i\langle p,x\rangle} e^{-\frac{i\|p\|^2 t}{2m}} \tilde{\psi}_0(p)\, dp. \tag{3.41}$$

Then

a) $\psi_t \in \mathcal{S}(\mathbb{R}^d)$ with $\|\psi_t\| = \|\psi_0\|$ for all $t \in \mathbb{R}$.
b) $\psi(0,x) = \psi_0(x)$ (with ψ_0 the prescribed function)
c) $\psi \in C^\infty(\mathbb{R}^{d+1})$, and (3.37) holds.
d) $t \mapsto \psi_t \in L^2(\mathbb{R}^d)$ is differentiable, and (3.36) holds with $\frac{d}{dt}$ interpreted as the derivative of an L^2-valued function.[a]

[a] One can also show that $t \mapsto \psi_t \in \mathcal{S}(\mathbb{R}^d)$ is differentiable in the locally convex topology of $\mathcal{S}(\mathbb{R}^d)$.

Proof. a) With $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$, also $\tilde{\psi}_0 \in \mathcal{S}(\mathbb{R}^d)$ lies in Schwartz space by Fourier transform. The function $\tilde{\psi}_t : p \mapsto e^{-\frac{i\|p\|^2 t}{2m}} \tilde{\psi}_0(p)$ is clearly smooth. Its multiderivaties $\partial_p^\beta \tilde{\psi}_t$ are of the form $\sum_{\gamma \le \beta} f_\gamma(p) e^{-\frac{i\|p\|^2 t}{2m}} \partial_p^\gamma \tilde{\psi}_0(p)$, with $f_\gamma \in \mathrm{Pol}(\mathbb{R}^d)$ polynomials, because derivatives of the exponential function are of this form. Hence $\sup_{p\in\mathbb{R}^d} |p^\alpha \partial_p^\beta \tilde{\psi}_t(p)| < \infty$ for any two multiindices α, β, which shows $\tilde{\psi}_t \in \mathcal{S}(\mathbb{R}^d)$. By the invariance of Schwartz space under Fourier transform, $\psi_t \in \mathcal{S}(\mathbb{R}^d)$ follows.

The norm can be computed with the help of the Plancherel Theorem in momentum space:

$$\|\psi_t\|^2 = \|\tilde{\psi}_t\|^2 = \int_{\mathbb{R}^d} \left| e^{-\frac{i\|p\|^2 t}{2m}} \tilde{\psi}_0(p) \right|^2 dp = \int_{\mathbb{R}^d} |\tilde{\psi}_0(p)|^2\, dp = \|\tilde{\psi}_0\|^2 = \|\psi_0\|^2.$$

b) follows immediately by inserting $t = 0$ and using the Fourier inversion theorem.

c) In order to prove smoothness of ψ, we note that the integrand $(t,x) \mapsto e^{i\langle p,x\rangle} e^{-\frac{i\|p\|^2 t}{2m}} \tilde{\psi}_0(p)$ is smooth (clear), and for each mixed partial derivative $\partial_t^n \partial_x^\alpha$ of the integrand, there exists a (t,x)-independent integrable upper bound. This again follows by the polynomial factors produced by the derivatives which are integrable in p against the rapidly decreasing $\tilde{\psi}_0$.

We may therefore differentiate under the integral sign. In particular, the equation (3.37) holds because the integrand satisfies

$$\left(i\partial_t + \frac{1}{2m}\Delta_x \right) e^{i\langle p,x\rangle} e^{-\frac{i\|p\|^2 t}{2m}} = 0.$$

d) Let $t, \varepsilon \in \mathbb{R}$. We have to show that

$$\begin{aligned} \left\| \varepsilon^{-1}(\psi_{t+\varepsilon} - \psi_t) + iH_0\psi_t \right\|^2 &= \left\| \mathcal{F}\left(\varepsilon^{-1}(\psi_{t+\varepsilon} - \psi_t) + iH_0\psi_t \right) \right\|^2 \\ &= \int_{\mathbb{R}^d} \left| \frac{1}{\varepsilon} \left(e^{i\frac{\|p\|^2(t+\varepsilon)}{2m}} - e^{i\frac{\|p\|^2 t}{2m}} \right) + i\frac{\|p\|^2}{2m} \right|^2 |\tilde{\psi}_0(p)|^2\, dp \end{aligned}$$

converges to zero as $\varepsilon \to 0$. The term in $|\cdot|^2$ converges to zero for every $p \in \mathbb{R}^d$ because it is a difference quotient and $i\frac{\|p\|^2}{2m}$ is the derivative. Furthermore, it is bounded by a

t-independent polynomial, which is integrable against the rapidly decaying function $|\tilde{\psi}_0|^2$. Hence dominated convergence yields the claimed result. □

We note that the preservation of the normalization established in part a) implies that ψ_t defines vector states of DiffPol($\mathbb{R}^d$) for all $t \in \mathbb{R}^d$, in agreement of what we found in matrix mechanics.

Furthermore, the solution (3.41) for $t \neq 0$ can be expressed in the form

$$\psi(t,x) = \Big(\frac{m}{2\pi i t}\Big)^{d/2} \int_{\mathbb{R}^d} e^{i\frac{m\|x-y\|^2}{2t}} \psi_0(y)\, dy. \tag{3.42}$$

This follows from an explicit calculation (exercise). We note that as a consequence of this representation, we have

$$\sup_{x\in\mathbb{R}^d} |\psi(t,x)| \leq \Big(\frac{m}{2\pi t}\Big)^{d/2} \|\psi_0\|_{L^1}, \qquad t \neq 0. \tag{3.43}$$

In particular, $\sup_{x\in\mathbb{R}^d} |\psi(t,x)| \to 0$ for $t \to \pm\infty$. As the L^2-norm of the solution is constant, this means that it spreads in space, similar to what we had observed in classical mechanics (Example 1.32). Further investigations of spreading wave packets will be carried out in the exercises.

In classical and matrix mechanics, we observed spreading of wave packets in some states, and oscillatory or even stationary behaviour in other states (see Example 2.34). While the free dynamics considered above exhibits only spreading in all states, for different Hamiltonians all these phenomena also occur in quantum mechanics. We give an example of this below.

Example 3.12 (Dynamics of states for the harmonic oscillator). Consider the oscillator Hamiltonian $H = \frac{1}{2}(P^2 + X^2)$ (for simplicity, we have omitted the constants $m, k > 0$ and work in dimension $d = 1$), and $A := \frac{1}{\sqrt{2}}(X + iP)$, with adjoint $A^* = \frac{1}{\sqrt{2}}(X - iP)$. Then

$$A^*A = \frac{1}{2}(X - iP)(X + iP) = \frac{1}{2}(X^2 + P^2) + \frac{i}{2}[X,P] = H - \frac{1}{2}\mathbb{1},$$

which implies $H = A^*A + \frac{1}{2}\mathbb{1}$. The Schwartz function

$$\psi_0(x) := e^{-x^2/2} \tag{3.44}$$

satisfies $iP\psi_0 = \psi_0' = -X\psi_0$, hence $A\psi_0 = 0$ and $H\psi_0 = \frac{1}{2}\psi_0$.

Hence the Schrödinger Equation with initial condition ψ_0 is solved by

$$\psi_t = e^{-it/2}\psi_0, \tag{3.45}$$

because $i\partial_t\psi_t = \frac{1}{2}\psi_t = H\psi_t$.

In this case, $|\psi_t(x)| = |\psi_0(x)|$ is independent of time, in stark contrast to the spreading observed for the free Schrödinger Equation.

These observations can be explained in terms of the spectra of the free and oscillator Hamiltonians: In the oscillator example, H has eigenvalues, and ψ_0 was chosen to be an eigenvector, bnut H_0 does not have any eigenvalues. In the next section, we recall the necessary mathematics for that.

3.2 The general Hilbert space formulation of quantum mechanics

We now develop the general formulation of quantum mechanics. So far, we have operated with differential operators on Schwartz space $\mathcal{S}(\mathbb{R}^d)$. However, we have already seen at several stages that it might be more important to use the fact that $\mathcal{S}(\mathbb{R}^d)$ is included in the Hilbert space $L^2(\mathbb{R}^d)$: The normalization condition of states needs the L^2-norm (Proposition 3.6), the *-operation on $\mathrm{DiffPol}(\mathbb{R}^d)$ is connected to the L^2-scalar product (Proposition 3.5 b)), and eigenvalues of Hamiltonians lead to different dynamical behaviour than absence of eigenvalues (Example 3.12). As the natural framework for spectral theory is given by Hilbert spaces, this is another indication that we should consider a Hilbert space framework.

We will, however, often need to go beyond the technically simplest setting of *bounded* operators on Hilbert spaces, and rather consider unbounded (densely defined) operators. This is explained at the example of the Schrödinger position and momentum operators in the following example, which also serves to recall some notions of operator theory (see Appendix M.4).

Note that we need to pass from the maps $X_j, P_j : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ to closed and selfadjoint versions because the spectrum of a non-closed unbounded operator is ill-behaved (Definition M.49) and only selfadjoint operators admit a spectral resolution (Theorem M.56), and generators of strongly continuous unitary one-parameter groups (i.e., dynamics) are selfadjoint (Theorem M.59)

Example 3.13 (Operator properties of the position and momentum operators). We consider the position and momentum operators

$$X_j, P_j : \mathcal{D} := \mathcal{S}(\mathbb{R}^d) \subset L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$$

as operators in the Hilbert space $\mathcal{H} := L^2(\mathbb{R}^d)$.

- The operators X_j, P_j are not defined on the full Hilbert space $L^2(\mathbb{R}^d)$, and the naive definitions $(X_j\psi)(x) = x_j\psi(x)$ and $(P_j\psi)(x) = -i\partial_j\psi(x)$ do not work for general $\psi \in L^2(\mathbb{R}^d)$ because $x \mapsto x_j\psi(x)$ may fail to lie in $L^2(\mathbb{R}^d)$, and for general $\psi \in L^2(\mathbb{R}^d)$, partial derivatives do not exist.
  However, X_j, P_j are *densely defined*, meaning that their domain of definition

$$\operatorname{dom} X_j = \operatorname{dom} P_j = \mathcal{S}(\mathbb{R}^d) \tag{3.46}$$

  is a dense subspace of $L^2(\mathbb{R}^d)$. As a special feature of these operators, we note that they have an *invariant domain*, namely they map their domain $\mathcal{S}(\mathbb{R}^d)$ into itself.

- X_j and P_j are *symmetric*, meaning that

$$\langle \psi, X_j \varphi \rangle = \langle X_j \psi, \varphi \rangle, \qquad \langle \psi, P_j \varphi \rangle = \langle P_j \psi, \varphi \rangle, \qquad \psi, \varphi \in \mathcal{S}(\mathbb{R}^d). \quad (3.47)$$

  This was shown in Proposition 3.5 b).
- Both X_j and P_j are *unbounded*, meaning that

$$\|X_j\| := \sup_{\psi \in \mathcal{S}(\mathbb{R}^d) \setminus \{0\}} \frac{\|X_j \psi\|}{\|\psi\|} = +\infty, \quad (3.48)$$

  and similarly $\|P_j\| = \infty$. To see this for X_j, consider a normalized $\psi \in \mathcal{S}(\mathbb{R}^d)$ such that $x_j > n$ for all $x \in \operatorname{supp} \psi$ (for given $n \in \mathbb{N}$). Then

$$\|X_j \psi\|^2 = \int_{\mathbb{R}^d} x_j^2 |\psi(x)|^2 \, dx \geq n^2 \int_{\mathbb{R}^d} |\psi(x)|^2 \, dx = n^2,$$

  and therefore $\|X_j\| = \sup_{\xi \in \mathcal{S}(\mathbb{R}^d), \|\xi\|=1} \|X_j \xi\| \geq \sup_n n = \infty$. For P_j, we similarly obtain with the help of the Plancherel Theorem

$$\|P_j\| = \sup_{\xi \in \mathcal{S}(\mathbb{R}^d), \|\xi\|=1} \|P_j \xi\| = \sup_{\xi \in \mathcal{S}(\mathbb{R}^d), \|\xi\|=1} \|X_j \xi\| = \infty.$$

- As densely defined operators, X_j and P_j have *adjoints* X_j^*, P_j^*. These symbols now deviate from their previous meaning, where they denoted the *-operation in $\operatorname{DiffPol}(\mathbb{R}^d)$. For clarity, we will from now on denote that operation by $D \mapsto D^\dagger$. Recall that the domain of X_j^* is by definition

$$\operatorname{dom} X_j^* = \{\psi \in L^2(\mathbb{R}^d) \ : \ \exists c > 0 \ : \ |\langle \psi, X\varphi \rangle| \leq c \|\varphi\| \ \forall \varphi \in \mathcal{S}(\mathbb{R}^d)\},$$

  and X_j^* is then defined by (using Riesz' Lemma)

$$\langle X_j^* \psi, \varphi \rangle := \langle \psi, X_j \varphi \rangle, \qquad \psi \in \operatorname{dom} X_j^*, \ \varphi \in \operatorname{dom} X_j.$$

  By symmetry, we know $\operatorname{dom} X_j^* \supset \mathcal{S}(\mathbb{R}^d)$ and $X_j^*|_{\mathcal{S}(\mathbb{R}^d)} = X_j^\dagger$, i.e. $X_j^* \supset X_j$ is an extension of X_j.
- X_j and P_j are both *not selfadjoint*. Selfadjointness of X_j would mean $\operatorname{dom} X_j^* = \operatorname{dom} X_j$ and $X_j^* \psi = X_j \psi$ for all $\psi \in \operatorname{dom} X_j$. To see that, let $f \in L^2(\mathbb{R}^d)$ such that $g : x \mapsto x_j f(x)$ lies in $L^2(\mathbb{R}^d)$, for example $f(x) = (1 + \|x\|^2)^{-d^2}$ (this function does not lie in $\mathcal{S}(\mathbb{R}^d)$). Then, for any $\varphi \in \mathcal{S}(\mathbb{R}^d)$

$$|\langle f, X_j \varphi \rangle| = \left| \int_{\mathbb{R}^d} \overline{f(x)} x_j \varphi(x) \, dx \right| = |\langle g, \varphi \rangle| \leq \|g\| \cdot \|\varphi\|,$$

  where we have used the Cauchy-Schwarz Inequality. Hence $g \in \operatorname{dom} X_j^*$ but $g \notin \operatorname{dom} X_j$, which shows that X_j is not selfadjoint. The argument for P_j is similar.

- However, X_j (and P_j) is *essentially selfadjoint*. An operator A is called essentially selfadjoint if it is *closable* (symmetric operators are always closable), and its closure $\overline{A}$ is selfadjoint. This is the same as saying that A has a unique selfadjoint extension, and is equivalent to the ranges $\mathrm{Ran}(A\pm i) = \{(A\pm i)\psi \ : \ \psi \in \mathrm{dom}\, A\}$ being dense for both choices of the $\pm$ sign (see Appendix M.4).
  To check this criterion for $A = X_j$, let $\psi \perp \mathrm{Ran}(X_j \pm i)$, i.e.

$$0 = \int_{\mathbb{R}^d} \overline{\psi(x)}(x_j \pm i) f(x)\, dx \qquad \forall f \in \mathcal{S}(\mathbb{R}^d).$$

  As Schwartz space is dense in $L^2(\mathbb{R}^d)$, for any bounded Borel set $B \subset \mathbb{R}^d$ we can find a sequence $(g_n) \subset \mathcal{S}(\mathbb{R}^d)$ such that $g_n \to \chi_B$ in $L^2(\mathbb{R}^d)$, which implies $0 = \int_B \overline{\psi(x)}(x_j \pm i) f(x)$ for all $f \in \mathcal{S}(\mathbb{R}^d)$. Again by density of the Schwartz space this implies $(x_j \mp i)\psi(x) = 0$ for almost all $x \in B$, hence $\psi(x) = 0$ for almost all $x \in B$. As B was arbitrary, we conclude $\psi = 0$ in $L^2(\mathbb{R}^d)$. Thus the orthogonal complement of $\mathrm{Ran}(X_j \pm i)$ is $\{0\}$, which is equivalent to $\mathrm{Ran}(X_j \pm i) \subset L^2(\mathbb{R}^d)$ being dense. The argument for P_j is again similar.
- The *unique selfadjoint extension* $\overline{X_j}$ of X_j is

$$\mathrm{dom}\, \overline{X_j} = \left\{\psi \in L^2(\mathbb{R}^d) \ : \ \int x_j^2 |\psi(x)|^2 dx < \infty\right\},$$
$$(\overline{X_j}\psi)(x) = x_j \psi(x).$$

  Clearly the operator $(\overline{X}_j, \mathrm{dom}\, \overline{X}_j)$ defined above is an extension of X_j, and it is easy to check that it is symmetric. To show that it is selfadjoint, we use the criterion that symmetric A is selfadjoint if and only if $\mathrm{Ran}(A \pm i)$ is the whole Hilbert space (See Appendix M.4). Thus we have to show that given $f \in L^2(\mathbb{R}^d)$, there exists $g \in \mathrm{dom}\, \overline{X}_j$ such that $(\overline{X}_j \pm i)g = f$. Define $g(x) := (x \pm i)^{-1} f(x)$. Then $g \in L^2(\mathbb{R}^d)$, and

$$\int_{\mathbb{R}^d} x_j^2\, |g(x)|^2\, dx = \int_{\mathbb{R}^d} \frac{x_j^2}{1 + x_j^2}\, |f(x)|^2\, dx \le \int_{\mathbb{R}^d} |f(x)|^2\, dx < \infty,$$

  hence $g \in \mathrm{dom}\, \overline{X}_j$. Since $(\overline{X}_j \pm i)g = f$, the proof is finished.
  Similarly one can show that the unique selfadjoint extension of P_j is

$$\mathrm{dom}\, \overline{P_j} = \left\{\psi \in L^2(\mathbb{R}^d) \ : \ \int p_j^2 |\tilde{\psi}(p)|^2 dp < \infty\right\},$$
$$(\mathcal{F}\overline{P_j}\psi)(p) = p_j (\mathcal{F}\psi)(p).$$

- Let us determine the *spectrum* of $\overline{X}_j$. We first note that this map is always injective: If $f \in \mathrm{dom}\, \overline{X}_j$ lies in the kernel, namely $\overline{X}_j f = \lambda f$, then $(x_j - \lambda) f(x) = 0$ for almost all $x \in \mathbb{R}^d$, hence $f(x) = 0$ for almost all $x \in \mathbb{R}^d$. This means $f = 0$ in the sense of $L^2(\mathbb{R}^d)$. To investigate surjectivity, we let $g \in L^2(\mathbb{R}^d)$ be arbitrary

and search for $f \in \operatorname{dom} \overline{X}_j$ such that

$$(x_j - \lambda) f(x) = g(x) \Rightarrow f(x) = \frac{g(x)}{x_j - \lambda}.$$

For $\operatorname{Im} \lambda \neq 0$, this function f lies in $\operatorname{dom} \overline{X}_j$ because

$$\int_{\mathbb{R}^d} \frac{x_j^2 |g(x)|^2}{|x_j - \lambda|^2}\, dx = \int_{\mathbb{R}^d} \frac{x_j^2 |g(x)|^2}{(x_j - \operatorname{Re}\lambda)^2 + (\operatorname{Im}\lambda)^2}\, dx \leq (\operatorname{Im}\lambda)^{-2} \|\overline{X}_j g\|^2 < \infty,$$

which implies $\sigma(\overline{X}_j) \subset \mathbb{R}$ (this is true for any selfadjoint operator). For $\operatorname{Im}\lambda = 0$, however, the above integral converges only if g vanishes in a neighbourhood of $x_j = \lambda$. Thus not every $g \in L^2(\mathbb{R}^d)$ lies in the range of $\overline{X}_j - \lambda$, and we conclude $\sigma(\overline{X}_j) = \mathbb{R}$.

- As a selfadjoint operator, $\overline{X}_j$ is of the form $\overline{X}_j = \int \lambda\, dE^{\overline{X}_j}(\lambda)$ for a (uniquely determined) projection-valued measure $E := E^{\overline{X}_j}$. We claim that this measure is (for $B \subset \mathbb{R}$ a Borel set)

$$(E(B)\psi)(x) = \chi_B(x_j)\psi(x). \tag{3.49}$$

We first note that the operators $E(B)$ are orthogonal projections, with $E(\varnothing) = 0$, $E(\mathbb{R}) = 1$. For a countable family (B_n) of mutually disjoint Borel sets B_n, we set $B := \bigcup_{n=1}^{\infty} B_n$. Then we have $E(B_n)E(B_m) = \delta_{n,m}E(B_n)$ because the B_n are mutually disjoint, and $E(B)E(B_n) = E(B_n) = E(B_n)E(B)$. For any $\psi \in \mathcal{H}$, this gives

$$\begin{aligned}
\left\| \sum_{n=1}^{N} E(B_n)\psi - E(B)\psi \right\|^2 &= \|E(B)\psi\|^2 - \sum_{n=1}^{N} \|E(B_n)\psi\|^2 \\
&= \int_{B^j} |\psi(x)|^2\, dx - \sum_{n=1}^{N} \int_{B_n^j} |\psi(x)|^2\, dx \\
&= \int_{B^j \setminus \bigcup_{n=1}^{N} B_n^j} |\psi(x)|^2\, dx,
\end{aligned}$$

where we have set $B_n^j := \{x \in \mathbb{R}^d \,:\, x_j \in B_n\}$. As $N \to 0$, the above integral converges to 0 by dominated convergence. This shows that E is a projection-valued measure.
For $\psi \in L^2(\mathbb{R}^d)$ and $B \subset \mathbb{R}$ a Borel set, we have

$$\langle \psi, E(B)\psi \rangle = \int_{\mathbb{R}} \chi_B(x_j) |\psi(x)|^2 dx = \int_B \int_{\mathbb{R}^{d-1}} |\psi(x_j, x_j^{\perp})|^2\, dx_j^{\perp}\, dx_j,$$

where $x_j^\perp$ stands for the coordinates $x_1, \dots, x_d$ except x_j. This shows that the measure E_ψ has the density $\psi_j(x_j) := \int_{\mathbb{R}^{d-1}} |\psi(x_j, x_j^\perp)|^2 \, dx_j^\perp$ w.r.t. Lebesgue measure, i.e. $dE_\psi(x_j) = \psi_j(x_j)\, dx_j$.
We now want to understand the integral $A := \int_{\mathbb{R}} \lambda \, dE(\lambda)$. Its domain consists of all $\psi \in L^2(\mathbb{R}^d)$ for which $\int_{\mathbb{R}} \lambda^2 \, dE_\psi(\lambda) < \infty$ is finite. This integral is

$$\int_{\mathbb{R}} \lambda^2 \, dE_\psi(\lambda) = \int_{\mathbb{R}} \lambda^2 \int_{\mathbb{R}^{d-1}} |\psi(\lambda, x_j^\perp)|^2 \, dx_j^\perp d\lambda = \int_{\mathbb{R}^d} x_j^2 |\psi(x)|^2 \, dx,$$

exactly as in the domain of $\overline{X}_j$. Hence $\operatorname{dom} A = \operatorname{dom} \overline{X}_j$.
To show that the operators A and $\overline{X}_j$ act in the same way, we first note that for simple functions $f = \sum_{k=1}^n \alpha_k \chi_{B_k}$ (the $\alpha_k \in \mathbb{C}$ are coefficients, the $B_k \subset \mathbb{R}$ Borel sets), the operator integral $I_E(f) := \int_{\mathbb{R}} f(\lambda) dE(\lambda)$ acts according to

$$(I_E(f)\psi)(x) = f(x_j)\psi(x), \qquad \psi \in \operatorname{dom} A.$$

This follows directly from the definition of E. For a bounded measurable function f, we find simple functions f_n such that $f_n \to f$ pointwise and $\|f_n\|_\infty \leq \|f\|_\infty$. By dominated convergence, this gives $(I_E(f)\psi)(x) = f(x_j)\psi(x)$ also for bounded measurable f. In particular, it holds for $f(\lambda) = \lambda \chi_{[-R,R]}(\lambda)$. Another approximation ($R \to \infty$) shows that this formula also holds for unbounded measurable f, and in particular for $f(\lambda) = \lambda$. Thus

$$\left(\int_{\mathbb{R}} \lambda \, dE(\lambda)\psi \right)(x) = x_j \psi(x) = (\overline{X}_j \psi)(x) = \left(\int_{\mathbb{R}} \lambda \, dE^{\overline{X}_j}(\lambda)\psi \right)(x).$$

The uniqueness of the spectral measure now implies $E = E^{\overline{X}_j}$.

- Since the momentum operators P_j are unitarily equivalent to the position operators X_j, the preceding analysis also applies to the P_j. One finds $\sigma(\overline{P}_j) = \mathbb{R}$ and

$$(\mathcal{F} E^{\overline{P}_j}(B)\psi)(p) = \chi_B(p_j)(\mathcal{F}\psi)(p). \tag{3.50}$$

- Stone's Theorem tells us that any selfadjoint operator A defines a strongly continuous unitary one-parameter group $U(t) = e^{itA}$, $t \in \mathbb{R}$. For $A = \overline{X}_j$ and $A = \overline{P}_j$ this these unitaries can be written down explicitly.
We have

$$(e^{it\overline{X}_j}\psi)(x) = \left(\int_{\mathbb{R}} e^{it\lambda} \, dE^{\overline{X}_j}\psi \right)(x) = e^{itx_j}\psi(x) \tag{3.51}$$

and

$$(\mathcal{F}e^{it\overline{P}_j}\psi)(p) = \left(\mathcal{F}\int_{\mathbb{R}} e^{it\lambda}\, dE^{\overline{P}_j}\psi\right)(p) = e^{itp_j}(\mathcal{F}\psi)(p)$$
$$\Longrightarrow\ (e^{it\overline{P}_j}\psi)(x) = (2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{i\langle p,x\rangle}e^{itp_j}(\mathcal{F}\psi)(p)\,dp$$
$$= \psi(x + te_j),$$

namely the unitary one-parameter group generator by momentum $\overline{P}_j$ generates translations in the e_j-direction, as in classical mechanics.

We isolate a few important points from this discussion and Appendix M.4.

Theorem 3.14. *Let H be a selfadjoint operator on a Hilbert space.*

a) The spectrum $\sigma(H) \subset \mathbb{R}$ is real and non-empty. It contains the point spectrum $\sigma_p(H)$ consisting of all eigenvalues of H, which may be empty.

b) There exists a unique projection-valued measure $E_H : \mathfrak{B}(\mathbb{R}) \to \mathcal{P}(\mathcal{H})$ such that

$$H = \int_{\mathbb{R}} \lambda\, dE^H(\lambda), \qquad \mathrm{dom}(H) = \left\{\psi \in \mathcal{H} \ :\ \int_{\mathbb{R}} \lambda^2 dE^H_\psi(\lambda) < \infty\right\}.$$

The support of E^H is the spectrum of H, and the number-valued measure $E^H_\psi : B \mapsto \|E^H(B)\psi\|^2$ is a finite Borel measure with $E^H_\psi(\mathbb{R}) = \|\psi\|^2$. For normalized ψ, it is a probability measure.

c) For a measurable function $f : \sigma(H) \to \mathbb{C}$, define the operator

$$f(H) := \int_{\mathbb{R}} f(\lambda)\, dE^H(\lambda), \qquad \mathrm{dom}\, f(H) = \left\{\psi \in \mathcal{H} \ :\ \int_{\mathbb{R}} |f(\lambda)^2| dE^H_\psi(\lambda) < \infty\right\}.$$

Then $f(H)$ is bounded for (essentially) bounded f, and $f \mapsto f(H)$ behaves like a *-algebra homomorphism (see Prop. M.55). As a map from bounded measurable functions to bounded operators, $f \mapsto f(H)$ is a unital *-algebra homomorphism.

d) The spectrum of $f(H)$ is

$$\sigma(f(H)) = \mathrm{ess\ ran}_{E^H}(f), \tag{3.52}$$

this consists of all $\lambda \in \mathbb{C}$ such that for every open neighbourhood N of λ, one has $E^H(f^{-1}(N)) \neq 0$. For continuous f, one has $\sigma(f(H)) = \overline{f(\sigma(H))}$.

e) The unitaries $U(t) := e^{itH}$ (i.e. $e_t(H)$ with $e_t(\lambda) := e^{it\lambda}$) form a strongly continuous one-parameter group. One has

$$\mathrm{dom}\, H = \{\psi \in \mathcal{H} \ :\ t \mapsto U(t)\psi \text{ is differentiable}\}, \tag{3.53}$$
$$H\psi = -i\frac{d}{dt}U(t)\psi, \qquad \psi \in \mathrm{dom}\, H. \tag{3.54}$$

Having recalled these facts, we now discuss the general Hilbert space formulation of quantum mechanics. As in classical and matrix mechanics, we will need to identify the *observables* and *states*, together with an appropriate probabilistic interpretation (involving

probabilities, expectation values, uncertainties), discuss *dynamics* in a Heisenberg and Schrödinger picture, *symmetries* and *constants of motion*.

Observables and states in quantum mechanics

The first datum needed to describe a quantum mechanical system is a *complex Hilbert space* $\mathcal{H}$. This space may be infinite-dimensional, but is typically separable. The Hilbert space $\mathcal{H}$ will be considered fixed for now.

The **observables will be identified with the selfadjoint operators on** $\mathcal{H}$. It must be noted that in general, these are not elements of a unital *-algebra because products and linear combinations of densely defined operators may fail to be densely defined. For example, no function $\psi \in L^2(\mathbb{R})$ except $\psi = 0$ has the property that

$$x \mapsto e^{x^2}\psi(x), \qquad p \mapsto e^{p^2}\tilde{\psi}(p) \tag{3.55}$$

both lie in[15] $L^2(\mathbb{R})$. Hence operators defined (on suitable domains) by $(A\psi)(x) = e^{x^2}\psi(x)$ and $\widetilde{(B\psi)}(p) = e^{p^2}\tilde{\psi}(p)$ have $\operatorname{dom} A \cap \operatorname{dom} B = \{0\}$. Thus algebraic expressions like AB, $A + B$, $[A, B]$ only make sense on the trivial vector $\psi = 0$ and do not carry any information about A, B.

Nonetheless, the *-algebraic approach will turn out to provide a clear picture. We denote by $B(\mathcal{H})$ the set of all *bounded* operators A on $\mathcal{H}$ (defined on the full Hilbert space, $\operatorname{dom} A = \mathcal{H}$). This is a unital *-algebra (with $A \mapsto A^*$ denoting the adjoint), so our general setting of Section 2.1 applies directly. It is even a C^*-algebra, namely closed in the norm topology given by the operator norm, and the C^*-property $\|A^*A\| = \|A\|^2$ holds. Furthermore, it is a von Neumann algebra, i.e. $B(\mathcal{H})$ is also closed in the strong operator topology, weak operator topology, and ultraweak topology (see Appendix M.4).

In case $\mathcal{H}$ is finite-dimensional, we recover $B(\mathcal{H}) \cong M_{\dim \mathcal{H}}(\mathbb{C})$ the matrix algebras considered before.

Note that we have several ways of turning an unbounded selfadjoint operator $(\operatorname{dom} A, A)$ into bounded ones, and vice versa: For any bounded measurable function $f : \sigma(A) \to \mathbb{C}$, the operator $f(A)$ is bounded with $\|f(A)\| = \operatorname{ess\,sup}_{\lambda\in\sigma(A)}|f(\lambda)|$. In particular, for $f = \chi_{\mathbb{B}}$ with $\mathbb{B} \subset \mathbb{R}$ Borel, $\chi_{\mathbb{B}}(A) = E^A(\mathbb{B})$ is an orthogonal projection, and for $t \in \mathbb{R}$, the operator e^{itA} is unitary. Another often used construction is to consider

$$A_n := AE^A([-n, n]) = E^A([-n, n])A, \qquad \|A_n\| \le n,$$

which satisfies

$$A_n\psi = E^A([-n, n])A\psi \to A\psi, \qquad \psi \in \operatorname{dom} A, \tag{3.56}$$

because $E^A([-n, n]) \to \mathbb{1}$ in SOT (this follows from the σ-additivity of E^A, exercise).

We will therefore be able to use bounded operators to study unbounded ones. We begin with considering the *states* of $B(\mathcal{H})$. Here our general analysis of Section 2.1 applies directly. We recall from functional analysis that any state $\omega \in S(B(\mathcal{H}))$ (a positive normalized linear functional $\omega : B(\mathcal{H}) \to \mathbb{C}$) satisfies

$$|\omega(A)| \le \|A\|, \qquad A \in B(\mathcal{H}), \tag{3.57}$$

[15] We do not give a formal proof here, but note that this fact is a consequence of the uncertainty relation.

and of course the Cauchy-Schwarz Inequality $|\omega(A^*B)|^2 \leq \omega(A^*A)\omega(B^*B)$. In particular, states are continuous when $B(\mathcal{H})$ is given the norm topology. Many states of $B(\mathcal{H})$ exist, for example the *vector states*

$$\omega_\psi \in S(B(\mathcal{H})), \qquad \omega_\psi(A) := \langle \psi, A\psi \rangle, \quad \psi \in \mathcal{H},\ \|\psi\| = 1. \tag{3.58}$$

As state spaces are convex, also convex combinations of these states are states. In infinite dimension, we can also consider limits of convex combinations: For an orthonormal family (ψ_j) and parameters $0 \leq \lambda_j \leq 1$, $\sum_j \lambda_j = 1$, also

$$\omega(A) := \sum_j \lambda_j \omega_{\psi_j}(A), \qquad A \in B(\mathcal{H}), \tag{3.59}$$

is a state: It is well-defined and normalized because of the summability condition $\sum_j \lambda_j = 1$, and positivity is clear by positivity of the vector states ω_{ψ_j} and $\lambda_j \geq 0$. When we extend the family (ψ_j) to an orthonormal basis and consider the Hilbert space trace Tr (see Appendix M.4), this state takes the form

$$\omega(A) = \mathrm{Tr}(\rho A), \qquad \rho = \sum_j \lambda_j P_{\psi_j}. \tag{3.60}$$

These observations motivate the following definitions. For simplicity, we restrict ourselves to separable Hilbert spaces here, although the generalization to non-separable Hilbert spaces is not difficult.

Definition 3.15 (Density matrices). Let $\mathcal{H}$ be a separable Hilbert space. A *density matrix* is a positive operator $\rho \in B(\mathcal{H})$ that is of the form

$$\rho = \sum_n \lambda_n P_{e_n}, \tag{3.61}$$

where $(\lambda_n) \subset [0,1]$ is a summable sequence of real numbers with $\sum_n \lambda_n = 1$, and (e_n) an orthonormal basis of $\mathcal{H}$. Equivalently, a density matrix is a positive trace class operator with trace $\mathrm{Tr}\,\rho = 1$. The set of all density matrices on $\mathcal{H}$ is denoted $\mathcal{D}(\mathcal{H})$.

Vector states are clearly even WOT- and SOT-continuous. For more general states, this will not be the case, but it will turn out to be important to isolate the following continuity property.

Definition 3.16 (Normal states). A state $\omega \in B(\mathcal{H})$ is called *normal* if for every SOT-convergent bounded net[a] (A_j), i.e. $A_j\psi \to A\psi$ for all $\psi \in \mathcal{H}$, and $\sup_j \|A_j\| < \infty$, one has $\lim_j \omega(A_j) = \omega(A)$.

[a] For SOT-convergent *sequences*, boundedness is automatic by the uniform boundedness principle.

Example 3.17. Let (e_n) be an orthonormal basis of $\mathcal{H}$, and define $Q_N := \sum_{n=1}^N P_{e_n}$. Then $\|Q_N\| = 1$ for all N, and $Q_N \to \mathbb{1}$ in SOT, but not in operator norm.
Proof: The Q_N are orthogonal projections, hence they have norm 1. For SOT-convergence,

we expand $\psi \in \mathcal{H}$ into the ONB, $\psi = \sum_{n=1}^{\infty} \langle e_n, \psi \rangle e_n$ (this converges in $\mathcal{H}$). Then

$$\begin{aligned} \|Q_N \psi - \psi\|^2 &= \left\| \sum_{n=1}^{N} \langle e_n, \psi \rangle e_n - \sum_{n=1}^{\infty} \langle e_n, \psi \rangle e_n \right\|^2 \\ &= \sum_{n>N} |\langle e_n, \psi \rangle|^2 \to 0 \qquad \text{as } N \to \infty \end{aligned}$$

However, $(Q_N - 1)e_{N+1} = -e_{N+1}$ has norm 1. Thus $\|Q_N - \mathbb{1}\| = 1 \not\to 0$. □

Recall from functional analysis that an increasing net converges to its supremum in strong operator topology (Vigier's Theorem).

The following result is the infinite-dimensional analogue of Prop. 2.11.

Proposition 3.18 (The normal state space of $B(\mathcal{H})$). *There is a bijection between the set $S_{\text{no}}(\mathcal{H})$ of normal states of $B(\mathcal{H})$ and the set of density matrices on $\mathcal{H}$, given by*

$$\mathcal{D}(\mathcal{H}) \to S_{\text{no}}(\mathcal{H}), \qquad \rho \mapsto \omega_\rho, \qquad \omega_\rho(A) := \text{Tr}(\rho A), \quad A \in B(\mathcal{H}).$$

Here Tr *denotes the Hilbert space trace.*

Proof. Let $\rho \in \mathcal{D}(\mathcal{H})$ be a density matrix and define $\omega_\rho(A) := \text{Tr}(\rho A)$. We have already explained $\omega_\rho \in S(B(\mathcal{H}))$ before, and now show that ω_ρ is normal. Let $\rho = \sum_n \lambda_n P_{e_n}$ be the spectral decomposition of ρ, and let $A_j \nearrow A$ be an SOT-convergent bounded net in $B(\mathcal{H})$. Then $Q_j e_n \to Q e_n$ in Hilbert space norm. This implies $\langle e_n, Q_j e_n \rangle \to \langle e_n, Q e_n \rangle$ for every n. Hence

$$\omega_\rho(Q_j) = \sum_j \lambda_j \langle e_n, Q_j e_n \rangle \to \sum_j \lambda_j \langle e_n, Q e_n \rangle = \omega_\rho(Q)$$

by dominated convergence (here the boundedness of the net enters), which shows that ω_ρ is normal.

Conversely, let $\omega \in S_{\text{no}}(\mathcal{H})$ be normal. The map $b : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$, $b(\psi, \xi) := \omega(|\psi\rangle\langle\xi|)$ is a sesquilinear form which is bounded because $|b(\psi, \xi)| \leq \||\psi\rangle\langle\xi|\| = \|\psi\|\|\xi\|$. Hence, by Riesz' Lemma, there exists a bounded operator $\rho \in B(\mathcal{H})$ such that $b(\psi, \xi) = \langle \psi, \rho \xi \rangle$. This operator is positive because $\omega(|\psi\rangle\langle\psi|) \geq 0$.

Now consider an orthonormal basis (e_n) of $\mathcal{H}$. Then the projections P_{e_n} are mutually orthogonal, and therefore $\sum_{n=1}^{N} P_{e_n} \leq 1$, so this sequence converges to 1 in SOT and is bounded. By normality,

$$\sum_{n=1}^{N} \langle e_n, \rho e_n \rangle = \omega \left(\sum_{n=1}^{N} P_{e_n} \right) \to \omega(\mathbb{1}) = 1.$$

As $\rho \geq 0$, this implies that ρ is trace class with trace 1, i.e. $\rho \in \mathcal{D}(\mathcal{H})$.

It remains to show $\omega = \omega_\rho$. By construction of ρ, the two states ω and ω_ρ agree on all finite rank operators. So it suffices to show that a normal state is fixed by its restriction to finite rank operators. But this is clear from normality by finite-rank approximation.

This shows that ω is normal if and only if it is given by a density matrix, and it remains to show that this density matrix ρ is uniquely determined by ρ. To that end, assume $\omega_{\rho_1} = \omega_{\rho_2}$

for $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$, and set $R := \rho_1 - \rho_2$, a bounded selfadjoint operator. Then, for any $\psi \in \mathcal{H}$, $\|\psi\| = 1$,

$$0 = \mathrm{Tr}\big(RP_\psi\big) = \langle \psi, R\psi \rangle.$$

By the polarization identity, this implies $R = 0$, and finishes the proof. □

In Exercise 3.7, you will prove that a normal state $\omega \in S_{\mathrm{no}}(\mathcal{H})$ is pure if and only if it is a vector state.

Example 3.19 (Gibbs state for the harmonic oscillator). We consider the Hamiltonian $H = \frac{1}{2}(X^2 + P^2)$ of the harmonic oscillator in one dimension, which has spectrum $\sigma(H) = \mathbb{N}_0 + \frac{1}{2}$ with spectral measure

$$E^H(\mathbb{B}) = \sum_{n \in \mathbb{N}, n+\frac{1}{2} \in \mathbb{B}} P_{e_n}, \tag{3.62}$$

where (e_n) is an orthonormal basis consisting of eigenvectors of H, namely $He_n = (n + \frac{1}{2})e_n$ (see Exercise 3.6).

Given $\beta > 0$, we consider the operator $e^{-\beta H} = \sum_{n=0}^{\infty} e^{-\beta(n+\frac{1}{2})} P_{e_n}$. This operator is positive and trace class because its eigenvalues are summable, $\sum_{n=0}^{\infty} e^{-\beta(n+\frac{1}{2})} < \infty$. When we normalize it by this trace, we find that

$$\rho := \frac{e^{-\beta H}}{\mathrm{Tr}\big(e^{-\beta H}\big)} \tag{3.63}$$

is a density matrix.

This density matrix models a state at thermal equilibrium at temperature $1/\beta$, and is called *Gibbs state*.

Remark 3.20 (Non-normal states exist). One can show that a state is normal if and only if it is ultraweakly continuous. If $\mathcal{H}$ is infinite-dimensional, not all states on $B(\mathcal{H})$ are normal. Non-normal states can be specified in terms of quotient constructions or finitely additive measures. We will mostly be interested in normal states because of their probability interpretation (Corollary 3.21).

For observables modelled by unbounded selfadjoint operators $(\mathrm{dom}\, A, A)$, the expectation value $\omega(A)$ is initially not defined. In this case, the idea is to approximate by bounded operators, as we explain now. Recall the notion of k-th moment $\mathbb{M}_k(\nu)$ of a Borel measure ν on $\mathbb{R}$ (Def. M.4).

Corollary 3.21 (Normal states and probability measures). *Let $(A, \mathrm{dom}\, A)$ be a selfadjoint operator with spectral measure $E^A : \mathfrak{B}(\mathbb{R}) \to \mathcal{P}(\mathcal{H})$, and let $\omega \in S_{\mathrm{no}}(\mathcal{H})$ be normal.*

a) The map

$$\mu_\omega^A := \omega \circ E^A : \mathfrak{B}(\mathbb{R}) \to [0,1], \qquad B \mapsto \omega(E^A(B)), \tag{3.64}$$

is a Borel probability measure, $\mu_\omega^A \in \mathcal{P}(\mathbb{R})$, with support $\operatorname{supp}\mu_\omega^A \subset \sigma(A)$.

b) If $f : \sigma(A) \to \mathbb{C}$ is a bounded measurable function, then

$$\omega(f(A)) = \int_{\mathbb{R}} f(\lambda)\, d\mu_\omega^A(\lambda). \tag{3.65}$$

c) Let $k \in \mathbb{N}$. If the k-th moment $\mathbb{M}_k(\mu_\omega^A)$ of μ_ω^A exists, then

$$\lim_{n\to\infty} \omega(A_n^k) = \mathbb{M}_k(\mu_\omega^A) = \int_{\mathbb{R}} \lambda^k\, d\mu_\omega^A(\lambda), \tag{3.66}$$

where $A_n := E^A([-n,n])A \in B(\mathcal{H})$.

Proof. a) The map μ_ω^A is well-defined and maps into $[0,1]$ by positivity of ω: On the one hand, $\mu_\omega^A(B) = \omega(E^A(B)) \geq 0$ because $E^A(B) \geq 0$ (projection) and ω is positive, and on the other hand, $\mu_\omega^A(B) = \omega(\mathbb{1} - E^A(B^c)) = 1 - \mu^A(B^c) \leq 1$, where $B^c = \mathbb{R} \setminus B$ is the complement of B in $\mathbb{R}$.

By properties of projection-valued measures, $\mu_\omega^A(\varnothing) = 0$. For σ-additivity, let $(B_n) \subset \mathbb{R}$ be mutually disjoint. Then $E^A(\bigcup_{n=1}^N B_n) = \sum_{n=1}^N E^A(B_n) \nearrow \sum_{n=1}^\infty E^A(B_n)$ as $N \to \infty$, where the convergence is in SOT. As this sequence is bounded (projections), normality of ω implies

$$\mu_\omega^A\left(\bigcup_{n=1}^\infty B_n\right) = \lim_{N\to\infty} \mu_\omega^A\left(\bigcup_{n=1}^\infty B_n\right) = \sum_{n=1}^\infty \mu_\omega^A(B_n),$$

i.e. μ_ω^A is σ-additive. The support property is clear because E^A vanishes outside of $\sigma(A)$.

b) We first consider the case that $f = s = \sum_j c_j \chi_{I_j}$ is a simple function on $\sigma(A)$. Then $f(A) = \sum_j c_j E^A(I_j)$, and the claim follows by definition of μ_ω^A.

For any bounded Borel function $f : \sigma(A) \to \mathbb{C}$, there exists a sequence (s_N) of simple functions such that $s_N \to f$ pointwise and $\|s_N\|_\infty \leq C$ for all N. Then $s_N(A) \to f(A)$ in SOT, and since this sequence is bounded, normality implies $\omega(s_N(A)) \to \omega(f(A))$. On the other hand, by dominated convergence,

$$\left|\omega_{s_N(A)} - \int_{\mathbb{R}} f(\lambda)\, d\mu_\omega^A(\lambda)\right| \leq \int_{\mathbb{R}} |s_N(\lambda) - f(\lambda)|\, d\mu_\omega^A(\lambda) \to 0.$$

Thus $\omega(A) = \int_{\mathbb{R}} f(\lambda)\, d\mu_\omega^A(\lambda)$.

c) By the previous part, we have $\omega(A_n^k) = \int_{\mathbb{R}} \lambda^k \chi_{[-n,n]}(\lambda)\, d\mu_\omega^A(\lambda)$, and dominated convergence yields the result. □

In view of this corollary, we make the following definitions.

Definition 3.22 (Probability measures, expectation values, and uncertainties for unbounded observables). Let $(A, \operatorname{dom} A)$ be a selfadjoint operator on $\mathcal{H}$ and $\omega \in S_{\mathrm{no}}(\mathcal{H})$ a normal state, defining the measure $\mu_\omega^A \in \mathcal{P}(\mathbb{R})$.

a) For a Borel set $\mathbb{B} \subset \mathbb{R}$, the probability that the value of A lies in $\mathbb{B}$ is

$$\Pr_\omega(A \in \mathbb{B}) = \mu_\omega^A(\mathbb{B}) = \omega(E^A(\mathbb{B})) = \int_{\mathbb{B}} d\mu_\omega^A. \tag{3.67}$$

b) If the first moment of μ_ω^A exists, the *expectation value of A in ω* is

$$\omega(A) := \int_{\mathbb{R}} \lambda \, d\mu_\omega^A(\lambda). \tag{3.68}$$

c) If the second moment of μ_ω^A exists, the *uncertainty value of A in ω* is

$$\Delta_\omega(A)^2 := \int_{\mathbb{R}} (\lambda - \omega(A))^2 \, d\mu_\omega^A(\lambda). \tag{3.69}$$

Example 3.23.

a) Let $(A, \operatorname{dom} A)$ be a bounded operator and $\psi \in \operatorname{dom} A$ a normalized vector. Then the first and second moments of $\mu_\psi^A := \langle \psi, E^A(\cdot)\psi \rangle$ exist, and

$$\omega_\psi(A) = \langle \psi, A\psi \rangle, \qquad \Delta_\psi(A)^2 = \|A\psi\|^2 - \omega_\psi(A)^2.$$

This will be shown in the exercises.

b) On $\mathcal{H} = L^2(\mathbb{R})$, let $\psi = 2^{-1/2}\chi_{[-1,1]}$. Then

$$\tilde{\psi}(p) = \frac{1}{2\sqrt{\pi}} \int_{-1}^{1} e^{-ipx} \, dx = \frac{1}{\sqrt{\pi}} \frac{\sin p}{p},$$

and the probability measure of the momentum operator P in the vector state ω_ψ is given by

$$d\mu_\psi^P(p) = |\tilde{\psi}(p)|^2 \, dp = \frac{1}{\pi} \frac{\sin^2 p}{p^2} dp.$$

The first moment of this measure does not exist, i.e. $\mathbb{M}_1(\mu_\psi^P) = \infty$, because $\int_{\mathbb{R}} \frac{|p| \sin^2 p}{p^2} dp = \infty$. Hence in this case, the momentum operator P has no well-defined expectation value or uncertainty, but probabilities like for example $\Pr_\psi(P \in \mathbb{R}_+) = \frac{1}{2}$ are well-defined.

We are now in a situation somewhat analogous to that of classical mechanics. There we modelled our states by probability measures on phase space, and only for compactly supported measures all observables had finite expectation values. Similarly, in quantum mechanics, some observables will have undefined expectation values depending on the state. However, we no longer have "finite states" which allow for finite expectation values of *all* observables (all selfadjoint operators) if $\dim \mathcal{H} = \infty$.

Nonetheless, the probability measures μ_ω^A are well defined for any selfadjoint A and any normal state ω, because their definition only involves bounded operators (spectral projections).

Dynamics and symmetries in quantum mechanics

To prepare our discussion of dynamics, we have a brief look at the automorphisms of $B(\mathcal{H})$. The following proposition is exactly analogous to Prop. 2.24 in matrix mechanics and states $\operatorname{Aut} B(\mathcal{H}) \cong \mathcal{U}(\mathcal{H})/\mathbb{T}$. As the proof is basically the same, we do not repeat it here.

Proposition 3.24 (Automorphisms of $B(\mathcal{H})$).
a) Let $U \in \mathcal{U}(\mathcal{H})$ be a unitary. Then

$$\alpha_U : B(\mathcal{H}) \to B(\mathcal{H}), \qquad \alpha_U(A) := UAU^*, \tag{3.70}$$

is a unital *-automorphism of $B(\mathcal{H})$.
b) Let $U, V \in \mathcal{U}(\mathcal{H})$. Then $\alpha_U = \alpha_V \iff U = zV$ for some $z \in \mathbb{C}$, $|z| = 1$.
c) Let $\alpha \in \operatorname{Aut} B(\mathcal{H})$. Then $\alpha = \alpha_U$ for some $U \in U(N)$.

As in matrix mechanics, we consider unitary one-parameter groups $(U(t))_{t\in\mathbb{R}}$ and the automorphisms $\alpha_t := \alpha_{U(t)}$ that they define as a model for time evolution (dynamics). Already in matrix mechanics, it was important to consider *continuous* one-parameter groups. The same applies in quantum mechanics, but we now have to more carefully specify which topology on $\mathcal{U}(\mathcal{H})$ (WOT, SOT, norm topology ...) continuity should refer to.

We want that expectation values depend continuously on the time parameter t, i.e. the functions $t \mapsto \omega(\alpha_t^H(A))$, $A \in B(\mathcal{H})$, $\omega \in S_{\text{no}}(\mathcal{H})$, should be continuous. For strongly continuous $U : \mathbb{R} \to \mathcal{U}(\mathcal{H})$, the expectation values $t \mapsto \langle\psi, U(t)AU(-t)\psi\rangle$ in vector states (for bounded A) are continuous. Hence, given a density matrix $\rho = \sum_n \lambda_n P_{e_n}$, also

$$t \mapsto \omega_\rho(\alpha_t^H(A)) = \sum_n \lambda_n \langle U(-t)e_n, A\, U(-t)e_n\rangle \tag{3.71}$$

is continuous (dominated convergence). Thus strongly continuous unitary one-parameter groups[16] match our demand of continuous expectation values. In particular, it is not necessary that $t \mapsto U(t)$ is norm continuous.

We therefore identify *dynamics* with strongly continuous unitary one-parameter groups $(U(t))_{t\in\mathbb{R}}$ of automorphisms. By Stone's Theorem, we know that any such dynamics defines a unique selfadjoint operator H, its *Hamiltonian*, and any selfadjoint operator H defines a dynamics via $U(t) = e^{itH}$. We will sometimes write U_H or α_t^H to indicate the Hamiltonian.

Definition 3.25 (Quantum mechanical system).
a) A *quantum mechanical system* is a pair consisting of a complex Hilbert space $\mathcal{H}$ and a selfadjoint operator $(\operatorname{dom} H, H)$ on $\mathcal{H}$, called the *Hamiltonian* of the system.
b) The strongly continuous unitary one-parameter group $U(t) = e^{itH}$, $t \in \mathbb{R}$, is called the *dynamics* of the system.
c) The maps, $t \in \mathbb{R}$,

$$\alpha_t^H : B(\mathcal{H}) \to B(\mathcal{H}), \qquad A \mapsto U(t)AU(-t), \tag{3.72}$$

$$(\alpha_t^H)^* : S_{\text{no}}(\mathcal{H}) \to S_{\text{no}}(\mathcal{H}), \qquad \omega \mapsto \omega \circ \alpha_t^H \tag{3.73}$$

[16] For a unitary one-parameter group, the notions of SOT and WOT continuity coincide.

are the *time evolution* in the Heisenberg and Schrödinger picture, respectively.

Note that the Schrödinger time evolution $(\alpha_t^H)^*$ is well-defined: If $\omega = \omega_\rho$ is normal, then the states

$$(\alpha_t^H)^*(\omega_\rho) = \omega_{\alpha_{-t}^H(\rho)} \tag{3.74}$$

are also normal, for every $t \in \mathbb{R}$. As in matrix mechanics, we have for any $A \in B(\mathcal{H})$

$$\omega(\alpha_t^H(A)) = (\alpha_t^H)^*(\omega)(A), \tag{3.75}$$

i.e. the Heisenberg and Schrödinger pictures are equivalent.

The Heisenberg dynamics α^H naturally extends to unbounded selfadjoint operators: If $(\operatorname{dom} A, A)$ is selfadjoint, also $(U_H(t)\operatorname{dom} A, U_H(t)AU_H(-t))$ is; this latter operator will also be denoted $\alpha_t^H(A)$.

As for unbounded operators A continuity and differentiability properties of $t \mapsto \alpha_t^H(A)$ are more complicated, we study time evolution in the Schrödinger picture. It is sufficient to consider the case of vector states here, which is governed by Schrödinger's Equation.

Definition 3.26 (Schrödinger Equation). Let $(\mathcal{H}, H)$ be a quantum mechanical system. The (*general*) *Schrödinger Equation* is the equation

$$i\frac{d}{dt}\psi(t) = H\psi(t) \tag{3.76}$$

for a differentiable[a] function $\psi : \mathbb{R} \to \mathcal{H}$ such that $\psi(t) \in \operatorname{dom} H$ for all $t \in \mathbb{R}$.

It is often supplemented by an *initial condition*

$$\psi(0) = \psi_0, \tag{3.77}$$

where $\psi_0 \in \operatorname{dom} H$ is a prescribed vector.

[a] differentiable is meant in the norm topology of $\mathcal{H}$, i.e. $\|\varepsilon^{-1}(\psi(t+\varepsilon) - \psi(t)) - i\frac{d}{dt}\psi(t)\| \to 0$ as $\varepsilon \to 0$.

Proposition 3.27. *Let $(\mathcal{H}, H)$ be a quantum mechanical system. Then for any $\psi_0 \in \operatorname{dom} H$, the initial value problem for the Schrödinger Equation has a unique solution, namely*

$$\psi(t) = e^{-itH}\psi_0. \tag{3.78}$$

Proof. The function $\psi(t) := U(t)\psi_0$, $U(t) := e^{-itH}$, satisfies the initial condition $\psi(0) = \psi_0$, and is differentiable at $t = 0$ by Stone's Theorem (because $\psi_0 \in \operatorname{dom} H$), with derivative $\frac{d}{dt}\big|_{t=0}\psi(t) = -iH\psi_0$.

For general $t \in \mathbb{R}$, the vector $U(t)\psi_0$ lies in $\operatorname{dom} H$, i.e. $\psi_0 \in \operatorname{dom}(HU(t))$, because

$$\int_{\mathbb{R}} |\lambda e^{i\lambda t}|^2 \, dE_{\psi_0}^H(\lambda) = \int_{\mathbb{R}} \lambda^2 \, dE_{\psi_0}^H(\lambda) = \|H\psi_0\|^2 < \infty.$$

Furthermore, H and $U(t)$ commute because they are functions of the same selfadjoint operator H. Hence we get for general $t \in \mathbb{R}$

$$\left\|\varepsilon^{-1}(U(t+\varepsilon)-U(t))\psi_0 + iHU(t)\psi_0\right\| = \left\|\varepsilon^{-1}(U(\varepsilon)\psi_0 - \psi_0) + iH\psi_0\right\| \to 0 \quad \varepsilon \to 0$$

which shows that $t \mapsto U(t)\psi$ solves Schrödinger's Equation.

It remains to show the uniqueness of the solution. Let $t \mapsto \phi(t) \in \operatorname{dom} H$ be another solution of the Schrödinger Equation, with the same initial condition $\phi(0) = \psi_0$. Then $\xi(t) := \psi(t) - \phi(t)$ also solves the Schrödinger Equation by linearity, with the initial condition $\xi(0) = 0$.

The function $f : t \mapsto \|\xi(t)\|^2 = \langle \xi(t), \xi(t)\rangle$ is differentiable, and has derivative $f'(t) = \langle \xi'(t), \xi(t)\rangle + \langle \xi(t), \xi'(t)\rangle$ (exercise). Hence

$$f'(t) = \langle -iH\xi(t), \xi(t)\rangle + \langle \xi(t), -iH\xi(t)\rangle = i\left(\langle H\xi(t), \xi(t)\rangle - \langle \xi(t), H\xi(t)\rangle\right) = 0,$$

where in the last step, we have used that H is symmetric (even selfadjoint) and $\xi(t) \in \operatorname{dom} H$. So f is constant, which because of the initial condition implies $\xi(t) = 0$ for all t. □

The solution of the Schrödinger Equation might appear to be almost trivial, as it is always given by $t \mapsto e^{-itH}\psi_0$. Note, however, the rather involved definition of the exponential function via spectral calclulus. Consider, for example, a Hamiltonian of the form (on $\mathcal{H} = L^2(\mathbb{R}^d)$)

$$(H\psi)(x) = -\frac{1}{2m}(\Delta\psi)(x) + V(x)\psi(x), \tag{3.79}$$

where Δ is the Laplace operator and $V : \mathbb{R}^d \to \mathbb{R}$ a suitable potential function. Suppose that we have verified that it is essentially selfadjoint on some convenient domain like for example $\operatorname{dom} H = \mathcal{S}(\mathbb{R}^d)$. Then we know on abstract grounds that the solution of the Schrödinger Equation with initial condition ψ_0 is $t \mapsto e^{-itH}\psi_0$. This does however not directly tell us what, for example, the expectation value of the position operator X_1 at time $t = 5$ is, for which we would need to calculate $\langle U(5)\psi_0, X_1 U(5)\psi_0\rangle$. This is not an easy matter in general.

Later, we will analyze such questions in concrete examples of interest, and also study how spectral properties of H govern qualititative properties of the solution.

As in matrix mechanics, we proceed with a discussion of constants of motion and stationary states (Def. 2.31).

Definition 3.28 (Constants of motion and stationary states in quantum mechanics). Let $(\mathcal{H}, H)$ be a quantum mechanical system.

a) A *constant of motion* is a selfadjoint operator $(\operatorname{dom} A, A)$ such that for all $t \in \mathbb{R}$

$$U_H(t)\operatorname{dom} A = \operatorname{dom} A, \qquad U_H(t)A\psi = AU_H(t)\psi, \qquad \psi \in \operatorname{dom} A. \tag{3.80}$$

b) A normal state $\omega \in S_{\mathrm{no}}(\mathcal{H})$ is called *stationary* if

$$\omega \circ \alpha_t^H = \omega, \qquad t \in \mathbb{R}. \tag{3.81}$$

Based on our experience in matrix mechanics, we expect that constants of motion A are characterized by the condition that A commutes with H (Prop. 2.32). However, a condition of the form "$AH = HA$" is problematic. The right hand side of this equation is only defined on those $\psi \in \operatorname{dom} A$ which satisfy $A\psi \in \operatorname{dom} H$, and the left hand side is only defined on those $\psi \in \operatorname{dom} H$ which satisfy $H\psi \in \operatorname{dom} A$. These domains can be different, small, or even $\{0\}$, in which case the condition becomes meaningless. Hence one needs another notion of commutation for unbounded operators.

Definition 3.29 (Commutation of unbounded operators). Two selfadjoint operators $(\operatorname{dom} A, A)$ and $(\operatorname{dom} H, H)$ are said to *commute* if and only if their spectral measures commute, i.e.

$$[E^A(\mathbb{B}_1), E^H(\mathbb{B}_2)] = 0 \tag{3.82}$$

for all Borel sets $\mathbb{B}_1, \mathbb{B}_2 \in \mathfrak{B}(\mathbb{R})$. This is equivalent to[a]

$$[e^{isA}, e^{itH}] = 0, \qquad t, s \in \mathbb{R}. \tag{3.83}$$

[a] see functional analysis 2.

Proposition 3.30 (Constants of motion, stationary states, and strong commutation). *Let $(\mathcal{H}, H)$ be a quantum mechanical system.*

a) A selfadjoint operator $(\operatorname{dom} A, A)$ is a constant of motion if and only if A and H commute strongly.

b) Suppose $\bar{H}$ is the selfadjoint closure of a symmetric operator $(\operatorname{dom} H, H)$ and

$$e^{isA} \operatorname{dom} H \subset \operatorname{dom} H, \qquad He^{isA}\psi = e^{isA}H\psi, \quad \psi \in \operatorname{dom} H, \tag{3.84}$$

holds for all $s \in \mathbb{R}$. Then A and $\bar{H}$ commute strongly, and A is a constant of motion for the dynamics defined by $\bar{H}$.

c) A normal state $\omega = \omega_\rho$, $\rho \in \mathcal{D}(\mathcal{H})$, is stationary if and only if $U_H(t)\rho U_H(t)^* = \rho$ for all $t \in \mathbb{R}$. In the case of a vector state, i.e. $\rho = P_\psi$ for some normalized vector ψ, the state is stationary if and only if ψ is an eigenvector of H.

Proof. a) Assume that A and H commute strongly. Then $E^A(\mathbb{B})$, $\mathbb{B} \subset \mathbb{R}$ Borel, commutes with any bounded function of H, and in particular with $U_H(t)$, $t \in \mathbb{R}$. As a vector $\psi \in \mathcal{H}$ lies in $\operatorname{dom} A$ if and only if $\int_{\mathbb{R}} \lambda^2 \, d\mu^A_{U_H(t)\psi}(\lambda) < \infty$, with the measure $\mu^A_{U(t)\psi}(\mathbb{B}) = \langle U_H(t)\psi, E^A(\mathbb{B})U_H(t)\psi\rangle = \langle \psi, E^A(\mathbb{B})\psi\rangle$, we see $U_H(t) \operatorname{dom} A = \operatorname{dom} A$. Moreover, using the spectral truncations $A_n := A\chi_{[-n,n]}(A)$, we get

$$A_n U_H(t)\psi = U_H(t)A_n\psi.$$

As $n \to \infty$, we have $A_n\xi \to A\xi$ for all $\xi \in \operatorname{dom} A$. Thus, with $\psi \in \operatorname{dom} A$, and therefore also $U_H(t)\psi \in \operatorname{dom} A$, we obtain $AU_H(t)\psi = U_H(t)A\psi$. This shows that strong commutation of A and H implies that A is a constant of motion.

Conversely, assume that A is a constant of motion. Then $A = U_H(t)AU_H(t)^*$ as selfadjoint operators (including equality of domains). The spectral measure of $U_H(t)AU_H(t)^*$

is $E^{U_H(t)AU_H(t)^*}(\mathbb{B}) = U_H(t)E^A(\mathbb{B})U_H(t)^*$ (this follows from uniqueness of spectral measures), hence $U_H(t)$ commutes with the spectral measure of A. This implies that the two unitary groups $U_H(t)$ and e^{isA} commute (for all $t, s \in \mathbb{R}$), and therefore A and H commute strongly.

b) For $\psi \in \operatorname{dom} H$, we have $e^{-isA}\psi \in \operatorname{dom} H$, and $e^{isA}He^{-isA}\psi = H\psi$, i.e. $e^{isA}He^{-isA} = H$ as an operator equality including domains. The closure $\bar{H}$ of H is given by the closure $\overline{\Gamma(H)} = \Gamma(\bar{H})$ of the graph of H. This implies $\bar{H} = \overline{e^{isA}He^{-isA}} = e^{isA}\bar{H}e^{-isA}$.

c) ω_ρ is stationary if and only if $\omega_\rho = \omega_\rho \circ \alpha_t^H = \omega_{\alpha_{-t}^H(\rho)}$, which is equivalent to $\rho = U_H(-t)\rho U_H(t)$ because the density matrix of a normal state is unique.

If $\rho = P_\psi$ is a rank one projection, we have $[U_H(t), P_\psi] = 0$ if and only if $U_H(t)\psi = c(t)\psi$ for some function $c : \mathbb{R} \to \mathbb{C}$. By the group property and strong continuity of U_H, we see that c is continuous and satisfies $c(t)c(s) = c(t+s)$, $c(0) = 1$. This implies $c(t) = e^{i\lambda t}$ for some $\lambda \in \mathbb{R}$. Hence $t \mapsto U_H(t)\psi$ is differentiable, so $\psi \in \operatorname{dom} H$, and taking the derivative shows $H\psi = \lambda\psi$. The implication from $\psi \in \operatorname{dom} H$ being an eigenvector of H to $[U_H(t), P_\psi] = 0$ is straightforward. □

While the first item gives a characterization of constants of motion, it is typically very difficult to check because one hardly ever knows the domain or spectral measure of a selfadjoint operator explicitly. Let us give an example of application of the second item. In this example, we consider so-called *Schrödinger operators*, namely Hamiltonians of the form

$$H = \frac{1}{2m}P^2 + V(X), \qquad (H\psi)(x) = -\frac{1}{2m}(\Delta\psi)(x) + V(x)\psi(x), \tag{3.85}$$

where $V : \mathbb{R}^d \to \mathbb{R}$ is a suitable potential function and the domain of H is not yet specified. These correspond precisely to classical Hamiltonians of the form "kinetic plus potential energy" as studied in classical mechanics.

There are various conditions that guarantee essential selfadjointness of H. For example, in Problem 3.4 you have shown that when V is continuous and bounded and the domain is $\operatorname{dom} H = \mathcal{S}(\mathbb{R}^d)$, essential selfadjointness holds, but there exist various other criteria with weaker assumptions.

Example 3.31 (Translation and rotation symmetry). Let $\mathcal{H} = L^2(\mathbb{R}^d)$ and consider a Schrödinger operator of the form (3.85), defined on Schwartz space $\operatorname{dom} H = \mathcal{S}(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$. Assume that H is essentially selfadjoint.

a) Assume that $V(x) = V(x + s\,e_k)$ for some $k \in \{1, \dots, d\}$, and all $s \in \mathbb{R}$. That is, V is constant in the direction of e_k. Then the (selfadjoint) k-th momentum operator P_k is a constant of motion.

b) Assume $d = 3$ and that $V(x) = V(Rx)$ for all $R \in \mathrm{SO}(3)$. Then the (selfadjoint) angular momenum operators L_1, L_2, L_3 are constants of motion.

Proof. a) We claim that the unitary one-parameter group generated by P_k acts by translations in e_k-direction, namely $(e^{isP_k}\psi)(x) = \psi(x + se_k)$, $\psi \in L^2(\mathbb{R}^d)$. To prove this, let $\psi \in \mathcal{S}(\mathbb{R}^d)$. Then

$$\begin{aligned} &(\mathcal{F}e^{isP_k}\psi)(p) = e^{isp_k}\tilde{\psi}(p), \\ \Longrightarrow\ &(e^{isP_k}\psi)(x) = (2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{i\langle p,x\rangle}e^{i\langle p,se_k\rangle}\tilde{\psi}(p)\,dp = \psi(x + se_k). \end{aligned}$$

This shows that the two unitaries e^{isP_k} and translation by se_k agree on the dense subspace $\mathcal{S}(\mathbb{R}^d)$, and hence on all of $L^2(\mathbb{R}^d)$.

We now use Proposition 3.30 b): The domain $\mathcal{S}(\mathbb{R}^d)$ is invariant under translations in the e_k-direction, i.e. $e^{isP_k}\,\mathcal{S}(\mathbb{R}^d) = \mathcal{S}(\mathbb{R}^d)$. Furthermore, for $\psi \in \mathcal{S}(\mathbb{R}^d)$,

$$(He^{isP_k}\psi)(x) = -\frac{1}{2m}(\partial_1^2 + \ldots + \partial_d^2)[x \mapsto \psi(x + se_k)] + V(x)\psi(x + se_k) = (H\psi)(x + se_k),$$

where we have used the translational invariance of V. Thus the conditions of Proposition 3.30 b) are satisfied, and the claim follows.

b) We fix $k \in \{1, 2, 3\}$, denote by $R_k(s) \in \mathrm{SO}(3)$ the rotation around the e_k-axis with angle t, e.g.

$$R_1(s) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos s & \sin s \\ 0 & -\sin s & \cos s \end{pmatrix}, \tag{3.86}$$

and consider the operator

$$(U(s)\psi)(x) = \psi(R_1(-s)x), \qquad \psi \in L^2(\mathbb{R}^3). \tag{3.87}$$

As rotations preserve Lebesgue measure, this is a well-defined unitary on $L^2(\mathbb{R}^3)$, and one directly sees $U(s)U(s') = U(s + s')$, i.e. U is a unitary one-parameter group. We claim that it is strongly continuous. To that end, let $\psi \in C_c(\mathbb{R}^3)$ (continuous functions of compact support). We have

$$\|U(s + \varepsilon)\psi - U(s)\psi\|^2 = \|U(\varepsilon)\psi - \psi\|^2 = \int_{\mathbb{R}^3} |\psi(R_1(-\varepsilon)x) - \psi(x)|^2\, dx.$$

As $\varepsilon \to 0$, the integrand converges to zero. Furthermore, thanks to the compact support of ψ, the supports of all the functions $U(\varepsilon)\psi$, $-1 < \varepsilon < 1$, are contained in a compact set. Hence we find an integrable majorant and may use dominated convergence to arrive at $\|U(s + \varepsilon)\psi - U(s)\psi\| \to 0$. As $C_c(\mathbb{R}^3) \subset L^2(\mathbb{R}^3)$ is dense, this result extends to a proof of the strong continuity of U.

By Stone's Theorem, U has a selfadjoint generator. In the exercises, you will prove that this is the first angular momentum operator L_1, and that L_1 is the closure of its restriction to $\mathcal{S}(\mathbb{R}^3)$.

Now we may argue as in the first part: $\mathcal{S}(\mathbb{R}^3)$ is invariant under rotations. Furthermore, on $\psi \in \mathcal{S}(\mathbb{R}^3)$, the Hamiltonian commutes with rotations: The kinetic term (Laplace operator) multiplies with $\frac{1}{2m}\|p\|^2$ in momentum space, which is clearly rotationally invariant, and the potential is rotationally invariant by assumption. This concludes the proof that the angular momentum operators L_k are constants of motion in this example. □

We have seen above that stationary vector states are given by eigenvectors of the Hamiltonian (which might or might not exist depending on H). A few more details on this are collected in the following lemma.

Example 3.32. Let $(\mathcal{H}, H)$ be a quantum mechanical system, and $\psi \in \mathrm{dom}\, H$ a normalized eigenvector of H, with eigenvalue $\lambda_0 \in \mathbb{R}$. Then

$$\mu_\psi^H = \delta_{\lambda_0} \tag{3.88}$$

is the Dirac measure at $\lambda = \lambda_0$, and $f(H)\psi = f(\lambda_0)\psi$ for all bounded Borel functions. Given any selfadjoint A with $\psi \in \operatorname{dom} A$, the measure $\mu^A_{U_H(t)\psi}$ is t-independent.

Proof: Since $H\psi = \lambda_0\psi$, we have

$$0 = \|(H - \lambda_0)\psi\|^2 = \int_{\mathbb{R}} (\lambda - \lambda_0)^2 \, d\mu^H_\psi(\lambda).$$

Hence μ^H_ψ is supported in $\{\lambda_0\}$, which proves the first claim because it is a probability measure. For any bounded function, we therefore have

$$\|f(H)\psi - f(\lambda_0)\psi\|^2 = \int_{\mathbb{R}} |f(\lambda) - f(\lambda_0)|^2 \, d\delta_{\lambda_0}(\lambda) = 0,$$

i.e. $f(H)\psi = f(\lambda_0)\psi$ and in particular $e^{-itH}\psi = e^{-it\lambda_0}\psi$, $t \in \mathbb{R}$. This implies that the vector state given by $e^{-itH}\psi$ is time-independent, and in particular $\mu^A_{U_H(t)\psi}$ is t-independent. □

This finishes the summary of the most important items of the mathematical framework of Hilbert space quantum mechanics.

3.3 Examples and spectral analysis of quantum Hamiltonians

In this section we consider quantum Hamiltonians defined on a subspace $\operatorname{dom} H$ of $\mathcal{H} = L^2(\mathbb{R}^d)$ that are of the form $H = \frac{1}{2m}P^2 + V(X)$, i.e.

$$(H\psi)(x) = -\frac{1}{2m}(\Delta\psi)(x) + V(x)\psi(x). \tag{3.89}$$

Here d is the spatial dimension, $m > 0$ the mass of the particle (we often set it to $m = \frac{1}{2}$) and $V : \mathbb{R}^d \to \mathbb{R}$ the potential of a conservative force field. Such Hamiltonians can be considered as the quantum counterparts of the Hamiltonian functions we studied in classical mechanics.

Two natural questions are: 1) Is H (essentially) selfadjoint? (answering this puts us in the framework discussed in the previous section) and 2) What can we say about the spectrum of H and the specral measure E^H? In particular, what can we say about the point spectrum? The spectral measure is essential for understanding the dynamics given by H, and the point spectrum informs us about the stationary states.

We first look at the simple case of a continuous compactly supported potential in one dimension.

Proposition 3.33. *Let $V : \mathbb{R} \to \mathbb{R}$ be continuous and compactly supported (in particular bounded), and set $V_- := \inf_x V(x)$.*

a) $H = P^2 + V(X)$ is essentially selfadjoint on $\mathcal{S}(\mathbb{R})$,

b) $(0, \infty) \subset \sigma_c(H) \subset \sigma(H)$,

c) $\sigma_p(H) \subset (V_-, 0]$

Proof. a) has been shown in Problem 3.4.

c) Let $\lambda \in \mathbb{R}$ be an eigenvalue of H, i.e. $H\psi = \lambda\psi$ for some $\psi \in \operatorname{dom} H = \operatorname{dom} P^2$, $\psi \neq 0$. We will present the proof under the simplifying assumption that ψ is twice differentiable (this is true in a generalized sense). Then $-\psi'' + V \cdot \psi = \lambda\psi$. Outside the compact support of ψ, this implies

$$\psi''(x) = -\lambda\psi(x) \implies \psi(x) = \alpha e^{i\sqrt{\lambda}x} + \beta e^{-i\sqrt{\lambda}x}, \qquad x \notin \operatorname{supp} V.$$

For $\lambda \geq 0$, these functions do not decay as $x \to \pm\infty$, and do not lie in $L^2(\mathbb{R})$. Hence only eigenvalues $\lambda < 0$ can exist.

Now suppose λ is an eigenvalue with $\lambda \leq V_-$, with normalized eigenvector ψ. Then

$$\begin{aligned} 0 &= \langle \psi, (H-\lambda)\psi \rangle \\ &= \langle \psi, P^2\psi \rangle + \langle \psi, (V(X)-\lambda)\psi \rangle \\ &= \|P\psi\|^2 + \int_{\mathbb{R}} [V(x)-\lambda]|\psi(x)|^2\, dx \\ &\geq 0, \end{aligned}$$

where we have estimated $V(x) - \lambda \geq V_- - \lambda \geq 0$ in the last step. This implies $P\psi = 0$ and hence $\psi = 0$, a contradiction.

b) It remains to show that any $\lambda > 0$ lies in the (continuous) spectrum of H. We set $k := \sqrt{\lambda}$, choose $\chi \in C_c^\infty(\mathbb{R})$ with $\|\chi\|_2 = 1$ and define, for $n \in \mathbb{N}$, the function

$$\psi_n(x) := \eta_n(x)\, e^{ikx}, \qquad \eta_n(x) := n^{-1/2}\chi\left(\frac{x-n^2}{n}\right).$$

Then $\psi_\in \operatorname{dom} H$ with $\|\psi_n\| = 1$ for all n. Because of the shift in the argument of χ, the support of ψ_n is disjoint from the support of V for large enough n. Hence

$$(H-\lambda)\psi_n(x) = -\psi_n''(x) - \lambda\psi_n(x) = (-\eta_n''(x) - 2ik\eta_n'(x))e^{ikx}.$$

In view of the estimates

$$\|\eta_n'\| = \frac{1}{n}\|\chi'\|_2, \qquad \|\eta_n''\| = \frac{1}{n^2}\|\chi''\|_2,$$

this results in

$$\|(H-\lambda)\psi_n\|_2 \leq \frac{1}{n^2}\|\chi''\|_2 + \frac{2k}{n}\|\chi'\|_2 \to 0$$

as $n \to \infty$. Thus we have constructed a sequence of normalized vectors ψ_n such that $\|(H-\lambda)\psi_n\| \to 0$ (This is called a Weyl sequence, see functional analysis). This is incompatible with $(H-\lambda)$ having a bounded inverse, i.e. λ must lie in the spectrum of H. □

For non-negative V, these results imply in particular that no eigenvalues exist. With more work, one can show that in case there exists $x \in \mathbb{R}$ with $V(x) < 0$, there exists at least one eigenvalue.

To illustrate an opposite situation in which only point spectrum exists, we also consider the case of a confining potential. An example of such a situation is the harmonic oscillator with potential $V(x) = x^2$.

Proposition 3.34. *Let $V : \mathbb{R} \to \mathbb{R}$ be such that $V(x) \geq 1$ and $\lim_{x\to\pm\infty} V(x) = \infty$, and assume that $H = P^2 + V(X)$ is essentially selfadjoint on $\mathcal{S}(\mathbb{R}^d)$. Then H^{-1} is compact and consequently there exists an orthonormal basis of eigenvectors of H.*

We now consider techniques that apply to more complicated but important potentials in $d = 3$ dimensions.

Definition 3.35 (Relatively bounded operators). Let A and B be two densely defined operators on a Hilbert space $\mathcal{H}$. We say that b is relatively bounded w.r.t. A if $\operatorname{dom} A \subset \operatorname{dom} B$ and there exist $a, b > 0$ such that

$$\|B\psi\| \leq a\|A\psi\| + b\|\psi\|, \qquad \psi \in \operatorname{dom} A. \tag{3.90}$$

The infimum over all $a > 0$ satisfying this equation is called the A-bound of B.

Theorem 3.36 (Kato-Rellich Theorem). *Let A, B be two densely defined symmetric operators such that B is relatively bounded w.r.t. A, with A-bound less than 1. If $(\operatorname{dom} A, A)$ is (essentially) selfadjoint, then $H := A + B : \operatorname{dom} A \to \mathcal{H}, \psi \mapsto A\psi + B\psi$, is (essentially) selfadjoint.*

Proof. We give the proof in the selfadjoint case, the essentially selfadjoint case is left as an exercise.

According to Theorem M.47, we need to show $\operatorname{Ran}(A + B \pm i) = \mathcal{H}$. This is clearly equivalent to showing that $\operatorname{Ran}(A + B \pm i\lambda) = \mathcal{H}$ for some $\lambda > 0$.

We begin by noting that for $\lambda > 0$, the operators $(A \pm i\lambda)^{-1} = f(A)$ and $A(A \pm i\lambda)^{-1} = g(A)$ are bounded (with $f, g : \mathbb{R} \to \mathbb{C}$, $f(x) = (x \pm i\lambda)^{-1}$ and $g(x) = x(x \pm i\lambda)^{-1}$), with the bounds $\|f(A)\| \leq \|f\|_\infty = \lambda^{-1}$ and $\|g(A)\| \leq \|g\|_\infty = 1$.

In view of the relative bound between B and A, we have for any $\psi \in \mathcal{H}$

$$\|B(A \pm i\lambda)^{-1}\psi\| \leq a\|A(A \pm i\lambda)^{-1}\psi\| + b\|(A \pm i\lambda)^{-1}\psi\| \leq \left(a + \frac{b}{\lambda}\right)\|\psi\|.$$

Hence, if we choose $\lambda > \frac{b}{1-a}$ (here the assumption $a < 1$ enters), we have

$$\|B(A \pm i\lambda)^{-1}\| \leq a + \frac{b}{\lambda} < 1.$$

By the geometric series (Neumann series), $1 + B(A \pm i\lambda)^{-1} : \mathcal{H} \to \mathcal{H}$ is bijective (the inverse is $\sum_{n=0}^{\infty}(B(A \pm i\lambda))^{-n}$). As also $(A \pm i\lambda)$ is bijective because A is selfadjoint, we find that

$$(A + B \pm i\lambda)\psi = (1 + B(A \pm i\lambda)^{-1})(A \pm i\lambda)\psi$$

ranges over all of $\mathcal{H}$ when ψ ranges over $\operatorname{dom} A$. This finishes the proof. □

Let us look at an important application of the Kato-Rellich Theorem to Schrödinger operators in dimension $d = 3$.

Corollary 3.37. *Let $V_1 \in L^2(\mathbb{R}^3)$, $V_2 \in L^\infty(\mathbb{R}^3)$, and $V := V_1 + V_2 : \mathbb{R}^3 \to \mathbb{R}$, be real. Then $H := P^2 + V(X) : \mathcal{S}(\mathbb{R}^3) \to L^2(\mathbb{R}^3)$ is essentially selfadjoint.*

Proof. As we know that $P^2 = -\Delta$ is essentially selfadjoint on $\mathcal{S}(\mathbb{R}^3)$, we only need to show that $V(X)$ relatively bounded w.r.t. P^2. For $\psi \in \mathcal{S}(\mathbb{R}^3)$, we begin with the estimate

$$\|V(X)\psi\| \leq \|V_1\|\|\psi\|_\infty + \|V_2\|_\infty\|\psi\|.$$

The second term $\|V_2\|_\infty\|\psi\|$ is harmless for the relative bound, but we need to estimate the first term $\|V_1\|\|\psi\|_\infty$. To that end, we use the Fourier transform and the Cauchy-Schwarz Inequality with the function $h : \mathbb{R}^d \to \mathbb{R}$, $h(p) = \|p\|^2$:

$$\begin{aligned}\|\psi\|_\infty &\leq (2\pi)^{d/2}\|\tilde{\psi}\|_1 \\ &= (2\pi)^{d/2}\|(h+1)^{-1}(h+1)\tilde{\psi}\|_1 \\ &\leq (2\pi)^{d/2}\left(\int_{\mathbb{R}^3}(\|p\|^2+1)^{-2}\,dp\right)^{1/2}\|(h+1)\tilde{\psi}\|_2\end{aligned}$$

Up to a multiplicative constant, this integral is $\int_0^\infty r^2(r^2+1)^{-2}dr < \infty$ by swichting to polar coordinates, hence finite. This results in the bound

$$\|\psi\|_\infty \leq C\|P^2\psi\|_2 + C'\|\psi\|_2$$

with numerical constants C, C'.

To produce a small relative bound, we now scale ψ: For $r > 0$, we consider $\tilde{\psi}_r(p) := r^3\tilde{\psi}(rp)$. Then one has $\|\tilde{\psi}_r\|_1 = \|\tilde{\psi}\|_1$ and $\|\tilde{\psi}_r\|_2 = r^{3/2}\|\tilde{\psi}\|_2$, as well as $\|P^2\tilde{\psi}_r\|_2 = r^{-1/2}\|P^2\tilde{\psi}\|_2$. Replacing $\|\tilde{\psi}\|_1$ by $\|\tilde{\psi}_r\|_1$ in the previous bound, this allows us to estimate for any $r > 0$

$$\|\psi\|_\infty \leq C\|P^2\psi_r\|_2 + C'\|\psi_r\|_2 = \frac{C}{\sqrt{r}}\|P^2\psi\|_2 + C'r^{3/2}\|\psi\|_2.$$

For large enough r, this gives a relative bound of V w.r.t. P^2 less than 1. □

The most important application of this result is to the Coulomb potential.

Example 3.38 (The Hamiltonian with Coulomb potential). Consider the Coulomb potential

$$V : \mathbb{R}^3 \to \mathbb{R}, \qquad V(x) := -\frac{Z}{\|x\|}, \qquad Z > 0,$$

and split it according to $V = V_1 + V_2$, with $V_1 := \chi_1 V$, $V_2 := (1-\chi_1)V$, where χ_1 is the characeristic function of the unit ball $\|x\| \leq 1$. Then $V_2 \in L^\infty(\mathbb{R}^3)$ and $V_1 \in L^2(\mathbb{R}^3)$ (integrate in polar coordinates). Hence the symmetric operator $V(X) : \mathcal{S}(\mathbb{R}^3) \to L^2(\mathbb{R}^3)$ satisfies the assumptions of the corollary, so $H = P^2 + V(X)$ is essentially selfadjoint on $\mathcal{S}(\mathbb{R}^d)$.

At this point, we have seen some useful criteria that allow us to decide about the question of (essential) selfadjointness of typical Hamiltonians. Let us now have a closer look at the spectrum of H and how it influences the dynamics. The idea in the following definition is that the measures μ_ψ^H can be quite different (for fixed H) depending on ψ.

Definition 3.39. Let μ be a (signed, complex) Borel measure on $\mathbb{R}$, and let ℓ denote Lebesgue measure on $\mathbb{R}$.

a) μ is called *pure point* if $\mu = \sum_n c_n \delta_{\lambda_n}$, a finite or countable sum of Dirac measures,

b) μ is called *absolutely continuous* (w.r.t. ℓ), written $\mu \ll \ell$, if every ℓ-zero set is a μ-zero set. (In this case, μ has a density w.r.t. ℓ (Thm. M.9)),

c) μ is called *singular continuous* (w.r.t. ℓ) if $\operatorname{supp}\mu$ is a Lebesgue zero set, but also $\mu(\{\lambda\}) = 0$ for all $\lambda \in \mathbb{R}$.

Definition 3.40 (Spectral types). Let $(\mathcal{H}, H)$ be a quantum-mechanical system. We define

$$\mathcal{H}^H_{\mathrm{pp}} := \{\psi \in \mathcal{H} \ : \ \mu^H_\psi \text{ is pure point}\}, \tag{3.91}$$

$$\mathcal{H}^H_{\mathrm{ac}} := \{\psi \in \mathcal{H} \ : \ \mu^H_\psi \text{ is absolutely continuous}\}, \tag{3.92}$$

$$\mathcal{H}^H_{\mathrm{sc}} := \{\psi \in \mathcal{H} \ : \ \mu^H_\psi \text{ is singular continuous}\}. \tag{3.93}$$

The three sets $\mathcal{H}^H_{\mathrm{pp}}, \mathcal{H}^H_{\mathrm{ac}}, \mathcal{H}^H_{\mathrm{sc}}$ are closed subspaces of $\mathcal{H}$. Let us prove this for $\mathcal{H}^H_{\mathrm{ac}}$: A vector $\psi \in \mathcal{H}$ lies in $\mathcal{H}^H_{\mathrm{ac}}$ if and only if $\mu^H_\psi(N) = 0$ for every Lebesgue null set $N \subset \mathbb{R}$. Since $\mu^H_\psi(N) = \langle \psi, E^H(N)\psi\rangle = \|E^H(N)\psi\|^2$, we see

$$\mathcal{H}^H_{\mathrm{ac}} = \bigcap_{N\,:\,\ell(N)=0} \ker E^H(N),$$

which is a closed linear subspace. In the exercises, you will prove

$$\mathcal{H}^H_{\mathrm{pp}} = \overline{\operatorname{span}\{E^H(\{\lambda\})\mathcal{H} \ : \ \lambda \in \mathbb{R}\}},$$

i.e. $\mathcal{H}^H_{\mathrm{pp}}$ is the closed subspace generated by all eigenspaces of H. Singular continuous spectrum is often absent[17], and we will focus mostly on the pure point and absolutely continuous parts. An example of a singular continuous measure is the Cantor measure.

It is clear by definition of these spaces that they are mutually orthogonal to each other. Without proof we state the following stronger result. See, for example, [Tes09, Lemma 3.19].

Lemma 3.41. *Let $(\mathcal{H}, H)$ be a quantum mechanical system, with $\mathcal{H}$ separable. Then*

$$\mathcal{H} = \mathcal{H}^H_{\mathrm{pp}} \oplus \mathcal{H}^H_{\mathrm{ac}} \oplus \mathcal{H}^H_{\mathrm{sc}}, \tag{3.94}$$

and there exist Borel sets $\mathbb{B}_{\mathrm{pp}}, \mathbb{B}_{\mathrm{ac}}, \mathbb{B}_{\mathrm{sc}} \subset \mathbb{R}$ such that

$$\mathcal{H}^H_{\mathrm{pp}} = E^H(\mathbb{B}_{\mathrm{pp}})\mathcal{H}, \qquad \mathcal{H}^H_{\mathrm{ac}} = E^H(\mathbb{B}_{\mathrm{ac}})\mathcal{H}, \qquad \mathcal{H}^H_{\mathrm{sc}} = E^H(\mathbb{B}_{\mathrm{sc}})\mathcal{H}. \tag{3.95}$$

[17] Singular continuous spectrum but may appear in Hamiltonians on discrete structures, for example in the so-called Fibonacci Hamiltonian on $\mathcal{H} = \ell^2_{\mathbb{Z}}$, defined by $(Hu)_n = u_{n+1} + u_{n-1} + \lambda V_n u_n$, where V_n is related to the Fibonacci sequence. Its spectrum is a Cantor-like set with spectral measures purely singular continuous for non-zero λ, it is a model of a quasicrystal.

Note that this result implies that these subspaces are invariant under the dynamics, e.g. $e^{-itH}\mathcal{H}^H_{\mathrm{pp}} = e^{-itH}E^H(\mathbb{B}_{\mathrm{pp}})\mathcal{H} = E^H(\mathbb{B}_{\mathrm{pp}})e^{-itH}\mathcal{H} = \mathcal{H}^H_{\mathrm{pp}}$. Furthermore, note that

$$\psi \in \mathcal{H}^H_{\mathrm{xx}} \implies \mu^H_{\varphi,\psi} \text{ is } xx \text{ for all } \varphi \in \mathcal{H}, \tag{3.96}$$

where "xx" stands for one of "pp", "ac", "sc". To prove (3.96) for the absolutely continuous case, let $N \subset \mathbb{R}$ be a Lebesgue null set and $\psi \in \mathcal{H}^H_{\mathrm{ac}}, \varphi \in \mathcal{H}$. Then $\mu^H_{\varphi,\psi}(N) = \langle \varphi, E^H(N)\psi\rangle = 0$, i.e. $\mu^H_{\varphi,\psi}$ is absolutely continuous w.r.t. Lebesgue. The corresponding proof for the pure point space is again carried out in the exercises.

We now want to understand the dynamical aspects of these types of spectra. For example, we might wonder how $|\langle \varphi, e^{-itH}\psi\rangle|^2$ looks like for large times and how this behaviour is influenced by ψ lying in one the three spectral spaces. This, on the one hand, informs us about the probability that when measuring whether the state of the system is φ in the state $e^{-itH}\psi$, the answer is "yes". On the other hand, for $\mathcal{H} = L^2(\mathbb{R}^d)$ and $\varphi = \chi_{\mathbb{B}}$ a characteristic function,

$$|\langle \varphi, e^{-itH}\psi\rangle|^2 = \left|\int_{\mathbb{B}} (e^{-itH}\psi)(x)\,dx\right|^2 \tag{3.97}$$

is sensitive to the part of $e^{-itH}\psi$ supported on $\mathbb{B}$, and can therefore serve as a tool to check the behaviour of $e^{-itH}\psi$, i.e. whether it stays in bounded regions, moves to infinity, etc.

We first consider the easiest case of the pure point spectrum.

Lemma 3.42 (Pure point vectors define bound states). *Let H be a selfadjoint Hamiltonian on $\mathcal{H} = L^2(\mathbb{R}^d)$, $\psi \in \mathcal{H}^H_{\mathrm{pp}}$, and $\varepsilon > 0$. Then there exists $R > 0$ such that*

$$\sup_{t\in\mathbb{R}} \Pr_{e^{-itH}\psi}(\|X\| > R) = \sup_{t\in\mathbb{R}} \int_{\|x\|>R} |(e^{-itH}\psi)(x)|^2\,dx < \varepsilon. \tag{3.98}$$

The interpretation of this result is the following: For $\psi \in \mathcal{H}^H_{\mathrm{pp}}$, there exists a finite ball $B_R := \{x : \|x\| \le R\}$ such that the probability that the position takes values in B_R is essentially one, at all times. Hence there is no movement to infinity in such a state, and one therefore calls it a *bound state*.

Proof. Let us first consider the special case that ψ is an eigenvector of H. Then $|(e^{-itH}\psi)(x)| = |\psi(x)|$ for all t, and since $\psi \in L^2(\mathbb{R}^d)$, we have $\int_{\|x\|>R} |\psi(x)|^2\,dx \to 0$ for $R \to \infty$. For any $\varepsilon > 0$, we can therefore choose R large enough such that $\int_{\|x\|>R} |\psi(x)|^2\,dx < \varepsilon$.

If $\psi = \sum_n c_n \psi_n$ is a finite linear combination of eigenvectors, we have

$$|(e^{-itH}\psi)(x)| = \left|\sum_n c_n\, e^{-it\lambda_n}\, \psi_n(x)\right| \le \sum_n |c_n|\,|\psi_n(x)|,$$

and the same argument works again. Now let ψ be a general vector in $\mathcal{H}^H_{\mathrm{pp}}$, i.e. $\psi = \lim_n \xi_n$ (i.e. $\|\psi - \xi_n\| \to 0$) with each ξ_n a linear combination of eigenvectors. Let χ_R denote the characteristic function of B_R. Then

$$\begin{aligned}\|(1-\chi_R)e^{-itH}\psi\| &\le \|(1-\chi_R)e^{-itH}\xi_n\| + \|(1-\chi_R)e^{-itH}(\psi-\xi_n)\| \\ &\le \|(1-\chi_R)e^{-itH}\xi_n\| + \|\psi-\xi_n\|.\end{aligned}$$

Given $\varepsilon > 0$, we choose $n = n_\varepsilon$ such that $\|\psi - \xi_n\| < \varepsilon$. For this fixed n_ε, we then choose R large enough such that $\|(1-\chi_R)e^{-itH}\xi_n\| < \varepsilon$. Taking the square then results in $\sup_{t\in\mathbb{R}} \Pr_{e^{-itH}\psi}(\|X\| > R) < 4\varepsilon^2$. □

This relatively simple result gives us a good qualititative understanding of the dynamics of pure point vectors as bound states. To also understand the continuous spectrum, the following result will be a useful tool.

Proposition 3.43 (Wiener's Theorem). *Let μ be a finite complex Borel measure on $\mathbb{R}$, and define the Fourier transform of μ as*

$$\tilde{\mu} : \mathbb{R} \to \mathbb{C}, \qquad \tilde{\mu}(t) := \int_{\mathbb{R}} e^{-it\lambda}\, d\mu(\lambda), \tag{3.99}$$

a) Then

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T |\tilde{\mu}(t)|^2\, dt = \sum_{\lambda\in\mathbb{R}} |\mu(\{\lambda\})|^2, \tag{3.100}$$

where the sum on the right hand side is finite.

b) If μ is absolutely or singular continuous w.r.t. Lebesgue measure,

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T |\tilde{\mu}(t)|^2\, dt = 0, \tag{3.101}$$

c) If μ is absolutely continuous w.r.t. Lebesgue measure,

$$\lim_{t\to\infty} \tilde{\mu}(t) = 0. \tag{3.102}$$

Proof. a) We calculate with Fubini's Theorem

$$\begin{aligned}\frac{1}{T}\int_0^T |\tilde{\mu}(t)|^2\, dt &= \frac{1}{T}\int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}} e^{-i(x-y)t}\, d\mu(x)\, d\overline{\mu}(y)\, dt \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} \left(\frac{1}{T}\int_0^T e^{-i(x-y)t} dt\right) d\mu(x)\, d\overline{\mu}(y).\end{aligned}$$

The function in brackets is bounded by 1 and converges pointwise to $\chi_{\{0\}}(x-y)$ as $T\to\infty$. Hence we may apply dominated convergence and see that the limit $T\to\infty$ of the above integral is

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \chi_{\{0\}}(x-y) d\mu(x)\, d\overline{\mu}(y) = \int_{\mathbb{R}} \mu(\{y\})\, d\overline{\mu}(y) = \sum_{y\in\mathbb{R}} |\mu(\{y\})|^2.$$

b) follows directly from a) and the definition of absolutely/singular continuous measure. c) is the Riemann-Lebesgue lemma. □

Example 3.44. Let H be a selfadjoint Hamiltonian on $\mathcal{H} = L^2(\mathbb{R}^d)$, and let $\psi \in \mathcal{H}^H_{\mathrm{ac}}$, $\|\psi\| = 1$. Then $e^{-itH}\psi \to 0$ weakly as $t \to \pm\infty$. In particular, for any bounded Borel set $\mathbb{B} \subset \mathbb{R}^d$,

$$\lim_{t\to\pm\infty} \int_{\mathbb{B}} (e^{-itH}\psi)(x)\, dx = 0. \tag{3.103}$$

Proof: As $\psi \in \mathcal{H}^H_{\mathrm{ac}}$, the complex Borel measure $\mu^H_{\varphi,\psi}$ is absolutely continuous w.r.t. Lebesgue measure, for every $\varphi \in \mathcal{H}$. Hence

$$\langle \varphi, e^{-itH}\psi\rangle = \int_{\mathbb{R}} e^{-it\lambda}\, d\mu^H_{\varphi,\psi}(\lambda) \to 0, \qquad t \to \pm\infty,$$

by the preceding proposition. The second statement follows by taking $\varphi = \chi_{\mathbb{B}}$. □

We would like to upgrade the limit (3.103) to the stronger statement that

$$\mathrm{Pr}_{e^{-itH}\psi}(X \in \mathbb{B}) = \int_{\mathbb{B}} |(e^{-itH}\psi)(x)|^2\, dx \tag{3.104}$$

converges to 0 for $t \to \pm\infty$. This has the clear interpretation that the probability to find the particle described by ψ in the region $\mathbb{B}$ goes to zero as $t \to \pm\infty$.

The abstract mechanism that allows for such an upgrade is described in the following theorem. We call it the "little RAGE theorem" because it is a weakened version of the so-called RAGE Theorem (abbreviating the authors Ruelle-Amrein-Georgescu-Enss). The full RAGE Theorem gives a full characterization of the continuous and pure point subspaces in terms of sequences of relatively compact operators. For us, the following version will be sufficient.

Theorem 3.45 (The little RAGE theorem). *Let $(\mathcal{H}, H)$ be a quantum-mechanical system, $\psi \in \mathcal{H}$ and $K \in B(\mathcal{H})$ a bounded operator such that $K(H+i)^{-1}$ is compact. Then, with P_{c}, P_{ac} denoting the orthogonal projections onto $\mathcal{H}^H_{\mathrm{ac}} \oplus \mathcal{H}^H_{\mathrm{sc}}$ and $\mathcal{H}^H_{\mathrm{ac}}$, respectively,*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \|Ke^{-itH}P_{\mathrm{c}}\psi\|^2\, dt = 0, \qquad \lim_{t\to\infty} \|Ke^{-itH}P_{\mathrm{ac}}\psi\| = 0. \tag{3.105}$$

Proof. We first give the proof under the much stronger assumption that K is of finite rank. Choosing an orthonormal basis $\{\psi_1, \dots, \psi_n\}$ of $\mathrm{Ran}\, K$, we have $K\xi = \sum_{k=1}^n \langle K^*\psi_k, \xi\rangle \psi_k$ and $\|K\xi\|^2 = \sum_{k=1}^n |\langle K^*\psi_k, \xi\rangle|^2$, and hence

$$\begin{aligned}
\frac{1}{T}\int_0^T \|Ke^{-itH}P_{\mathrm{c}}\psi\|^2\, dt &= \frac{1}{T}\sum_{k=1}^n \int_0^T |\langle K^*\psi_k, e^{-itH}P_{\mathrm{c}}\psi|^2\, dt \\
&= \sum_{k=1}^n \frac{1}{T}\int_0^T |\tilde{\mu}^H_{K^*\psi_k, P_{\mathrm{c}}\psi}(t)|^2\, dt \longrightarrow 0 \quad \text{as } t \to \infty \\
\|Ke^{-itH}P_{\mathrm{ac}}\psi\|^2 &= \sum_{k=1}^n |\langle K^*\psi_k, e^{-itH}P_{\mathrm{ac}}\psi|^2 \\
&= \sum_{k=1}^n |\tilde{\mu}^H_{K^*\psi_k, P_{\mathrm{c}}\psi}(t)|^2 \longrightarrow 0 \quad \text{as } t \to \infty
\end{aligned}$$

by Wiener's Theorem 3.43. If K is compact, then there exists a sequence (K_n) of finite rank operators such that $\|K - K_n\| \leq \frac{1}{n}$. Estimating $\|Ke^{-itH}P_c\psi\| \leq \|K_n e^{-itH}P_c\psi\| + \frac{1}{n}\|P_c\psi\|$ (and similarly for P_{ac}) shows that the result holds if K is compact.

Now we pass to the general situation in which only $K(H+i)^{-1}$ is assumed to be compact. For $\psi \in \mathcal{H}$, let $\psi_n \in \operatorname{dom} H$ such that $\|\psi - \psi_n\| \leq \frac{1}{n}$. Then

$$\|Ke^{-itH}P_c\psi\| \leq \|Ke^{-itH}P_c\psi_n\| + \frac{\|K\|}{n} = \|K(H+i)^{-1}e^{-itH}P_c(H+i)\psi_n\| + \frac{\|K\|}{n}.$$

Using the same argument as before for the compact operator $K(H+i)^{-1}$ finishes the proof. □

To apply this theorem, we need to find meaningful operators K such that $K(H+i)^{-1}$ is compact (in this situation, one also says that K is *relatively compact w.r.t. H*). We first collect some preliminaries on Hilbert-Schmidt operators.

Definition 3.46 (Hilbert-Schmidt operators). An operator $K : \mathcal{H} \to \mathcal{H}$ is called *Hilbert-Schmidt* if for some orthonormal basis (e_n) of $\mathcal{H}$, the *Hilbert-Schmidt norm*

$$\|K\|_2^2 := \sum_n \|Ke_n\|^2 < \infty \tag{3.106}$$

is finite. The set of all Hilbert-Schmidt operators on $\mathcal{H}$ is denoted $\mathcal{I}_2(\mathcal{H})$.

By realizing $\sum_n \|Ke_n\|^2 = \operatorname{Tr}(K^*K)$, one sees that the Hilbert-Schmidt norm and the condition (3.106) is independent of the considered orthonormal basis. We mention without proof here the fact that $\mathcal{I}_2(\mathcal{H})$ is a Hilbert space w.r.t. the scalar product

$$\langle K, L\rangle := \operatorname{Tr}(K^*L), \tag{3.107}$$

and a two-sided *-ideal in $B(\mathcal{H})$, namely with $K \in \mathcal{I}_2(\mathcal{H})$, also $K^*, AK, KA \in \mathcal{I}_2(\mathcal{H})$, for any $A \in B(\mathcal{H})$.

Proposition 3.47 (Some facts about Hilbert-Schmidt operators).
a) $\|K\| \leq \|K\|_2$ *for any* $K \in \mathcal{I}_2(\mathcal{H})$.
b) Hilbert-Schmidt operators are compact.
c) Let (X,μ) *is a measure space,* $\mathcal{H} = L^2(X,\mu)$*, and let* $k \in L^2(X\times X, \mu\times\mu)$*. Then*

$$K : L^2(X,\mu) \to L^2(X,\mu), \qquad (Kf)(x) = \int_X k(x,y)f(y)\,d\mu(y)$$

is a Hilbert-Schmidt operator, with $\|K\|_2 = \|k\|_{L^2(X\times X,\mu\times\mu)}$.

Proof. a) Let (e_n) be an orthonormal basis of $\mathcal{H}$ and $\psi \in \mathcal{H}$. Then $\psi = \sum_n c_n e_n$ (with $c_n = \langle e_n, \psi\rangle$ square summable), and

$$\|K\psi\|^2 = \left\|\sum_n c_n Ke_n\right\|^2 \leq \sum_n |c_n|^2 \cdot \sum_m \|Ke_m\|^2 = \|\psi\|^2\|K\|_2^2,$$

which implies the claimed result.

b) Let (e_n) be an orthonormal basis of $\mathcal{H}$, and P_N the orthogonal projection onto $\operatorname{span}\{e_1, \dots, e_N\}$. For $K \in \mathcal{I}_2(\mathcal{H})$, the operator $K_N := KP_N$ has finite rank and is therefore compact. We have

$$\|K - K_N\|^2 \le \|K - K_N\|_2^2 = \sum_{n>N} \|Ke_n\|^2 \to 0 \quad \text{as } N \to \infty.$$

Hence K lies in the norm closure of the finite rank operators, i.e. K is compact.

c) Let (e_n) be an orthonormal basis of $L^2(X,\mu)$. The function $k_x(y) := k(x,y)$ lies in $L^2(X,\mu)$ for almost every $x \in X$. For such x, we have $(Ke_n)(x) = \int_X k(x,y)e_n(y)d\mu(y) = \langle k_x, e_n\rangle$, and

$$\sum_n |(Ke_n)(x)|^2 = \sum_n |\langle k_x, e_n\rangle|^2 = \|k_x\|^2 = \int_X |k(x,y)|^2 \, d\mu(y).$$

This implies via Tonelli's Theorem

$$\begin{aligned}\|K\|_2^2 &= \sum_n \int_X |(Ke_n)(x)|^2 d\mu(x) = \int_X \sum_n |(Ke_n)(x)|^2 d\mu(x) \\ &= \int_X \int_X |k(x,y)|^2 \, d\mu(x)\, d\mu(y) = \|k\|_{L^2(X\times X)}.\end{aligned}$$

□

Example 3.48. Let $\chi \in C_c^\infty(\mathbb{R}^d)$ be a smooth function of compact support, $K := \chi(X) \in B(L^2(\mathbb{R}^d))$ the operator multiplying with it, and let $H_0 = P^2$ denote the selfadjoint extension of the free Hamiltonian on $\mathcal{S}(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$. Then $K(H_0+i)^{-1}$ is compact.

Proof: Let χ_n denote the characeristic function of the ball $B_n \subset \mathbb{R}^d$ with radius n and origin $p = 0$. Then, for any $\psi \in L^2(\mathbb{R}^d)$, we have

$$\begin{aligned}(K(H_0+i)^{-1}\chi_n(P)\psi)(x) &= \chi(x)((H_0+i)^{-1}\chi_n(P)\psi)(x) \\ &= (2\pi)^{-d/2} \int_{B_n} \frac{\chi(x)e^{i\langle p,x\rangle}}{\|p\|^2+i} \tilde{\psi}(p)\, dp.\end{aligned}$$

As the function $(x,p) \mapsto \frac{\chi(x)e^{i\langle p,x\rangle}}{\|p\|^2+i}\chi_n(p)$ lies in $L^2(\mathbb{R}^d \times \mathbb{R}^d, dx\, dp)$, we see that $K(H_0+i)^{-1}\chi_n(P)$ is Hilbert-Schmidt and in particular compact.

Furthermore, the operator $(H_0+i)^{-1}(1-\chi_n(P))$ acts in momentum space by multiplication with $p \mapsto (\|p\|^2+i)^{-1}(1-\chi_n(p))$, and therefore has norm

$$\|(H_0+i)^{-1}\chi_n(P)\| = (n^4+1)^{-1/2}.$$

This implies $K(H_0+i)^{-1}\chi_n(P) \to K(H_0+i)^{-1}$ in operator norm as $n \to \infty$, and it follows that $K(H_0+i)^{-1}$ is compact. □

Proposition 3.49 (Dynamical properties of (absolutely) continous spectrum).

a) Let $(\operatorname{dom} H_0, H_0)$ *be a selfadjoint operator and* V *relatively bounded w.r.t.* H_0*, with relative bound less than* 1*. Then, if* $K \in B(\mathcal{H})$ *is relatively compact w.r.t.* H_0*, it is also relatively compact w.r.t.* $H = H_0 + V$*.*

b) On $\mathcal{H} = L^2(\mathbb{R}^d)$*, consider a potential* $V : \mathbb{R}^d \to \mathbb{R}$ *such that* V *is relatively bounded w.r.t.* $H_0 = P^2$ *on* $\operatorname{dom} H_0 = \mathcal{S}(\mathbb{R}^d)$*, with relative bounded less than* 1*. Then, for every* $\chi \in C_c^\infty(\mathbb{R}^d)$*, the operator* $\chi(X)(H+i)^{-1}$ *is compact.*

c) In the situation of b), we have for any bounded Borel set $\mathbb{B} \subset \mathbb{R}^d$

$$\frac{1}{T}\int_0^T \Pr_{e^{-itH}\psi}(X \in \mathbb{B})\, dt \longrightarrow 0, \qquad \psi \in \mathcal{H}_{\mathrm{ac}}^H \oplus \mathcal{H}_{\mathrm{c}}^H, \tag{3.108}$$

$$\Pr_{e^{-itH}\psi}(X \in \mathbb{B}) \longrightarrow 0, \qquad \psi \in \mathcal{H}_{\mathrm{ac}}^H. \tag{3.109}$$

Proof. a) We recall that the assumptions guarantee that $H = H_0 + V$ is selfadjoint on $\operatorname{dom} H_0$ by the Kato-Rellich Theorem. Furthermore, $(H_0+i)^{-1}\xi \in \operatorname{dom} H_0$ for all $\xi \in \mathcal{H}$, and the relative bound implies

$$\|V(H_0+i)^{-1}\xi\| \le a\|H_0(H_0+i)^{-1}\xi\| + b\|(H_0+i)^{-1}\xi\| \le (a+b)\|x\|$$

as in the proof of the Kato-Rellich Theorem. Hence $V(H_0+i)^{-1}$ is bounded. We claim that the *resolvent identity*

$$(H+i)^{-1} = (H_0+i)^{-1} - (H_0+i)^{-1}V(H+i)^{-1} \tag{3.110}$$

holds (as an equation in $B(\mathcal{H})$). Indeed, for every $\xi \in \mathcal{H}$, the vector $\psi := (H+i)^{-1}\xi$ satisfies $\xi = (H+i)\psi = (H_0+i)\psi + V\psi$, which implies

$$\begin{aligned}(H+i)^{-1}\xi &= \psi \\ &= (H_0+i)^{-1}\xi - (H_0+i)^{-1}V\psi \\ &= (H_0+i)^{-1}\xi - (H_0+i)^{-1}V(H+i)^{-1}\xi,\end{aligned}$$

namely (3.110). Multiplying (3.110) from the left by K shows that

$$K(H+i)^{-1} = K(H_0+i)^{-1} - K(H_0+i)^{-1}V(H+i)^{-1}$$

is compact as a sum of compact operators (note that $V(H+i)^{-1}$ is bounded, so also $K(H_0+i)^{-1}V(H+i)^{-1}$ is compact).

b) The relative compactness of $K := \chi(X)$ w.r.t. P^2 has been shown in Example 3.48. Corollary 3.37 gives examples of this situation.

c) Let $\chi \in C_c^\infty(\mathbb{R}^d)$ be such that $\chi(x) = 1$ for all $x \in \mathbb{B}$, and $0 \le \chi(x) \le 1$ for all $x \in \mathbb{R}^d$. Then

$$\Pr_{e^{-itH}\psi}(X \in \mathbb{B}) = \int_{\mathbb{B}} |(e^{-itH}\psi)(x)|^2\, dx \le \|\chi(X)e^{-itH}\psi\|^2,$$

and the claims follow from Theorem 3.45. □

This result gives a qualitative understanding of the dynamics of vectors in the (absolutely) continuous subspace: For any bounded set $\mathbb{B}$, the probability that the position of the particle

is in $\mathbb{B}$, converges to zero as $t \to \infty$. This limit holds only in an averaged sense for singular continuous spectrum, and as a limit without averages in the case of absolutely continuous spectrum. Such vector states are called *scattering states* because they behave like waves flowing to infinity at asymptotic times. They form the basis of quantum scattering theory, where scattering phenomena are investigated in more detail.

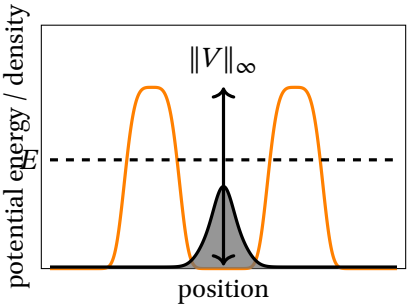
Scattering states behave exactly opposite to bound states that are given by vectors from the pure point space (Lemma 3.42).

Example 3.50 (Tunneling). Let us consider, for simplicity in $d = 1$ dimension, a compactly supported potential $V \in C_c(\mathbb{R})$ that is positive and consists of two potential barriers trapping a particle in the middle region $\mathbb{B}$ between the two barriers.

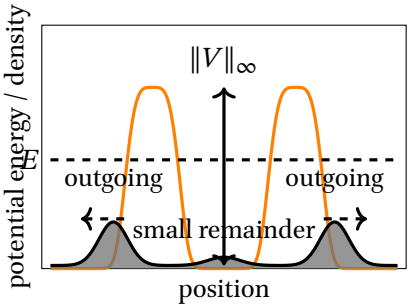
In this situation, we know that $H = P^2 + V(X)$ is essentially selfadjoint on Schwartz space (the potential is even bounded), and that H has no point spectrum (see Proposition 3.33). Hence Proposition 3.49 tells us that the probability that the position of the particle is in the trapped region $\mathbb{B}$ goes to zero as $t \to \infty$ (for singular continuous spectrum, this would only hold in an averaged sense, but here we have no singular continuous spectrum).

This means no matter how we design the initial vector state $\psi = \psi_0 \in L^2(\mathbb{R})$ at $t = 0$, the particle will escape the middle region $\mathbb{B}$ in the long run. This applies in particular to ψ such that $\tilde{\psi}$ has compact support, say such that $\Pr_\psi(H \leq V_0) = 1$, i.e. the probability that the energy of the particle is less than the potential barrier is 1.

(a) Initial state



(b) Typical late-time situation



In classical physics, in such a situation the portion of the probability distribution supported in $\mathbb{B}$ would remain in $\mathbb{B}$ for all times, for energy conservation reasons (see Corollary 1.10). In quantum mechanics, this is no longer the case: Although energy is conserved in the sense that H is a constant of motion, a quantum particle can "tunnel through the potential barrier", which is classically impossible.

This is the qualititative description of the *quantum tunneling effect*. It is clear from our discussion that it also applies in higher-dimensional systems.

3.4 Angular momentum on $L^2(\mathbb{R}^3)$

In this section, we prepare the ground for our subsequent analysis of Hamiltonians with spherically symmetric potentials, including in particular the hydrogen atom Hamiltonian, i.e. Schrödinger operators of the form

$$H = P^2 + V(X)$$

where $V : \mathbb{R}^3 \to \mathbb{R}$ is a potential satisfying $V(x) = v(\|x\|)$, $x \in \mathbb{R}^3$. (The hydrogen atom case is given by $V(x) = -\frac{Z}{\|x\|}$, see Example 3.38.) If H is essentially selfadjoint on,

say, $\mathcal{S}(\mathbb{R}^3)$, then it follows from Example 3.31 and Exercise 3.11 that its selfadjoint extension has the angular momentum operators L_1, L_2, L_3 (reviewed below) as constants of motion because H commutes with the unitary strongly continuous representation

$$U : \mathrm{SO}(3) \to \mathcal{U}(L^2(\mathbb{R}^3)), \qquad (U(R)\psi)(x) = \psi(R^{-1}x). \tag{3.111}$$

Clearly, rotations only affect the angular part. For example, if $\psi(x) = g(\phi,\theta)f(r)$ in polar coordinates (ϕ,θ) (angles) and $r = \|x\|$ (radius), then in $U(R)\psi$, the rotation only acts on g and not on f. We might therefore present our representation Hilbert space in the unitarily equivalent form

$$L^2(\mathbb{R}^3) \cong L^2(S^2, d\omega) \otimes L^2(\mathbb{R}_+, r^2\,dr), \tag{3.112}$$
$$U(R) \cong \tilde{U}(R) \otimes 1, \tag{3.113}$$

given by passing from cartesian to polar coordinates (here $d\omega(\phi,\theta) = \sin\theta\, d\phi\, d\theta$ in the usual polar coordinates).

In order to avoid unpleasant polar coordinates calculations, we will study the angular dependence in terms of certain polynomials in cartesian coordinates.

Definition 3.51 (Homogeneous harmonic polynomials).

a) A polynomial p in three variables is called *homogeneous of degree* $l \in \mathbb{N}_0$ if

$$p(\lambda x) = \lambda^l p(x), \qquad x \in \mathbb{R}^3,\ \lambda > 0. \tag{3.114}$$

The space of all these polynomials is denoted $\mathrm{Pol}_l(\mathbb{R}^3) \subset \mathrm{Pol}(\mathbb{R}^3)$.

b) A polynomial p in three variables is called *harmonic* if $\Delta p = 0$, where $\Delta = \partial_1^2 + \partial_2^2 + \partial_3^2$ denotes the Laplace differential operator. The space of all harmonic polynomials that are homogeneous of degree l is denoted $V_l \subset \mathrm{Pol}_l(\mathbb{R}^3)$.

The relation of these polynomials to functions in $L^2(S^2)$ is the following: Given any function $g : S^2 \to \mathbb{C}$ and degree $l \in \mathbb{N}_0$, we may extend g uniquely to an l-homogeneous function $g_l : \mathbb{R}^3\backslash\{0\} \to \mathbb{C}$, namely by

$$g_l(x) := \|x\|^l g\left(\frac{x}{\|x\|}\right), \qquad x \neq 0.$$

Conversely, any continuous function $g : \mathbb{R}^3 \to \mathbb{C}$ restricts to a function on S^2, and if g is homogeneous of some degree $l \in \mathbb{N}_0$, then the restriction map is injective. We may therefore think of homogeneous polynomials either as homogeneous polynomials on $\mathbb{R}^3$, or as polynomials on S^2. In particular, we may view V_l as a subspace of $L^2(S^2)$ by restriction.

The *spherical harmonics* are the (properly normalized) restrictions of homogeneous harmonic polynomials to the sphere.

Example 3.52. Let us give some examples of harmonic homogeneous polynomials in order to convince ourselves that many of them exist:

- Any constant polynomial is harmonic and homogeneous of degree $l = 0$,
- In degree $l = 1$, the polynomials x_1, x_2, x_3 are harmonic,
- In degree $l = 2$, the polynomials $x_k x_l$ (with $k \neq l$), $x_1^2 - x_2^2$, and $x_1^2 + x_2^2 - 2x_3^2$ are harmonic,
- In degree $l \in \mathbb{N}_0$, the two polynomials $\mathrm{Re}(x_1 + ix_2)^l$ and $\mathrm{Im}(x_1 + ix_2)^l$ are

homogeneous (clear) and harmonic. The latter property follows from

$$\begin{aligned}\Delta(x_1+ix_2)^l &= l\partial_1(x_1+ix_2)^{l-1} + il\partial_2(x_1+ix_2)^{l-1}\\ &= l(l-1)(x_1+ix_2)^{l-1} - l(l-1)(x_1+ix_2)^{l-2}\\ &= 0.\end{aligned}$$

The dimensions of the spaces Pol_l (we drop the $\mathbb{R}^3$ here and in the following whenever we consider polynomials of *three* variables) and $V_l \subset \mathrm{Pol}_l$ can be computed as follows.

Lemma 3.53. *Let $l \in \mathbb{N}_0$. Then*

$$\dim \mathrm{Pol}_l = \frac{1}{2}(l+1)(l+2), \qquad \dim V_l = 2l+1. \tag{3.115}$$

Proof. The space Pol_l is spanned by the monomials $x_1^{l_1}x_2^{l_2}x_3^{l_3}$ with $l_1, l_2, l_3 \in \mathbb{N}_0$ satisfying $l_1+l_2+l_3 = l$. For fixed $l_3 \in \{0, \dots, l\}$, there are $l-l_3+1$ many possibilities for choosing l_1, l_2. Thus

$$\dim \mathrm{Pol}_l = \sum_{l_3=0}^{l}(l+1-l_3) = (l+1)^2 - \frac{1}{2}(l+1)l = \frac{1}{2}(l+1)(l+2). \tag{3.116}$$

Let $p \in \mathrm{Pol}_l$, and write it sorted by powers of x_3, i.e.

$$p_3(x) = \sum_{k=0}^{l} x_3^k p_k(x_1, x_2),$$

with p_k polynomials of two variables (homogeneous of degree $l-k$). Then p is harmonic if and only if

$$\begin{aligned}0 = \Delta p(x) &= \sum_{k=0}^{l} x_3^k(\partial_1^2+\partial_2^2)p_k(x_1,x_2) + \sum_{k=2}^{l} k(k-1)x_3^{k-2}p_k(x_1,x_2)\\ &= \sum_{k=0}^{l} x_3^k(\partial_1^2+\partial_2^2)p_k(x_1,x_2) + \sum_{k=0}^{l-2}(k+1)(k+2)x_3^k p_{k+2}(x_1,x_2)\end{aligned}$$

vanishes. Comparing coefficients, this is equivalent to the recursion relation

$$p_{k+2} = -(k+1)^{-1}(k+2)^{-1}(\partial_1^2+\partial_2^2)p_k.$$

Hence for even (odd) k, the polynomial p_k is uniquely determined by p_0 (by p_1). Conversely, any choice of $p_0 \in \mathrm{Pol}_l(\mathbb{R}^2)$ and $p_1 \in \mathrm{Pol}_{l-1}(\mathbb{R}^2)$ uniquely determines a harmonic polynomial in V_l by the recursion relation.

As $\dim \mathrm{Pol}_l(\mathbb{R}^2) = l+1$ and $\dim \mathrm{Pol}_{l-1}(\mathbb{R}^2)$ (same argument as for the first dimension formula), we get $\dim V_l = \dim \mathrm{Pol}_l(\mathbb{R}^2) + \dim \mathrm{Pol}_{l-1}(\mathbb{R}^2) = 2l+1$. □

We now explain how the spaces V_l are related to rotations and angular momentum.

Lemma 3.54. *Consider the representation* $\rho : \mathrm{SO}(3) \to \mathrm{Aut}\,\mathrm{Pol}(\mathbb{R}^3)$, $(\rho(R)p)(x) := p(R^{-1}x)$. *Then* V_l *is an invariant subspace for* ρ, *i.e.* $p \in V_l \implies \rho(R)p \in V_l$ *for all* $l \in \mathbb{N}_0$, $R \in \mathrm{SO}(3)$.

Proof. It is clear that $\rho(R)$ maps polynomials to polynomials. It preserves Pol_l, $l \in \mathbb{N}_0$, because

$$(\rho(R)p)(\lambda x) = p(R^{-1}(\lambda x)) = p(\lambda R^{-1}x) = \lambda^l(\rho(R)p)(x).$$

Furthermore, Δ commutes with $\rho(R)$ (we had seen this already in Example 3.31 in a similar setting). Hence $\Delta p = 0$ implies $\Delta\rho(R)p = \rho(R)\Delta p = 0$, which finishes the proof. □

By restricting to the finite-dimensional space V_l, we have a finite-dimensional representation

$$\rho_l : \mathrm{SO}(3) \to \mathrm{Aut}\, V_l, \tag{3.117}$$

just as in matrix mechanics (Section 2.4). The generators of the one-parameter groups of rotations around the three coordinate axis are

$$L_1 = -ix_2\partial_3 + ix_3\partial_2, \qquad L_2 = -ix_3\partial_1 + ix_1\partial_3, \qquad L_3 = -ix_1\partial_2 + ix_2\partial_1, \tag{3.118}$$

considered as maps $V_l \to V_l$. This follows by explicit differentiation as in Exercise 3.11. The so defined maps satisfy the familar angular momentum exchange relations, i.e. they form angular momentum operators in the sense of Definition 2.37, with total angular momentum $L^2 := L_1^2 + L_2^2 + L_3^2$. We also define $L_\pm := L_1 \pm iL_2$ as in Lemma 2.39.

Proposition 3.55. *Define*

$$K := x_1\partial_1 + x_2\partial_2 + x_3\partial_3 \tag{3.119}$$

as a linear map $\mathrm{Pol}(\mathbb{R}^3) \to \mathrm{Pol}(\mathbb{R}^3)$, *where* x_j *means multiplication by* x_j, $j = 1, 2, 3$.

a) $L^2 = K(K+1) - \|x\|^2\Delta$ *as maps* $\mathrm{Pol} \to \mathrm{Pol}$,

b) For every $l \in \mathbb{N}_0$, *we have* $L^2|_{V_l} = l(l+1)\,\mathrm{id}_{V_l}$.

c) The system of angular momentum operators L_1, L_2, L_3 *on* V_l *is irreducible.*

d) As subspaces $V_l \subset L^2(S^2)$, *we have*

$$L^2(S^2) = \bigoplus_{l=0}^{\infty} V_l. \tag{3.120}$$

Proof. a) We set $W_{ij} := x_i\partial_j - x_j\partial_i$, so that $L_1 = -iW_{23}$, $L_2 = iW_{13}$, $L_3 = -iW_{12}$. Then

$$\begin{aligned} L^2 &= L_1^2 + L_2^2 + L_3^2 \\ &= -\sum_{1\le i<j\le 3} W_{ij}^2 \\ &= -\sum_{1\le i<j\le 3} (x_i^2\partial_j^2 + x_j^2\partial_i^2 - 2x_ix_j\partial_i\partial_j - x_i\partial_i - x_j\partial_j). \end{aligned}$$

The sum over the second order part simplifies as follows: Since the summand is symmetric in i, j, and vanishes for $i = j$, we have

$$\begin{aligned} -\sum_{1\le i<j\le 3}(x_i^2\partial_j^2 + x_j^2\partial_i^2 - 2x_ix_j\partial_i\partial_j) &= -\frac{1}{2}\sum_{i,j}(x_i^2\partial_j^2 + x_j^2\partial_i^2 - 2x_ix_j\partial_i\partial_j) \\ &= -\|x\|^2\Delta + \sum_{i,j} x_ix_j\partial_i\partial_j. \end{aligned}$$

The sum over the first order part is

$$\sum_{1\le i<j\le 3}(x_i\partial_i + x_j\partial_j) = 2K,$$

which yields $L^2 = -\|x\|^2\Delta + \sum_{i,j} x_ix_j\partial_i\partial_j + 2K$. But

$$\begin{aligned} K^2 &= \left(\sum_i x_i\partial_i\right)^2 \\ &= \sum_{i,j} x_i\partial_i x_j\partial_j \\ &= \sum_{i,j} x_ix_j\partial_i\partial_j + \sum_{i,j} x_i[\partial_i, x_j]\partial_j \\ &= \sum_{i,j} x_ix_j\partial_i\partial_j + K. \end{aligned}$$

Inserting this into the previous result yields $L^2 = -\|x\|^2\Delta + K(K+1)$.

b) On the monomial $p(x) = x_1^{l_1}x_2^{l_2}x_3^{l_3}$, the maps $K_j := x_j\partial_j$ act according to $K_jp = l_jp$. Hence $Kp = (l_1 + l_2 + l_3)p$, and we conclude that $K|_{\mathrm{Pol}_l} = l\,\mathrm{id}_{\mathrm{Pol}_l}$. As Δ vanishes on harmonic polynomials by definition, the claim follows from a).

c) We now consider the polynomial

$$p_l(x) := (x_1 + ix_2)^l, \tag{3.121}$$

and note

$$\begin{aligned} L_3p_l &= (-ix_1\partial_2 + ix_2\partial_1)(x_1 + ix_2)^l = l\cdot p_l, \\ L_+p_l &= (ix_3\partial_2 + x_3\partial_1)(x_1 + ix_2)^l = 0. \end{aligned}$$

As in the proof of Theorem 2.40 (which is based on the commutation relations recorded in Lemma 2.39), this implies that the polynomial $L_-^kp_l \in V_l$ satisfies $L_3L_-^kp_l = (l-k)p_l$, and $L_-^kp_l \neq 0$ for $0 \le k \le 2l+1$. As these vectors are eigenvectors of L_3, with different eigenvalues, they are linearly independent.

Now suppose that $W_l \subset V_l$ is a non-zero invariant subspace. As L_3 is diagonalizable on W_l, this space contains an eigenvector, i.e. one of the $L_-^kp_l$. But then acting with $L_\pm$ shows that this space contains all $(2l+1)$ of them, which span the whole space V_l because $\dim V_l = 2l+1$.

d) We have to show two things: (i) For $l \neq l'$, the spaces V_l and $V_{l'}$ are orthogonal as subspaces of $L^2(S^2)$, and (ii) the span of all V_l, $l \in \mathbb{N}_0$, is dense in $L^2(S^2)$.

Point (i) follows because V_l is the $l(l+1)$-eigenspace of the symmetric operator L^2, and eigenspaces for different eigenvalues are orthogonal. For point (ii), one first checks that

$$\mathrm{Pol}_l = V_l \oplus \|x\|^2 \,\mathrm{Pol}_{l-2}, \tag{3.122}$$

i.e. any polynomial p that is homogeneous of degree l can be written uniquely as $p = h + \|x\|^2 q$ with $h \in V_l$ harmonic and $q \in \mathrm{Pol}_{l-2}$. Proving (3.122) is left as an exercise. Once (3.122) is established, one can use it iteratively on the homogeneous components of an arbitrary polynomial. As the factors $\|x\|^2$ become trivial in restriction to S^2, this shows that given any polynomial p, there exists a (not necessarily homogeneous) harmonic polynomial q such that $p|_{S^2} = q|_{S^2}$. As $\bigoplus_{l=0}^{\infty} V_l$ spans all homogeneity degrees l, and as polynomials are dense in $L^2(S^2)$ by a version of Weierstraß' Theorem, this finishes the proof. □

As a representation of $\mathrm{SO}(3)$, the canonical representation U (3.111) decomposes as

$$L^2(\mathbb{R}^3) \cong \bigoplus_{l=0}^{\infty} (V_l \otimes L^2(\mathbb{R}_+, r^2\,dr)) \tag{3.123}$$

$$U \cong \bigoplus_{l=0}^{\infty} (U_l \otimes \mathrm{id}_{L^2(\mathbb{R}_+, r^2\,dr)}), \tag{3.124}$$

where U_l is the irreducible spin-l representation. Note that only integer spin values $l \in \mathbb{N}_0$ appear here. An operator H commuting with U will therefore only act non-trivially on the radial space $L^2(\mathbb{R}_+, r^2\,dr)$.

The following lemma is a preparation for evaluating Hamiltonians on the radial space.

Lemma 3.56. *Let $p \in V_l$ be a non-zero harmonic polynomial homogeneous of degree $l \in \mathbb{N}_0$, and $f : (0,\infty) \to \mathbb{C}$ measurable.*

a) The function

$$\psi : \mathbb{R}^3 \to \mathbb{C}, \qquad \psi(x) := p(x) f(\|x\|) \tag{3.125}$$

lies in $L^2(\mathbb{R}^3)$ if and only if

$$\int_0^\infty |f(r)|^2\, r^{2l+2}\, dr < \infty. \tag{3.126}$$

b) For smooth f, we have

$$(\Delta\psi)(x) = p(x)\left[f''(\|x\|) + \frac{2(l+1)}{\|x\|} f'(\|x\|)\right]. \tag{3.127}$$

c) The function ψ lies in $\mathrm{dom}\, P^2$ if and only if[a] $f, f' \in \mathrm{AC}(\mathbb{R}_+)$, $f \in L^2(\mathbb{R}_+, r^{2l+2}\,dr)$, and

$$\int_0^\infty \left| f''(r) + \frac{2(l+1)}{r} f'(r) \right|^2 r^{2l+2}\, dr < \infty. \tag{3.128}$$

For $l = 0$, one must in addition require $\lim_{r\to 0} r f(r) = 0$.

[a] For an interval I, the space $\mathrm{AC}(I)$ of absolutely continuous functions $I \to \mathbb{C}$ consists of all continuous function that can be written as the integral of some locally integrable function. Functions in $\mathrm{AC}(I)$ are almost everywhere differentiable.

Proof. a) The first claim is a direct calculation: Setting $c := \|p|_{S^2}\|_{L^2(S^2)} > 0$, we have

$$\begin{aligned}\|\psi\|^2_{L^2(\mathbb{R}^3)} &= \int_{\mathbb{R}^3} |p(x)|^2 |f(\|x\|)|^2 \, dx \\ &= \int_{\mathbb{R}^3} \|x\|^{2l} \left| p\left(\frac{x}{\|x\|}\right)\right|^2 |f(\|x\|)|^2 \, dx \\ &= c \int_0^\infty r^{2l+2} |f(r)|^2 \, dr,\end{aligned}$$

which shows that $\psi \in L^2(\mathbb{R}^3)$ is equivalent to (3.126).

b) To evaluate the Laplace operator on ψ, we first note that by the chain rule,

$$\partial_j f(\|x\|) = \frac{\partial \|x\|}{\partial x_j} f'(\|x\|) = \frac{x_j}{\|x\|} f'(\|x\|),$$

and

$$\langle x, \nabla p(x)\rangle = l p(x). \tag{3.129}$$

To prove the latter equality, note that differentiating $p(\lambda x) = \lambda^l p(x)$ w.r.t. λ at $\lambda = 1$ gives $\langle x, \nabla p(x)\rangle$ for the left hand side and $l p(x)$ for the right hand side.

With this preparation, we calculate with the Leibniz rule

$$\begin{aligned}\Delta(p(x) f(\|x\|)) &= (\Delta p)(x) + 2 \sum_{j=1}^3 \left[\partial_j p(x)\, \partial_j f(\|x\|) + \partial_j^2 f(\|x\|)\right] \\ &= 2 \left\langle \nabla p(x), \frac{x}{\|x\|} f(\|x\|)\right\rangle + p(x) \sum_{j=1}^3 \partial_j \left(\frac{x_j}{\|x\|} f'(\|x\|)\right) \\ &= \frac{2l}{\|x\|} p(x) f(\|x\|) + \left(\frac{3p(x)}{\|x\|} - \frac{p(x)}{\|x\|}\right) f'(\|x\|) + p(x) f''(\|x\|) \\ &= p(x)\left(f''(\|x\|) + \frac{2(l+1)}{\|x\|} f'(\|x\|)\right).\end{aligned}$$

c) The claims in this part can be verified by checking when the Fourier transform $\tilde\psi$ of ψ satisfies $\int_{\mathbb{R}^3} \|p\|^4 |\tilde\psi(p)|^2 dp < \infty$. We do not give the details of this computation here. □

To understand why the extra limit condition is required for $l = 0$ in c), consider $\chi \in C_c^\infty(\mathbb{R}_+)$ with $\chi(r) = 1$ for $r \le 1$, and $f(r) := \chi(r)/r$. Then all conditions listed in c), except the limit condition, hold. One may check that in terms of distributions, $P^2 f = 4\pi\delta_0$, the delta distribution at $x = 0$. As this does not lie in $L^2(\mathbb{R}^3)$, we see that an additional condition is necessary. For $l > 0$, this does not happen.

3.5 Hamiltonians with central potential

In this section, we consider a potential $V : \mathbb{R}^3 \to \mathbb{R}$ such that:

- V is rotationally symmetric, i.e. $V(x) = v(\|x\|)$ for some $v : \mathbb{R}_+ \to \mathbb{R}$,
- $V = V_\infty + V_2$ with $V_\infty \in L^\infty(\mathbb{R}^3)$ and $V_2 \in L^2(\mathbb{R}^3)$ both rotationally symmetric, i.e. $V_\infty(v) = v_\infty(\|x\|)$ and $V_2(x) = v_2(\|x\|)$ for some $v_\infty \in L^\infty(\mathbb{R}_+)$ and $v_2 : \mathbb{R}_+ \to \mathbb{R}$ satisfying

$$\int_0^\infty r^2 |v_2(r)|^2 \, dr < \infty. \tag{3.130}$$

Our most important example is the Coulomb potential, given by

$$v(r) = -\frac{Z}{r}, \qquad Z > 0. \tag{3.131}$$

See Example 3.38 for the proof that this potential satisfies the above assumptions.

Atomic numbers. The Hamiltonian with potential (3.131) models the electrostatic repulsion between a positively and a negatively charged particle, for example a proton and an electron. In atomic units, where the electron mass is $m = 1$, Planck's constant is $\hbar = 1$, and the electron charge is also $e = 1$, this corresponds to choosing $Z = 1$ in (3.131).

For a larger nucleus consisting of $Z \in \mathbb{N}$ protons (and some neutrons), the Coulomb potential models the force that this nucleus exerts on a single electron. The parameter $Z \in \mathbb{N}$ is called the *atomic number*. The symbol Z derives from the German "Zahl".

Given any such potential V, we consider the Hamiltonian

$$H := P^2 + V(X), \qquad (H\psi)(x) = -(\Delta\psi)(x) + v(\|x\|)\psi(x). \tag{3.132}$$

Recall that H is essentially selfadjoint on $\mathcal{S}(\mathbb{R}^3)$ and selfadjoint on $\operatorname{dom} H = \operatorname{dom} P^2$. This was proven in Corollary 3.37 by an application of the Kato-Rellich Theorem 3.36, namely $V(X)$ is relatively bounded w.r.t. P^2, with relative bound less than 1.

We now have a closer look at H and begin with its spectrum. The arguments given below apply to more general situations and do not yet use the rotational symmetry of H.

Proposition 3.57 (Spectra of Hamiltonians). *Let H be a Hamiltonian of the form (3.132).*

a) $\sigma(H)$ is unbounded above.

b) If $\sup_{\|x\| \le r} |V(x)| \to 0$ for $r \to \infty$, we have $[0, \infty) \subset \sigma(H)$.

c) $\sigma(H)$ is bounded below, i.e. $\sigma(H) \subset [-\lambda_0, \infty)$ for some $\lambda_0 \in \mathbb{R}$.

Proof. a) Let $\psi \in C_c^\infty(\mathbb{R}^3) \subset \operatorname{dom} H$, and consider $\psi_n(x) := e^{in\langle k, x\rangle}\psi(x)$, where $k \in \mathbb{R}^3 \setminus \{0\}$ is a fixed vector, and $n \in \mathbb{N}$. Then $\tilde{\psi}_n(p) = \tilde{\psi}(p - nk)$, and

$$\begin{aligned}
\langle \psi_n, H\psi_n \rangle &= \langle \psi_n, P^2\psi_n \rangle + \langle \psi_n, V(X)\psi_n \rangle \\
&= \int_{\mathbb{R}^3} \|p\|^2 |\tilde{\psi}(p - nk)|^2 \, dp + \int_{\mathbb{R}^3} V(x)|\psi_n(x)|^2 \, dx \\
&= \int_{\mathbb{R}^3} \|p + nk\|^2 |\tilde{\psi}(p)|^2 \, dp + \int_{\mathbb{R}^3} V(x)|\psi(x)|^2 \, dx.
\end{aligned}$$

Because of $\|p+nk\|^2 = \|p\|^2 + 2n\langle k,p\rangle + n^2\|k\|^2$, the first integral goes to $+\infty$ as $n \to \infty$, whereas the second integral does not depend on n. If $\sigma(H)$ was bounded above, say $\sigma(H) = \operatorname{supp} E^H \subset (-\infty, C]$ for some $C > 0$, then we would have

$$\langle \psi, H\psi\rangle = \int_{(-\infty,C]} \lambda \, d\mu_\psi^H(\lambda) \leq C\|\psi\|^2$$

for all $\psi \in \operatorname{dom} H$; a contradiction.

This argument used the fact that by multiplying with the phases $e^{in\langle k,x\rangle}$, we can arbitrarily increase the kinetic energy while keeping the potential energy constant. We now give a more refined argument based on the same idea, making use of the additional assumption $V(x) \to 0$ uniformly for $\|x\| \to \infty$ (as in Prop. 3.33).

b) Consider $\lambda \geq 0$ and pick $k \in \mathbb{R}^3$ with $\|k\|^2 = \lambda$. Pick a smooth function $\chi \in C_c^\infty(\mathbb{R}^3)$, $\|\chi\|_2 = 1$, with support in the unit ball $\mathbb{B} = \{x \,:\, \|x\| \leq 1\}$, and consider the sequence

$$\psi_n(x) := n^{-3/2}\chi\left(\frac{x - n^2 e_1}{n}\right) e^{i\langle k,x\rangle}.$$

We note that $\|\psi_n\|_2 = 1$ for all $n \in \mathbb{N}$, and that the support of ψ_n is contained in the ball $\mathbb{B}_n := B_n(n^2 e_1)$ of radius n and center $n^2 e_1$. This implies

$$\|V\psi_n\|_2^2 = \int_{\mathbb{B}_n} |V(x)|^2 |\psi_n(x)|^2 \, dx \to 0 \quad \text{as } n \to \infty.$$

For the kinetic energy, we have $\|(P^2 - \lambda)\psi_n\|^2 \to 0$ as $n \to \infty$. The argument for that is as in Prop. 3.33: Taking derivatives by the product rule, the term in which $P^2 = -\Delta$ acts on $e^{i\langle k,x\rangle}$ produces a factors $\|k\|^2 = \lambda$, canceling the explicit $-\lambda\psi_n$. Terms in which derivatives act on χ produce additional factors of n^{-1} and therefore vanish as $n \to \infty$.

These arguments show $\|(H - \lambda)\psi_n\| \to 0$, $\|\psi_n\| = 1$, and imply $\lambda \in \sigma(H)$.

c) The following argument works in much more generality; it only relies on the relative bound $\|V(X)\psi\| \leq a\|P^2\psi\| + b\|\psi\|$ (with fixed $a, b > 0$, $a < 1$, for all $\psi \in \operatorname{dom} P^2$). As P^2 is positive ($\sigma(P^2) = [0,\infty)$), $P^2 + \lambda$ is invertible for every $\lambda > 0$. Hence

$$\|V(X)(P^2+\lambda)^{-1}\psi\| \leq a\|P^2(P^2+\lambda)^{-1}\psi\| + b\|(P^2+\lambda)^{-1}\psi\| \leq \left(a + \frac{b}{\lambda}\right)\|\psi\|,$$

and by choosing λ large enough, we get $\|V(X)(P^2+\lambda)^{-1}\| < 1$. This implies that

$$H + \lambda = V(X) + P^2 + \lambda = (1 + V(X)(P^2+\lambda)^{-1})(P^2+\lambda)$$

is invertible (the first bracket is invertible by a geometric series), i.e. $-\lambda \in \rho(H)$ for λ large enough. This shows that $\sigma(H)$ is bounded below. □

Let us emphasize the importance of the lower bound c). This bound is remarkable because the potentials in our class may be *unbounded below*, like the Coulomb potential. While the possible values of the classical Coulomb Hamiltonian function

$$H^{\mathrm{cl}}(x,p) = \|p\|^2 - \frac{Z}{\|x\|} \tag{3.133}$$

are unbounded below (for $Z > 0$), namely $H^{\mathrm{cl}}(\mathbb{R}^2) = \mathbb{R}$, the possible values of the quantum Coulomb Hamiltonian operator are bounded below. This is already a strong indication that the quantum mechanical model might improve upon the classical one, because the classical model predicts unstable behaviour (the electron crashes into the proton) because of the unbounded below energy values.

In comparison to Prop. 3.33 for compactly supported potentials, we here have not yet a statement that the point spectrum is negative. Several such statements under different hypothesis exist. Due to the singular and long range nature of the Coulomb potential, not all of them apply to it. We therefore provide an argument that is designed specifically for the Coulomb potential.

Proposition 3.58 (Virial theorem and negative point spectrum for Coulomb pot.). *Consider the Hamiltonian $H = P^2 + V(X)$ with Coulomb potential $V(x) = -\frac{Z}{\|x\|}$. Then the* Virial theorem *holds, namely*

$$\langle \psi, P^2 \psi \rangle = -\frac{1}{2} \langle \psi, V(X)\psi \rangle \tag{3.134}$$

for any eigenvector ψ of H. As a consequence, all eigenvalues of H are negative,

$$\sigma_p(H) \subset (-\infty, 0). \tag{3.135}$$

Proof. We consider the dilation operators

$$D(s) : L^2(\mathbb{R}^3) \to L^2(\mathbb{R}^3), \qquad (D(s)\psi)(x) := e^{-3s/2}\psi(e^{-s}x).$$

It is easy to check that D is a strongly continuous unitary one-parameter group (exercise), and we have

$$D(s)HD(-s) = e^{2s}P^2 + e^{s}V(X), \qquad \mathrm{dom}(D(s)HD(-s)) = \mathrm{dom}\, H = \mathrm{dom}\, P^2.$$

Let λ be an eigenvalue of H, i.e. $H\psi = \lambda\psi$ for some $\psi \in \mathrm{dom}(H)\backslash\{0\}$. Then

$$\langle \psi, [D(s), H]\psi \rangle = \lambda \langle \psi, D(s)\psi \rangle - \lambda \langle \psi, D(s)\psi \rangle = 0.$$

Hence the limit $\lim_{s\to 0} \frac{1}{s}\langle \psi, [D(s), H]\psi \rangle$ exists and vanishes. Using the scaling behaviour of the kinetic and potential energy, this limit can also be computed as

$$\begin{aligned}
0 &= \lim_{s\to 0} \frac{1}{s}\langle \psi, [D(s), H]\psi \rangle \\
&= \lim_{s\to 0} \langle D(-s)\psi, \frac{1}{s}(H - D(-s)HD(s))\psi \rangle \\
&= \lim_{s\to 0} \langle D(-s)\psi, [\frac{1}{s}(P^2 - e^{-2s}P^2) + \frac{1}{s}(V(X) - e^{-s}V(X))]\psi \rangle \\
&= \langle \psi, [2P^2 + V(X)]\psi \rangle.
\end{aligned}$$

In the last step, we have realized the limit in the right entry of the scalar product as a derivative at $s = 0$, and used the continuity of D in the left entry.

This shows the virial theorem. As a simple consequence, the eigenvalue λ must be negative

$$\begin{aligned}\lambda\|\psi\|^2 &= \langle\psi, H\psi\rangle\\ &= \langle\psi, [P^2 + V(X)]\psi\rangle\\ &= \langle\psi, [2P^2 + V(X)]\psi\rangle - \langle\psi, P^2\psi\rangle\\ &= -\langle\psi, P^2\psi\rangle\\ &< 0.\end{aligned}$$

□

We now exploit the SO(3)-symmetry of H for general rotationally symmetric potential (in the class defined at the beginning of this section). Recall that the natural SO(3)-representation U on $L^2(\mathbb{R}^3)$, given by $(U(R)\psi)(x) = \psi(R^{-1}x)$, decomposes according to

$$L^2(\mathbb{R}^3) = \bigoplus_{l=0}^{\infty} [V_l \otimes L^2(\mathbb{R}_+, r^2\,dr)],$$
$$U(R) = \bigoplus_{l=0}^{\infty} [U_l(R) \otimes \mathrm{id}_{L^2(\mathbb{R}_+, r^2\,dr)}],$$

where U_l is the irreducible spin-l representation, V_l is the $(2l+1)$-dimensional space of spherical harmonics of angular momentum l (the restriction of the harmonic polynomials homogeneous of degree l to S^2), and we have

$$V_l = \mathrm{span}\{Y_{lm} \;:\; m \in \{-l, \dots, l\}\}, \tag{3.136}$$

where $Y_{ll}(x) = (x_1 + ix_2)^l$, and $Y_{l,l-k} = L_-^k Y_{ll}$, up to normalization.

Also recall from Lemma 3.56 c) the conditions on f such that $\psi(x) := p(x)f(\|x\|)$, with $p \in V_l$, lies in $\mathrm{dom}\, P^2$. We denote the domain of all these functions $\mathcal{D}_l \subset L^2(\mathbb{R}_+, r^{2l+2}\,dr)$.

Proposition 3.59 (The radial Hamiltonians). *Let H be a Hamiltonian of the form (3.132).*

a) Each space of the form $V_l \otimes L^2(\mathbb{R}_+, r^2\,dr)$ is invariant under the spectral projections of H.

b) Let $l \in \mathbb{N}_0$ and $f \in L^2(\mathbb{R}_+, r^{2l+2}\,dr)$. Then $\psi(x) := p(x)f(\|x\|)$ lies in $\mathrm{dom}\, H$ if and only if $f \in \mathcal{D}_l$.

c) For $f \in \mathcal{D}_l$, we have

$$(H\psi)(x) = p(x)(H_l f)(\|x\|), \tag{3.137}$$
$$(H_l f)(r) := -f''(r) - \frac{2(l+1)}{r} f'(r) + v(r)f(r), \qquad r > 0. \tag{3.138}$$

Proof. a) We know that the SO(3)-representation U commutes with the dynamics e^{itH}, $t \in \mathbb{R}$ (Example 3.31). For the following argument, we can use any bounded operator $B \in B(L^2(\mathbb{R}^3))$ commuting with U. For $l \in \mathbb{N}_0$, consider the orthogonal projection P_l onto $V_l \otimes L^2(\mathbb{R}_+, r^2\,dr)$, and $B_{kl} := P_k B P_l$.

Let $f, g \in L^2(\mathbb{R}_+, r^2\,dr)$ and define $B_{kl}^{f,g} : V_l \to V_k$ by

$$\langle\eta, B_{kl}^{f,g}\xi\rangle := \langle f \otimes \eta, B_{kl}(g \otimes \xi)\rangle, \qquad \eta \in V_k,\ \xi \in V_l. \tag{3.139}$$

These maps satisfy $B^{f,g}_{kl} U_l(R) = U_k(R) B^{f,g}_{kl}$ for every $R \in \mathrm{SO}(3)$. By Schur's Lemma, it follows that $B^{f,g}_{kl} = 0$ for $k \neq l$, i.e. $0 = \langle f \otimes \eta, B_{kl}(g \otimes \xi)\rangle$ for $k \neq l$. As f, g, η, ξ were arbitrary, we arrive at $B_{kl} = 0$, $k \neq l$.

For $k = l$, the operator $B^{f,g}_{ll}$ commutes with the irreducible representation U_l. This implies that $B^{f,g}_{ll}$ is a multiple of the identity, $B^{f,g}_{ll} = b_l(f,g)\,\mathrm{id}_{V_l}$. Thus $B_{ll} = \mathrm{id}_{V_l} \otimes B_l$ for some bounded operator $B_l : L^2(\mathbb{R}_+, r^2\,dr) \to L^2(\mathbb{R}_+, r^2\,dr)$, and we conclude that B can be written in the form $B = \bigoplus_l [\mathrm{id}_{V_l} \otimes B_l]$.

Applied to $B = e^{itH}$, it follows that the spaces $V_l \otimes L^2(\mathbb{R}_+, r^2\,dr)$ are invariant under e^{itH} and hence under the spectral projections of H.

b) By Kato-Rellich, we have $\operatorname{dom} H = \operatorname{dom} P^2$. Hence the claim follows immediately from Lemma 3.56.

c) For smooth ψ, this formula has been shown in Lemma 3.56 (note that V is rotationally symmetric and therefore only acts on the radial part). This formula extends to general $\psi \in \operatorname{dom} H = \operatorname{dom} P^2$. □

This proposition has used rotational symmetry to simplify the partial differential operator H in three variables to the l-dependent family H_l of ordinary differential operators, the so-called *radial Hamiltonians*.

For general V, this is however still too complicated to allow for explicit determination of the point spectrum. In the case of the Coulomb potential, however, we can say much more.

Theorem 3.60 (The point spectrum of the hydrogen Hamiltonian). *Let H be the Hamiltonian given by $H = P^2 + V(X)$ on $\operatorname{dom} P^2$, with $V(x) = -\frac{Z}{\|x\|}$, $Z > 0$, and let*

$$\lambda_N := -\frac{Z^2}{4(N+1)^2}, \qquad N \in \mathbb{N}_0. \tag{3.140}$$

a) $\sigma_p(H) = \{\lambda_N : N \in \mathbb{N}_0\}$.

b) An orthonormal basis of the λ_N-eigenspace is given by

$$\psi_{N,l,m}(x) = C_{N,m,l} Y_{lm}(x) Q^{2l+1}_{N-l}\left(\frac{Z\|x\|}{N+1}\right) e^{-\frac{Z\|x\|}{2(N+1)}}, \tag{3.141}$$

$$l \in \mathbb{N}_0, \quad m \in \mathbb{Z}, \qquad |m| \leq l \leq N, \tag{3.142}$$

where Y_{lm} are the spherical harmonics, Q^{2l+1}_{N-l} are the generalized Laguerre polynomials[a], and $C_{N,m,l}$ suitable normalization constants.

c) $\dim \ker (H - \lambda_N) = (N+1)^2$.

[a] These polynomials will be constructed in the proof.

Proof. a) According to the virial theorem, we may restrict our search for eigenvalues λ of H to $\lambda < 0$. By decomposing H into its angular and radial parts, we consider angular momentum $l \in \mathbb{N}_0$ and have to determine for which $\lambda < 0$ the eigenvalue equation $H_l f = \lambda f$ has a solution $f \in \mathcal{D}_l$. Explicitly,

$$(H_l f)(r) = -f''(r) - \frac{2(l+1)}{r} f'(r) - \frac{Z}{r} f(r) = \lambda f(r).$$

To simplify the equation, we set $\lambda = -\kappa^2$ and introduce the function g via $f(r) = e^{-\kappa r} g(2\kappa r)$, $r > 0$. In terms of g and the variable $s := 2\kappa r$, the eigenvalue equation takes the form of Kummer's equation

$$sg''(s) + [2(l+1) - s]g'(s) + \left[\frac{Z}{2\kappa} - (l+1)\right] g(s) = 0. \tag{3.143}$$

To find solutions of this equation, we first make the assumption that g is a power series, namely $g(s) = \sum_{k=0}^{\infty} c_k s^k$ (this assumption will be justified later). With the abbreviations $a := 2l+1$ and $b := \frac{Z}{2\kappa} - (l+1)$, the power series ansatz yields

$$\begin{aligned} 0 &= \sum_{k=2}^{\infty} c_k k(k-1) s^{k-1} + \sum_{k=1}^{\infty} c_k k(a+1-s) s^{k-1} + \sum_{k=0}^{\infty} c_k b s^k \\ &= \sum_{k=0}^{\infty} \left[c_{k+1} k(k+1) + (k+1)(a+1) c_{k+1} - k c_k + c_k b \right] s^k. \end{aligned}$$

By comparison of the coefficients, we find the recursion relation

$$c_{k+1} = \frac{k-b}{(k+1)(k+a+1)} \cdot c_k, \qquad k \in \mathbb{N}_0,$$

with the solution

$$c_k = c_0 \frac{1}{k!} \prod_{j=0}^{k-1} \frac{j-b}{j+a+1}.$$

We now distinguish the two cases $b \in \mathbb{N}_0$ and $b \notin \mathbb{N}_0$.

For $b = n \in \mathbb{N}_0$, we have $c_k = 0$ for $k > b$, and g is a polynomial. This polynomial is, up to a multiplicative constant, the generalized Laguerre polynomial Q_{N-l}^{2l+1}. In this case, it is easy to check that $f \in \mathcal{D}_l$, and hence $\psi \in \operatorname{dom} H$. Consequently

$$\lambda_{n+l} = -\frac{1}{(n+l+1)^2} \frac{Z^2}{4}, \qquad n, l \in \mathbb{N}_0, \tag{3.144}$$

are eigenvalues of H, with corresponding eigenfunctions

$$\psi(x) = p_l(x) Q_{N-l}^{2l+1}(2\sqrt{|\lambda_N|}\|x\|)\, e^{-\sqrt{|\lambda_N|}\|x\|}, \tag{3.145}$$

where $N := n + l$. This already shows that all λ_N are eigenvalues of H.

For $b \notin \mathbb{N}_0$, all coefficients c_k are non-zero, and $\frac{c_{k+1}}{c_k} \geq \frac{1/2}{k+1}$ for sufficient large k. This leads to an exponential recurrence relation, i.e. up to adding a polynomial, $g(r) \geq e^{r/2}$. But this function does not lie in $L^2(\mathbb{R}_+, r^2\, dr)$, so it does not lead to eigenvalues.

It remains to justify the power series ansatz for g. We have seen above that for any $\lambda < 0$, this ansatz produces a solution $s \mapsto g_1(s)$ of Kummer's equation (3.143). As a linear ODE of second order, it has two linearly independent solutions. One may check by explicit differentiation that

$$g_2(s) := g_1(s) \int_{s_0}^{s} \frac{e^{-\int_{s_0}^{s'} \left(\frac{2(l+1)}{s''} - 1\right) ds''}}{g_1(s')^2}\, ds' \tag{3.146}$$

is another solution of (3.143), and that every solution of (3.143) is a linear combination of g_1 and g_2 (use the Wronskian). Here $s_0, s_0' > 0$ are fixed reference points. One may then check that g_2 behaves as $g_2(s) \sim c\, s^{-2l-1}$ for $s \to 0$. This behaviour is in conflict with the domain conditions of Lemma 3.56. Hence the second solution does not contribute eigenvalues of H.

b), c) The eigenvalues of H are labelled by $N = n + l, n, l \in \mathbb{N}_0$. For fixed $N \in \mathbb{N}_0$, the (n,l)-pairs $(N,0), (N-1,1), \dots, (0,N)$ contribute to the λ_N-eigenspace. There are $N+1$ such pairs. At fixed l, the L_3-eigenvalue m may take the $2l+1$ different values $m = -l, \dots, l$, hence the dimension of the λ_N-eigenspace is $\sum_{l=0}^{N}(2l+1) = N(N+1)+(N+1) = (N+1)^2$.

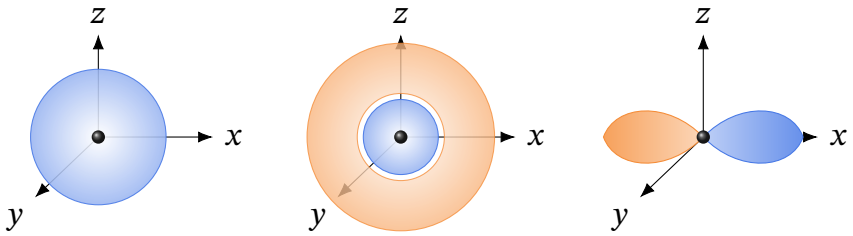
The formula for the eigenfunctions $\psi_{N,l,m}$ follows by inserting the solution of the radial equation. □

⚛ **Electron orbitals.** Let us discuss the structure of the point spectrum and eigenfunctions of H. At energy eigenvalue λ_N, $N \in \mathbb{N}_0$, angular momentum eigenfunctions with l (the eigenvalue of L^2 is $l(l+1)$) ranging over $l \in \{0, \dots, N\}$ and m (the eigenvalue of L_3) ranging over $m \in \{-l, \dots, l\}$ contribute. The angular momentum eigenvalues (l,m) fix the angular dependence of the eigenfunction $\psi_{N,l,m}$, and the radial dependence is then fixed as explained above.

For example, the lowest eigenvalue of H is $\lambda_0 = -\frac{Z^2}{4}$, the ground state energy. It is non-degenerate, i.e. there exists a unique eigenfunction (up to multiplicative constants) with this eigenvalue. It has $l = 0$ and $m = 0$, i.e. it is constant in the angles. Chemists call this an "s-orbital", more precisely the $1s$-orbital.

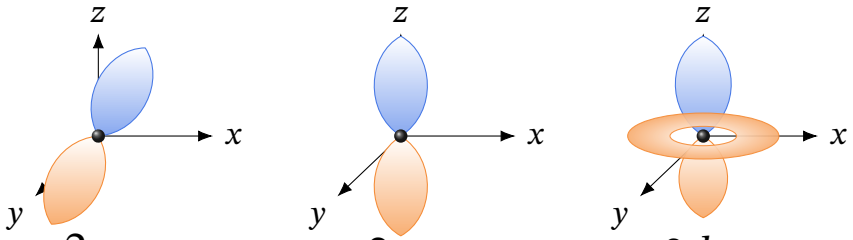
The first excited energy above the ground state energy is $\lambda_1 = -\frac{Z^2}{8}$. Its eigenspace is four-dimensional, with one s-orbital ($2s$), i.e. $l = m = 0$, and also solutions of angular momentum $l = 1$, and $m = -1, 0, 1$. Chemists call these "p-orbitals", more precisely $2p_x, 2p_y, 2p_z$ standing for $N = 1, l = 1, m = -1, 0, 1$. These historic names derive from atomic spectroscopy and stand for "sharp" ($s, l = 0$), "principal" ($p, l = 1$), "diffuse" (d, $l = 2$), "fundamental" (f,$l = 3$). After f, one proceeds alphabetically $g, h, \dots$.

The following pictures of the spherical harmonics depict the value of $|Y_{lm}(x)|$, $x \in S^2$, and give an impression of the shape of the orbitals. Together with the radial dependence, they determine the probability distribution for the position of the electron bound to the nucleus.



$1s$ $2s$ $2p_x$

$N = 0,\ l = 0,\ m = 0$ $N = 1,\ l = 0,\ m = 0$ $N = 0,\ l = 1$, real combination of $m = \pm 1$



$2p_y$ $2p_z$ $3d_{z^2}$

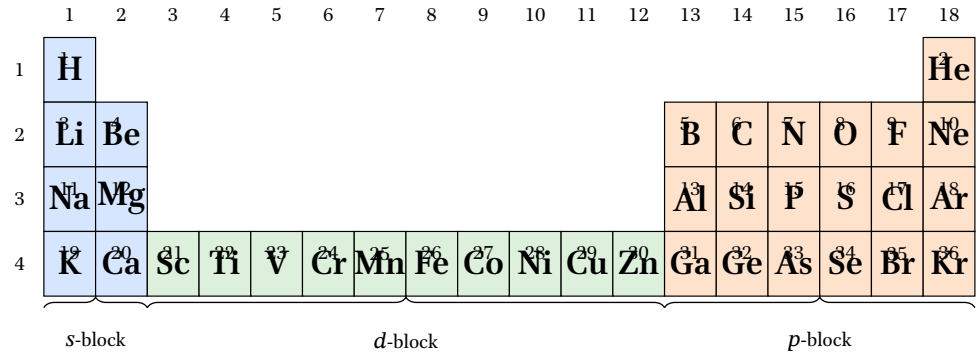
$N = 0,\ l = 1$, real combination of $m = \pm 1$ $N = 0,\ l = 1,\ m = 0$ $N = 0,\ l = 2,\ m = 0$

The determination of the bound states of the hydrogen atom and the corresponding absorption spectra were an early success of quantum mechanics. It led to good agreement with experimental observations, in contrast to classical physics.

⚛ **Atoms and the periodic table.** Larger atoms than hydrogen have $Z > 1$ protons (plus neutrons) in the nucleus, and Z electrons in the shell (for neutral atoms). For example, Helium consists of $Z = 2$ protons and one (or two, depending on isotope) neutrons in its nucleus and $Z = 2$ electrons.

The simplest idea is to treat the atom as one positively charged particle of charge Ze (with e the electron charge), and Z electrons around them. Neglecting all electron-electron interaction, one might think that the ground state of such an atom is given by filling the shells $N = 0, 1, 2, \ldots$ from the bottom. This rough idea suggests that chemical properties of atoms may be similar when the top shell is completely filled, has just one electron, etc.

As the eigenspace corresponding to λ_N has dimension N^2, one might expect a periodic table with periods $1, 4, 9, \ldots$. This does however not match with what is actually observed in nature. The picture becomes much better when including spin, namely working on $L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$ of two-component wave functions to model the spin $\frac{1}{2}$ of the electrons. This suggests periods $2N^2$, i.e. $2, 8, 18, \ldots$, and already partially agrees with the periodic table, whose first four rows have lengths $2, 8, 8, 18$:



This discrepancy is essentially due to the electron-electron repulsion that we have ignored here.

This discussion brings us to the end of the lecture. Needless to say, there are many directions in which mathematical quantum physics extends beyond what has been covered here. Let us mention at least three such directions:

- **Many-body quantum mechanics.** In many-body quantum mechanics, one studies systems consisting of many quantum particles, extending the formalism of composite systems from matrix mechanics to quantum mechanics. This involves in particular many-body Schrödinger operators which are often best studied with approximations (e.g. mean-field limits) and other perturbation methods from analysis. Applications arise in particular in the description of gases of quantum particles and phenomena like Bose-Einstein condensation.
- **Quantum field theory.** In quantum field theory, the rules of special relativity are incorporated into quantum physics, which leads to a structure of the observable algebra in which localization in different regions of spacetime have to be taken into account in order to comply with causality. Instead of the Hamiltonian, local quantum fields, their correlation functions, or the observable algebras they generate, are the main object of interest, and studied with tools from functional analysis. Applications arise in particular in the description of elementary particles.
- **Quantum information theory.** In quantum information theory, one studies information-theoretic questions about quantum states, extending our discussion of entanglement. Central objects are quantum channels, which describe the physically allowed transformations of states, and measurements, which convert quantum information into classical data. Typical questions concern the distinguishability of states, quantum communication and cryptography, and the use of entanglement as a resource. Mathematically, the subject draws on matrix analysis, operator inequalities, convex optimization and probability theory.

3.6 Exercises and Problems

Exercise 3.1 (Polynomial differential operators). Consider the set $\mathrm{DiffPol}(\mathbb{R}^d)$ of polynomial differential operators on $\mathbb{R}^d$. Show the following:

a) Every $D \in \mathrm{DiffPol}(\mathbb{R}^d)$ is a well-defined linear map $\mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$.

b) DiffPol($\mathbb{R}^d$) is a unital *-algebra with the *-operation given by

$$\left(\sum_\alpha M_{Q_\alpha}\partial^\alpha\right)^* := \sum_\alpha (-1)^{|\alpha|}\partial^\alpha \circ M_{\overline{Q_\alpha}}. \tag{3.147}$$

c) For $\psi, \varphi \in \mathcal{S}(\mathbb{R}^d)$, one has

$$\langle \psi, D\varphi\rangle = \langle D^*\psi, \varphi\rangle, \tag{3.148}$$

where $\langle\cdot,\cdot\rangle$ is the L^2-scalar product.

Exercise 3.2 (The Fourier transform on $\mathcal{S}(\mathbb{R}^d)$). Prove the following:

a) Let $f \in \mathcal{S}(\mathbb{R}^d)$. Then $\tilde{f} \in \mathcal{S}(\mathbb{R}^d)$.

b) The Fourier transform on $\mathcal{S}(\mathbb{R}^d)$, i.e. the map

$$\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d), \quad f \mapsto \tilde{f}, \tag{3.149}$$

satisfies for any $k \in \{1, \dots, d\}$,

$$\mathcal{F}P_k = X_k\mathcal{F}, \qquad \mathcal{F}^{-1}X_k = P_k\mathcal{F}^{-1}, \tag{3.150}$$

where X_k, P_k are the Schrödinger position/momentum operators.

c) The Fourier transform on $\mathcal{S}(\mathbb{R}^d)$ is a linear bijection. Its inverse (the *inverse Fourier transform*) is given by

$$\mathcal{F}^{-1} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d), \tag{3.151}$$

$$(\mathcal{F}^{-1}f)(x) = (2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{i\langle p,x\rangle} f(p)\,dp = \tilde{f}(-x). \tag{3.152}$$

Exercise 3.3 (Expectation values and uncertainties for position and momentum). Considering vector states ω_ψ on DiffPol($\mathbb{R}$) given by vectors $\psi \in \mathcal{S}(\mathbb{R})$, $\|\psi\|_2 = 1$ (in dimension $d = 1$ for simplicity), show the following:

a) Given any $x_0 \in \mathbb{R}$ and any $\delta_X > 0$, there exists ψ as above such that $\omega_\psi(X) = x_0$ and $\Delta_\psi(X) = \delta_X$. Show that for no ψ, one gets $\Delta_\psi(X) = 0$.

b) Given any $p_0 \in \mathbb{R}$ and any $\delta_P > 0$, there exists ψ as above such that $\omega_\psi(P) = p_0$ and $\Delta_\psi(P) = \delta_P$. Show that for no ψ, one gets $\Delta_\psi(P) = 0$.

c) Given arbitrary $x_0, p_0 \in \mathbb{R}$ and $\delta_X, \delta_P > 0$, does there exist ψ as above satisfying the requirements in a) and b)?

Exercise 3.4 (Quantum harmonic oscillator). Consider the functions $\psi_0, \psi_1 \in \mathcal{S}(\mathbb{R})$ given by

$$\psi_0(x) = \pi^{-\frac{1}{4}}e^{-\frac{1}{2}x^2} \quad \text{and} \quad \psi_1(x) = \sqrt{2}\pi^{-\frac{1}{4}}xe^{-\frac{1}{2}x^2} \tag{3.153}$$

as well as $\phi_{c_0,c_1} := c_0\psi_0 + c_1\psi_1$ for $c_0, c_1 \in \mathbb{C}$. Let H be the Hamiltonian of the

harmonic oscillator

$$H := \frac{1}{2}P^2 + \frac{1}{2}X^2. \tag{3.154}$$

a) Show that ψ_0 and ψ_1 are normalized w.r.t. the L^2-scalar product.
b) Show that ψ_0 and ψ_1 are eigenvectors of H with eigenvalues

$$\lambda_0 = \tfrac{1}{2} \quad \text{and} \quad \lambda_1 = \tfrac{3}{2}. \tag{3.155}$$

c) Show that for $n = 0, 1$ it holds

$$\omega_{\psi_n}(X) = \omega_{\psi_n}(P) = 0 \quad \text{and} \quad \omega_{\psi_n}(X^2) = \omega_{\psi_n}(P^2) = n + \frac{1}{2}. \tag{3.156}$$

d) Consider $c_n = \frac{e^{i\lambda_n t}}{\sqrt{2}}$ for $t \in \mathbb{R}$ and show

$$\omega_{\phi_{c_0,c_1}}(X) = \frac{1}{\sqrt{2}}\cos t, \quad \omega_{\phi_{c_0,c_1}}(P) = \frac{1}{\sqrt{2}}\sin t, \tag{3.157}$$

$$\omega_{\phi_{c_0,c_1}}(X^2) = 1, \quad \omega_{\phi_{c_0,c_1}}(P^2) = 1. \tag{3.158}$$

Exercise 3.5 (Diagonal Operator). Consider a Hilbert space $\mathcal{H}$ with orthonormal basis $(e_n)_{n\in\mathbb{N}}$. Let $(A, \mathrm{dom}(A))$ be a densely defined unbounded operator with $\mathrm{dom}(A) = \mathrm{span}\{e_n \,:\, n \in \mathbb{N}\}$ and

$$Ae_n = \lambda_n e_n, \quad \lambda_n \in \mathbb{R}, n \in \mathbb{N}. \tag{3.159}$$

Show the following:
a) A is essentially selfadjoint.
b) The operator $(\overline{A}, \mathrm{dom}(\overline{A}))$ is given by

$$\mathrm{dom}(\overline{A}) = \{\sum_{n\in\mathbb{N}} c_n e_n \in \mathcal{H} \,:\, \sum_{n\in\mathbb{N}} |\lambda_n c_n|^2 < \infty\}, \tag{3.160}$$

$$\overline{A}(\sum_{n\in\mathbb{N}} c_n e_n) = \sum_{n\in\mathbb{N}} \lambda_n c_n e_n. \tag{3.161}$$

c) The spectrum of $\overline{A}$ is given by

$$\sigma(\overline{A}) = \overline{\{\lambda_n \,:\, n \in \mathbb{N}\}}. \tag{3.162}$$

For simplicity assume that $\lambda_n \neq \lambda_m$ for $n \neq m$. Show that the spectral measure $E^{\overline{A}}$ of $\overline{A}$ is given by

$$E^{\overline{A}}(S) = \sum_{\lambda_n \in S} P_{e_n}, \quad S \subset \sigma(\overline{A}) \text{ Borel set.} \tag{3.163}$$

Exercise 3.6 (Quantum Harmonic Oscillator 2). Consider on the Schwartz space $\mathcal{S}(\mathbb{R})$ the Hamiltonian of the harmonic oscillator

$$H := \frac{1}{2}P^2 + \frac{1}{2}X^2 = A^*A + \frac{1}{2}, \tag{3.164}$$

where $A = 2^{-\frac{1}{2}}(X + iP)$ and $A^* = 2^{-\frac{1}{2}}(X - iP)$ on $\mathcal{S}(\mathbb{R})$. Define for $n \in \mathbb{N}_0$ the functions $\psi_n \in \mathcal{S}(\mathbb{R})$ by

$$\psi_0(x) := e^{-\frac{x^2}{2}} \quad \text{and} \quad \psi_n := (A^*)^n\psi_0. \tag{3.165}$$

a) Show that

$$\psi_n \neq 0, \quad A\psi_n = n\psi_{n-1} \quad \text{and} \quad A\psi_0 = 0, \quad n \in \mathbb{N}_0. \tag{3.166}$$

b) Show that

$$H\psi_n = (n + \frac{1}{2})\psi_n, \quad n \in \mathbb{N}_0. \tag{3.167}$$

c) Denote the space of polynomials on $\mathbb{R}$ by $\mathrm{Pol}(\mathbb{R})$. Show that

$$\mathrm{span}\{\psi_n \ : \ n \in \mathbb{N}_0\} = \mathrm{Pol}(\mathbb{R}) \cdot \psi_0. \tag{3.168}$$

d) Define the function

$$f_z(x) := e^{izx}e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}, z \in \mathbb{C} \tag{3.169}$$

and show that $f_z \in L^2(\mathbb{R})$ for all $z \in \mathbb{C}$. Prove moreover that in L^2-norm

$$f_z = \lim_{N\to\infty} f_{z,N}, \quad \text{where} \quad f_{z,N}(x) = \sum_{n=0}^{N} \frac{(zx)^n}{n!}e^{-\frac{x^2}{2}}. \tag{3.170}$$

e) Show that if $g \in L^2(\mathbb{R})$ is orthogonal to $(\psi_n)_{n\in\mathbb{N}_0}$, then

$$\int_{\mathbb{R}} e^{ikx}g(x)e^{-\frac{x^2}{2}}\,dx = 0. \tag{3.171}$$

Conclude that $(\psi_n)_{n\in\mathbb{N}_0}$ is an orthogonal basis of $L^2(\mathbb{R})$.

f) Show that H defined on $\mathcal{S}(\mathbb{R})$ is essentially selfadjoint and compute its spectrum and spectral measure.

Exercise 3.7 (Pure Normal States of $\mathcal{B}(\mathcal{H})$). Show that a normal state $\omega \in \mathcal{S}_{\mathrm{no}}(\mathcal{H})$ is pure if and only if ω is a vector state.

Exercise 3.8 (Differentiable $\mathcal{H}$-valued functions). Consider a function $\psi : \mathbb{R} \to \mathcal{H}$ that is differentiable w.r.t. the norm of $\mathcal{H}$. Show that the function

$$f : \mathbb{R} \to [0, \infty), \quad f(t) := \|\psi(t)\|^2 \tag{3.172}$$

is differentiable with

$$\frac{d}{dt} f(t) = \langle \psi'(t), \psi(t) \rangle + \langle \psi(t), \psi'(t) \rangle. \tag{3.173}$$

Exercise 3.9 (Moments and Domains). Consider a selfadjoint operator $(A, \mathrm{dom}(A))$ and recall Definition 3.22 for the moments of μ_ψ^A for $\psi \in \mathcal{H}$.

a) Let $\psi \in \mathrm{dom}(A)$ with $\|\psi\| = 1$. Show that

$$\omega_\psi(A) = \langle \psi, A\psi \rangle, \quad \Delta_\psi(A) = \|A\psi - \omega_\psi(A)\|. \tag{3.174}$$

b) Find an example of a selfadjoint operator A and $\psi \in \mathcal{H}$ s.t. $\psi \notin \mathrm{dom}(A)$ but the first moment of μ_ψ^A exists.

Exercise 3.10 (Spectral Decomposition). Let $(\mathcal{H}, H)$ be a quantum-mechanical system and recall the definitions of $\mathcal{H}_{\mathrm{pp}}^H, \mathcal{H}_{\mathrm{ac}}^H, \mathcal{H}_{\mathrm{sc}}^H$, see Def. 3.40.

a) Show that

$$\mathcal{H}_{\mathrm{pp}}^H = \overline{\mathrm{span}\{E^H(\{\lambda\})\mathcal{H} \ : \ \lambda \in \mathbb{R}\}}.$$

b) Show that if $\psi \in \mathcal{H}_{\mathrm{xx}}^H$, then $\mu_{\phi,\psi}^H$ is xx for all $\phi \in \mathcal{H}$.

c) Show that the subspaces $\mathcal{H}_{\mathrm{pp}}^H, \mathcal{H}_{\mathrm{ac}}^H, \mathcal{H}_{\mathrm{sc}}^H$ are orthogonal to each other.

Exercise 3.11 (Angular Momentum Operators in Quantum Mechanics). Consider $\mathcal{S}(\mathbb{R}^3) \subset L^2(\mathbb{R}^3)$ and a Hamiltonian H that is essentially selfadjoint on $\mathcal{S}(\mathbb{R}^3)$ and rotationally symmetric, i.e. for $\psi \in \mathcal{S}(\mathbb{R}^3)$

$$HU(R)\psi = U(R)H\psi, \quad (U(R)\psi)(x) = \psi(R^{-1}x), \ R \in \mathrm{SO}(3). \tag{3.175}$$

a) Show that the selfadjoint generator L_1 of

$$(U_1(s)\psi)(x) := \psi(R_1(s)^{-1}x), \quad R_1(s) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(s) & \sin(+s) \\ 0 & \sin(-s) & \cos(s) \end{pmatrix} \tag{3.176}$$

satisfies $L_1 = X_2P_3 - X_3P_2$ on $\mathcal{S}(\mathbb{R}^3)$, see Example 3.31.

b) Use the following Theorem to show that L_i on $\mathcal{S}(\mathbb{R}^3)$ is essentially selfadjoint:
Theorem: Let A be selfadjoint. If a dense subspace $D \subset \mathrm{dom}\, A$ is invariant under $e^{itA} : D \to D$ for all $t \in \mathbb{R}$, then $A : D \to \mathcal{H}$ is essentially selfadjoint, see Reed&Simon 1 Theorem 8.11.

c) Show that the selfadjoint angular momentum operators L_1, L_2, L_3 are constants of motion and satisfy the angular momentum relations on $\mathcal{S}(\mathbb{R}^3)$.

Consider the isomorphism $L^2(\mathbb{R}^3) \simeq L^2(S^2, d\omega) \otimes L^2((0,\infty), r^2 dr)$, where $d\omega(\phi,\theta) = \sin(\theta) d\phi d\theta$. Let $(e_n)_{n\in\mathbb{N}}$ be an ONB of $\mathcal{H}_r := L^2((0,\infty), r^2 dr)$ and consider the ONB of $\mathcal{H}_\omega := L^2(S^2, d\omega)$ given by the spherical harmonics for $l \in \mathbb{N}_0$, $m \in \mathbb{Z}$ with $-l \le m \le l$

$$Y_l^m(\theta,\phi) := c_{m,l} e^{im\phi} P_l^m(\cos\theta), \theta \in [0,\pi], \phi \in [0, 2\pi], \tag{3.177}$$

where $c_{n,m}$ are normalization constants and P_l^m is the associated Legendre polynomial that satisfies the equation

$$(1-x^2)\frac{d^2}{dx^2}P_l^m(x) - 2x\frac{d}{dx}P_l^m(x) + [l(l+1) - \frac{m^2}{1-x^2}]P_l^m(x) = 0. \tag{3.178}$$

d) Show that $\psi_{n,l,m} := e_n \otimes Y_l^m, n \in \mathbb{N}, l \in \mathbb{N}_0, m \in \mathbb{Z}$ with $-l \le m \le l$, is an ONB of $\mathcal{H}$. Use that L^2 is given in spherical coordinates by $L^2 = -\frac{1}{\sin\theta}\partial_\theta \sin\theta \partial_\theta - \frac{1}{\sin^2\theta}\partial_\phi^2$ and show that

$$L^2 \psi_{n,l,m} = l(l+1)\psi_{n,l,m}. \tag{3.179}$$

Conclude that L^2 is essentially selfadjoint on $\mathcal{S}(\mathbb{R}^3)$ and a constant of motion.

Exercise 3.12 (Harmonic Polynomials). In this exercise we show that

$$\mathrm{Pol}_l = V_l \oplus \|x\|^2 \mathrm{Pol}_{l-2}, \quad l \in \mathbb{N}_0, \tag{3.180}$$

where V_l is the space of harmonic polynomials of homogeneous degree l. Proceed in the following steps:

Consider for $p = \sum_{\alpha\in\mathbb{N}_0^3} c_\alpha x^\alpha \in \mathrm{Pol}$ the differential operator

$$D(p) := \sum_{\alpha\in\mathbb{N}_0^3} \overline{c_\alpha}\partial_\alpha \in \mathrm{DiffPol}(\mathbb{R}^3) \tag{3.181}$$

acting on $\mathrm{Pol}(\mathbb{R}^3)$. Define the sesquilinear form

$$\langle\cdot,\cdot\rangle : \mathrm{Pol}(\mathbb{R}^3) \times \mathrm{Pol}(\mathbb{R}^3) \to \mathbb{C}, \quad \langle p, q\rangle := (D(p)q)(0) \tag{3.182}$$

by acting with $D(p)$ on q and then evaluating the resulting polynomial at 0.

a) Show that $\langle\cdot,\cdot\rangle$ is an scalar product on $\mathrm{Pol}(\mathbb{R}^3)$ and that Pol_l and Pol_k are orthogonal if $l \neq k$.

b) Consider the Laplace operator $\Delta = D(\|x\|^2) = \sum_{i=1}^3 \partial_i^2$ as a map

$$\Delta : \mathrm{Pol}(\mathbb{R}^3) \to \mathrm{Pol}(\mathbb{R}^3). \tag{3.183}$$

Show that $\Delta^* = \|x\|^2$ w.r.t. the scalar product introduced above.

c) Use the Kern-Range Lemma to show that

$$\mathrm{Pol}_l = V_l \oplus \|x\|^2 \,\mathrm{Pol}_{l-2}, \quad l \in \mathbb{N}_0. \tag{3.184}$$

Exercise 3.13 (Spherical Harmonics). For angular momentum $l = 0$ and $l = 1$, consider the spaces V_l of harmonic polynomials homogeneous of degree l as subspaces of $L^2(S^2)$ by restriction to S^2. Compute orthonormal bases of these two spaces (orthonormal w.r.t. the scalar product of $L^2(S^3)$), consisting of simultaneous eigenvectors of L^2 and L_3.

Problem 3.1 (Groenewold-van Hove Theorem). Consider a unital algebra $\mathcal{A}$ and a map

$$Q : \mathrm{Pol}(\mathbb{R}^{2d}) \to \mathcal{A}, \tag{3.185}$$

called polynomial quantization map, between the polynomial algebra over classical phase space $\mathbb{R}^{2d}$ and the algebra $\mathcal{A}$. Assume that Q satisfies the following properties:

a) Q is linear,
b) $Q(X^\alpha) = Q(X)^\alpha$ and $Q(P^\alpha) = Q(P)^\alpha$ for every $\alpha \in \mathbb{N}_0^d$,
c) Q is unital $Q(1) = 1$,
d) $Q(\{f, g\}) = -i[Q(f), Q(g)]$.

Show that no such map Q exists.

Hint: Consider $X^2P^2 = \frac{1}{9}\{X^3, P^3\} = \frac{1}{6}\{X^2P, XP^2\}$.

Problem 3.2 (Weyl Quantization). Consider the Weyl quantization map

$$\begin{aligned} W &: \mathrm{Pol}(\mathbb{R}^{2d}) \to \mathrm{DiffPol}(\mathbb{R}^d), \\ (W(f)\psi)(x) &:= (2\pi)^{-d} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{i\langle x-y,p\rangle} f\left(\tfrac{x+y}{2}, p\right)\psi(y)\,dy\,dp. \end{aligned} \tag{3.186}$$

and recall that $W(f_{\alpha\beta}) = 2^{-|\alpha|}\sum_{\mu\le\alpha}\binom{\alpha}{\mu} X^{\alpha-\mu}P^\beta X^\mu$ for the monomial function $f_{\alpha\beta}(x,p) = x^\alpha p^\beta$ with $\alpha, \beta \in \mathbb{N}_0^d$.

Show the following:

a) $W(\{X_n^{\mathrm{cl}}, P_m^{\mathrm{cl}}\}) = -i[W(X_n^{\mathrm{cl}}), W(P_m^{\mathrm{cl}})]$,
b) $W((X^{\mathrm{cl}})^\alpha) = X^\alpha$ and $W((P^{\mathrm{cl}})^\alpha) = P^\alpha$ for $\alpha \in \mathbb{N}_0$,
c) $W(f^*) = W(f)^*$ for all $f \in \mathrm{Pol}(\mathbb{R}^{2d})$,
d) the Weyl quantization map $W : \mathrm{Pol}(\mathbb{R}^{2d}) \to \mathrm{DiffPol}(\mathbb{R}^d)$ is bijective,
e) $L_j := W(L_j^{\mathrm{cl}})$, $j = 1, 2, 3$ satisfy the angular momentum relations

$$-i[L_1, L_2] = L_3, \quad \text{etc..} \tag{3.187}$$

Problem 3.3 (Free Schrödinger Equation). Consider the free Hamiltonian $H := P^2$ on $\mathcal{S}(\mathbb{R})$.

a) Show that H is symmetric and essentially selfadjoint.

b) Show that $\sigma(\overline{H}) = [0, \infty)$.

c) Denote for $S \subset [0, \infty]$ the set $\sqrt{S} := \{\sqrt{\lambda} \ : \ \lambda \in S\}$. Show that the spectral measure $E^{\overline{H}}$ is given by

$$E^{\overline{H}}(\cdot) = E^{\overline{P}}(\sqrt{\cdot}) + E^{\overline{P}}(-\sqrt{\cdot}). \tag{3.188}$$

Problem 3.4 (Bounded perturbation of selfadjoint operators). Let $(\mathrm{dom}A, A)$ be a densely defined closable operator on a Hilbert space $\mathcal{H}$ and $B \in \mathcal{B}(\mathcal{H})$. Show that:

a) $(\mathrm{dom}A, A + B)$ is closable with adjoint $(\mathrm{dom}A^*, A^* + B^*)$.

b) If $(\mathrm{dom}A, A)$ is (essentially) selfadjoint and B is selfadjoint, then $(\mathrm{dom}A, A+B)$ is (essentially) selfadjoint.

As an application, let $\mathcal{H} = L^2(\mathbb{R})$ and show that the operator

$$H = -\Delta + M_V \ : \ \mathcal{S}(\mathbb{R}) \to \mathcal{H}, \quad (Hf)(x) = -f''(x) + V(x)f(x)$$

is essentially selfadjoint, where $V \ : \ \mathbb{R} \to \mathbb{R}$ is a bounded continuous function.

Problem 3.5 (Free Schrödinger equation 2). Recall that for $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$ the solution of the free Schrödinger equation is given in Proposition 3.11 for $t \in \mathbb{R}$ by

$$\psi_t(x) := \psi(t, x) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{i\langle p, x\rangle} e^{-\frac{i\|p\|^2 t}{2m}} \tilde{\psi}_0(p)\, dp. \tag{3.189}$$

Show the following:

a) For $t \neq 0$ the solution can be expressed as

$$\psi(t, x) = \Big(\frac{m}{2\pi i t}\Big)^{d/2} \int_{\mathbb{R}^d} e^{i\frac{m\|x-y\|^2}{2t}} \psi_0(y)\, dy. \tag{3.190}$$

Hint: For $f, g \in L^1(\mathbb{R})$ it holds $(2\pi)^{-d/2}\mathcal{F}(f * g) = (\mathcal{F}f) \cdot (\mathcal{F}g)$, see Analysis 3 Lemma 3.10.

b) The L^2-norm is constant $\|\psi_t\| = \|\psi_0\|$ for all $t \in \mathbb{R}$ and

$$\sup_{x \in \mathbb{R}^d} |\psi(t, x)| \leq \Big(\frac{m}{2\pi t}\Big)^{d/2} \|\psi_0\|_{L^1}, \qquad t \neq 0. \tag{3.191}$$

Conclude that $\sup_{x\in\mathbb{R}^d} |\psi(t, x)| \to 0$ for $t \to \pm\infty$.

Problem 3.6 (Classical vs Quantum Position). Consider a continuous potential $V \ : \ \mathbb{R} \to \mathbb{R}$ with $V(x) \to \infty$ for $|x| \to \infty$.

a) Consider $H^{\mathrm{cl}}(x, p) := \frac{p^2}{2m} + V(x)$ and $\nu \in \mathcal{P}(\mathbb{R}^2)$ s.t. $H^{\mathrm{cl}}_* \nu$ has compact sup-

port in $\mathbb{R}$. Show that $P_*^{\mathrm{cl}}\nu$ and $X_*^{\mathrm{cl}}\nu$ have compact support.

b) Consider $H := \frac{P^2}{2m} + V(X)$ for $V(X) = X^2$. Consider $\psi \in \mathcal{H}$ normalized with $\operatorname{supp} \mu_\psi^H$ in the compact interval $[E_-, E_+]$. Show that μ_ψ^X does not have compact support. Interpret the result physically.

Problem 3.7 (Kato-Rellich Theorem). Prove the statement in the Kato-Rellich Theorem (Thm. 3.36) about essentially selfadjoint operators.

Problem 3.8 (Pöschl-Teller Potential). Consider for $l \in \mathbb{N}_0$ the Pöschl-Teller potential

$$V_l : \mathbb{R} \to \mathbb{R}, \quad V_l(x) := -\frac{l(l+1)}{\cosh^2(x)} \tag{3.192}$$

and the associated Schrödinger operator $H_l := P^2 + V_l(X)$ on $\mathcal{S}(\mathbb{R})$.

a) Show that for H_l is essentially selfadjoint for all $l \in \mathbb{N}_0$.

b) Show that for $l \in \mathbb{N}_0$

$$[0, \infty) \subset \sigma(\overline{H_l}). \tag{3.193}$$

c) Define the operator

$$a_l : \mathcal{S}(\mathbb{R}) \to L^2(\mathbb{R}), \quad (a_l\psi)(x) := (P\psi)(x) - il\tanh(x)\psi(x). \tag{3.194}$$

Show that a_l is closable and

$$a_l^* = \overline{a_l}^* = \overline{P} + ilM_{\tanh}. \tag{3.195}$$

d) Define the operators

$$A_l := a_l^* a_l \text{ on } \operatorname{dom} A_l \quad \text{and} \quad B_l := a_l a_l^* \text{ on } \operatorname{dom} B_l. \tag{3.196}$$

Show that for $l \in \mathbb{N}$ on $\mathcal{S}(\mathbb{R})$ it holds

$$H_l = A_l - l^2 \quad \text{and} \quad H_{l-1} = B_l - l^2. \tag{3.197}$$

e) Show that $a_l^*\psi \neq 0$ for all $\psi \in \mathcal{S}(\mathbb{R})$ and $l \in \mathbb{N}_0$ and that $\ker(\overline{a_l}) \supset \mathbb{C} \cdot \psi_0^l$ for $l \in \mathbb{N}$, where $\psi_0^l(x) = \cosh^{-l}(x)$.

f) Show that for $l \in \mathbb{N}$

$$\sigma_p(\overline{H_l}) \supset \bigcup_{j=1}^{l} \{-j^2\}. \tag{3.198}$$

M Appendix: Mathematical background

M.1 Measure theory

In this section a few facts from measure theory are collected. For more details, see for example [Els18]. We will typically consider measures on (subsets of) $\mathbb{R}^n$ or $\mathbb{C}^n$, which are in particular locally compact σ-compact Hausdorff spaces. In the following X, X_1, X_2 will always denote locally compact σ-compact Hausdorff spaces. We write $(X, \mathfrak{A})$ for measurable spaces, i.e. $\mathfrak{A}$ is a sigma algebra on X, and $\mathfrak{B}(X)$ for the Borel sigma algebra on X, the sigma algebra generated by all open subsets of X.

Definition M.1 (Borel measures).

a) A *Borel measure* on X is a measure $\mu : \mathfrak{B}(X) \to [0, \infty]$ that is *locally finite*, i.e. every point has an open neighbourhood $\mathcal{O}$ with $\mu(\mathcal{O}) < \infty$.

b) A *Borel probability measure* is a measure $\mu : \mathfrak{B}(X) \to [0, 1]$ with $\mu(X) = 1$. The set of all Borel probability measures on X is denoted $\mathcal{P}(X)$.

c) The *support* $\operatorname{supp} \mu$ of a Borel measure μ is the complement of the largest open set $B \in \mathfrak{B}(X)$ with $\mu(B) = 0$. The set of all compactly supported Borel probability measures on X is denoted $\mathcal{P}_c(X)$.

We note that every Borel probability measure is inner and outer regular and in particular a Radon measure.

Every continuous function $\varphi : X \to \mathbb{C}$ is Borel measurable. It is integrable in particular if it has compact support, i.e. $C_c(X) \subset L^1(X, d\mu)$ for Borel measures μ. If $\mu \in \mathcal{P}_c(X)$ has compact support, even $C(X) \subset L^1(X, d\mu)$.

There exist several versions of Riesz' Representation Theorems, characterizing positive linear functionals on function spaces in terms of measures (a linear functional on a function space is called positive if it maps functions that take only non-negative values to non-negative numbers). For the space of continuous functions $C(X)$ on X, one needs compactly supported measures because no growth condition at infinity is imposed.

Theorem M.2 (Riesz' Representation Theorem). *There exists a bijection between positive linear functionals $\omega : C(X) \to \mathbb{C}$ and compactly supported Borel measures $\mu \in \mathcal{P}_c(X)$ given by*

$$\omega(f) = \int_X f \, d\mu, \qquad f \in C(X). \tag{M.1}$$

A distinguished class of probability measures are Dirac measures.

Lemma M.3 (Probability measures concentrated at one point). *Let $\nu \in \mathcal{P}(X)$ with $\operatorname{supp} \nu = \{x\}$ for some $x \in X$. Then $\nu = \delta_x$.*

Proof. In any Hausdorff space, singletons $\{x\}$ are closed, so $X \backslash \{x\}$ is open and in particular Borel. By definition of support and $\operatorname{supp} \nu = \{x\}$, the Borel set $X \backslash \{x\}$ has measure 0. As ν is a probability measure, $\nu(\{x\}) = 1$. This implies $\nu = \delta_x$. □

Specifically for Borel probability measures on $\mathbb{R}$, we recall the following notions.

Definition M.4 (Moments, expectation values, variances). Let $\nu \in \mathcal{P}(\mathbb{R})$.

a) If $\lambda \mapsto \lambda^n$, $n \in \mathbb{N}$, is integrable w.r.t. ν (i.e. an element of $L^1(\mathbb{R}, d\nu)$, that is $\int_{\mathbb{R}} |\lambda|^n \, d\nu(\lambda) < \infty$), then the *n-th moment of* ν is defined as

$$\mathbb{M}_n(\nu) := \int_{\mathbb{R}} \lambda^n \, d\nu(\lambda). \tag{M.2}$$

b) If the first moment $\mathbb{M}_1(\nu)$ exists, it is called the *expectation value* (or *expected value*, *mean value*) of ν and denoted[a]

$$\mathbb{E}(\nu) := \mathbb{M}_1(\nu). \tag{M.3}$$

c) If the second moment exists, the *variance* of ν is defined as

$$\mathbb{V}(\nu) := \mathbb{M}_2(\nu) - \mathbb{M}_1(\nu)^2 \tag{M.4}$$

$$= \int_{\mathbb{R}} \lambda^2 \, d\nu(\lambda) - \left(\int_{\mathbb{R}} \lambda \, d\nu(\lambda) \right)^2 = \int_{\mathbb{R}} (\lambda - \mathbb{E}(\nu))^2 \, d\nu(\lambda). \tag{M.5}$$

[a] Usually this is written as $\mathbb{E}_\nu(X)$ for the random variable X; but we drop X in order to avoid confusion with position (operators) or topological spaces.

We note that in case the n-th moment exists, also all lower moments exist. This is the case because for $k \leq n$, we have $|\lambda|^k \leq 1 + |\lambda|^n$ for every $\lambda \in \mathbb{R}$, so $\int_{\mathbb{R}} |\lambda|^k \, d\nu(\lambda) \leq \int_{\mathbb{R}} d\nu(\lambda) + \int_{\mathbb{R}} |\lambda|^n d\nu(\lambda) = 1 + \int_{\mathbb{R}} |\lambda|^n d\nu(\lambda)$.

Pushforward measures

Definition M.5 (Pushforward measures). Let $(X_1, \mathfrak{A}_1), (X_2, \mathfrak{A}_2)$ be measurable spaces, $\varphi : (X_1, \mathfrak{A}_1) \to (X_1, \mathfrak{A}_2)$ measurable, and $\mu : \mathfrak{A}_1 \to [0, \infty]$ a measure.

The *pushforward measure* of μ by φ is

$$\varphi_* \mu : \mathfrak{A}_2 \to [0, \infty], \qquad (\varphi_* \mu)(B_2) := \mu(\varphi^{-1}(B_2)). \tag{M.6}$$

Proposition M.6 (Properties of pushforward measures). *Let μ be a Borel probability measure on X_1 and $\varphi : X_1 \to X_2$ Borel measurable.*

a) $\varphi_* \mu$ is a Borel probability measure on X_2.

b) $\operatorname{supp} \varphi_* \mu \subset \overline{\varphi(\operatorname{supp} \mu)}$. For continuous φ, equality holds.

c) A measurable function $g : X_2 \to \overline{\mathbb{R}}$ is integrable w.r.t. $\varphi_* \mu$ if and only if $g \circ \varphi$ is integrable w.r.t. μ. In this case,

$$\int_{X_2} g \, d(\varphi_* \mu) = \int_{X_1} (g \circ \varphi) \, d\mu. \tag{M.7}$$

Measures with densities

Definition M.7 (Measures with densities). Let μ be a measure on $(X, \mathfrak{A})$ and $\rho : X \to [0, \infty]$ measurable. Then

$$(\rho\mu)(A) := \int_A \rho \, d\mu \tag{M.8}$$

is a measure on $(X, \mathfrak{A})$, called the *measure with density ρ relative to μ*.

A measurable function $f : X \to \mathbb{C}$ is $\rho\mu$-integrable if and only if $f \cdot \rho$ is μ-integrable. In this case,

$$\int_X f \, d(\rho\mu) = \int_X f\rho \, d\mu. \tag{M.9}$$

We therefore also write

$$d(\rho\mu)(x) = \rho(x) \, d\mu(x) \tag{M.10}$$

for the measure $\rho\mu$.

Definition M.8. Let μ, ν be two measures on $(X, \mathfrak{A})$. Then ν is called *absolutely continuous w.r.t. μ*, written $\nu \ll \mu$, if every μ-zero set is a ν-zero set.

Theorem M.9 (Radon-Nikodým Theorem). *If μ is a σ-finite measure and $\nu \ll \mu$, then ν has a density ρ w.r.t. μ.*

M.2 Linear algebra

In this section a few facts from linear algebra are collected.

$\mathbb{C}^N$ as a Hilbert space

The complex linear space $\mathbb{C}^N$ carries the Euclidean scalar product $\langle \cdot, \cdot \rangle$ which makes it a Hilbert space. We will write $(e_k)_{k=1,\dots,N}$ for the canonical orthonormal basis of $\mathbb{C}^N$.

Given any orthonormal basis (b_k) of $\mathbb{C}^N$ and any $\psi, \xi \in \mathcal{H}$, norms and scalar products can be expressed as

$$\|\psi\|^2 = \sum_{k=1}^{N} |\langle b_k, \psi \rangle|^2, \tag{M.11}$$

$$\langle \psi, \xi \rangle = \sum_{k=1}^{N} \langle \psi, b_k \rangle \langle b_k, \xi \rangle. \tag{M.12}$$

$M_N(\mathbb{C})$ as a C^*-algebra

The algebra $M_N(\mathbb{C})$ of all complex $(N \times N)$-matrices is denoted $M_N(\mathbb{C})$. It has $(N \times N)$ identity matrix 1 as multiplicative unit, and the star operation $(A^*)_{kl} = \overline{A_{lk}}$. Equivalently,

$$\langle \psi, A\xi \rangle = \langle A^* \psi, \xi \rangle, \qquad \psi, \xi \in \mathbb{C}^N. \tag{M.13}$$

With these operations, $M_N(\mathbb{C})$ is a unital *-algebra. The operator norm

$$\|A\| := \sup_{\psi \in \mathbb{C}^N \setminus \{0\}} \frac{\|A\psi\|}{\|\psi\|} \tag{M.14}$$

satisfies the C^*-identity

$$\|A^* A\| = \|A\|^2, \qquad A \in M_N(\mathbb{C}). \tag{M.15}$$

As an algebra, $M_N(\mathbb{C})$ is noncommutative (for $N \geq 2$) and has trivial center: The only elements $A \in M_N(\mathbb{C})$ that commute with all $B \in M_N(\mathbb{C})$ are the multiples of the identity, $A = c \cdot 1, c \in \mathbb{C}$.

We will also often talk about positive elements.

Proposition M.10 (Positive matrices). *Let $A \in M_N(\mathbb{C})$. Then the following are equivalent:*

a) $A = A^*$ and all eigenvalues (see Def. M.13) λ of A are non-negative, $\lambda \geq 0$,
b) there exists $B \in M_N(\mathbb{C})$ such that $A = B^* B$,
c) $\langle \psi, A\psi \rangle \geq 0$ for all $\psi \in \mathbb{C}^N$.

A matrix satisfying these conditions is called positive, *written as $A \geq 0$.*

Dirac notation and matrix units

Given $\psi \in \mathbb{C}^N$, we consider the linear functional

$$\langle\psi| : \mathbb{C}^N \to \mathbb{C}, \qquad \xi \mapsto \langle\psi, \xi\rangle, \tag{M.16}$$

and given $\psi, \phi \in \mathbb{C}^N$, we consider the matrix (linear map)

$$|\phi\rangle\langle\psi| : \mathbb{C}^N \to \mathbb{C}^N, \qquad \xi \mapsto \langle\psi, \xi\rangle\phi. \tag{M.17}$$

These maps satisfy

$$(|\phi\rangle\langle\psi|)(|\phi'\rangle\langle\psi'|) = \langle\psi, \phi'\rangle \cdot |\phi\rangle\langle\psi'|, \tag{M.18}$$

$$(|\phi\rangle\langle\psi|)^* = |\psi\rangle\langle\phi|. \tag{M.19}$$

In particular, $P_\psi := |\psi\rangle\langle\psi|$ is the orthogonal rank one projection onto the subspace $\mathbb{C}\psi$.

Definition M.11 (Matrix units). A *system of matrix units* in $M_N(\mathbb{C})$ is a family $E_{kl} \in M_N(\mathbb{C})$, $k, l \in \{1, \dots, N\}$ (here the lower indices do not indicate rows/columns!), such that for all $i, j, k, l \in \{1, \dots, N\}$,

$$E_{kl}E_{ij} = \delta_{li}E_{kj}, \qquad (E_{kl})^* = E_{lk}, \qquad \sum_{k=1}^{N} E_{kk} = 1. \tag{M.20}$$

Lemma M.12. *Every orthonormal basis (b_k) of $\mathbb{C}^N$ defines a system of matrix units via $E_{kl} := |b_k\rangle\langle b_l|$, and every system of matrix units is of this form. A system of matrix units forms a linear basis of $M_N(\mathbb{C})$ (i.e. every $A \in M_N(\mathbb{C})$ is a unique linear combination of the E_{kl}).*

The matrix units defined by the canonical orthonormal basis of $\mathbb{C}^N$ are called the *canonical matrix units of $M_N(\mathbb{C})$*.

The spectral theorem and functional calculus for $M_N(\mathbb{C})$

Definition M.13 (Spectra of Matrices). Let $A \in M_N(\mathbb{C})$.

a) An eigenvalue of A is a complex number λ such that $(A - \lambda 1)$ is not invertible, i.e. such that there exists $\psi \in \mathbb{C}^N$, $\psi \neq 0$, with $A\psi = \lambda\psi$. The set of all eigenvalues of A is called the *spectrum of A* and denoted $\sigma(A)$.

b) Given $\lambda \in \sigma(A)$, the orthogonal projection onto the subspace spanned by of all eigenvectors of A with eigenvalue λ is denoted $E^A(\lambda) := E^A(\{\lambda\})$ and called the *spectral projection of A with eigenvalue λ*. For normal A, the map

$$\sigma(A) \ni \lambda \mapsto E^A(\lambda) \in M_N(\mathbb{C}) \tag{M.21}$$

is called the *spectral measure of A*.

The spectrum of $A \in M_N(\mathbb{C})$ consists of at most N points in $\mathbb{C}$ (the zeros of the characteristic polynomial $\lambda \mapsto \det(A - \lambda 1)$, which is of degree N). For selfadjoint A, the spectrum is real, i.e. $\sigma(A) \subset \mathbb{R}$. For unitary A, the spectrum lies on the unit circle, $\sigma(A) \subset \mathbb{T}$.

Theorem M.14 (The spectral theorem for $M_N(\mathbb{C})$). *Let $A \in M_N(\mathbb{C})$ be normal, i.e. $AA^* = A^*A$.*

a) *E^A is a* projection-valued measure*: The spectral projections of A are mutually orthogonal and sum to the identity, i.e. for $\lambda, \lambda' \in \sigma(A)$,*

$$E^A(\lambda)E^A(\lambda') = \delta_{\lambda,\lambda'}E^A(\lambda), \qquad \sum_{\lambda\in\sigma(A)} E^A(\lambda) = 1. \tag{M.22}$$

b) *A can be represented in spectral form, namely*

$$A = \sum_{\lambda\in\sigma(A)} \lambda E^A(\lambda). \tag{M.23}$$

The spectral measure is the only projection-valued measure such that (M.23) holds.

Note that the spectral theorem implies in particular that every normal $A \in M_N(\mathbb{C})$ admits an orthonormal basis (b_k) consisting only of eigenvectors of A (with eigenvalues λ_k), and can be written as

$$A = \sum_{k=1}^{N} \lambda_k |b_k\rangle\langle b_k|. \tag{M.24}$$

But in constrast to the spectral representation (M.23), this representation is not canonical because it requires choices of orthonormal bases in the eigenspaces of A.

The projection-valued measure carries its name for a reason: It takes values in orthogonal projections, and the map on the power set of $\sigma(A)$,

$$2^{\sigma(A)} \ni S \mapsto E^A(S) := \sum_{\lambda\in S} E^A(\{\lambda\}) \tag{M.25}$$

is normalized in the sense $E^A(\emptyset) = 0$, $E^A(\sigma(A)) = 1$, and additive: $E^A(S_1 \sqcup S_2) = E^A(S_1) + E^A(S_2)$ for disjoint $S_1, S_2 \subset \sigma(A)$. In particular, for every density matrix $\rho \in \mathcal{D}_N$ (Def. 2.10), the map

$$2^{\sigma(A)} \ni S \mapsto \omega_\rho(E^A(S)) \in [0,1] \tag{M.26}$$

is a probability measure on the spectrum $\sigma(A)$ of A.

The spectral theorem is very helpful for understanding functional calculus. Given $A \in M_N(\mathbb{C})$ and any function $f : \sigma(A) \to \mathbb{C}$, there exists a polynomial p such that $f|_{\sigma(A)} = p|_{\sigma(A)}$. For polynomials $p(\lambda) = \sum_{j=0}^m c_j\lambda^j$, we define $p(A) := \sum_{j=0}^m c_j A^j \in M_N(\mathbb{C})$.

Proposition M.15. *Let $A \in M_N(\mathbb{C})$ and p a polynomial.*

a) *The spectral mapping principle holds: $\sigma(p(A)) = p(\sigma(A))$.*

b) *If A is normal: $\|p(A)\| = \sup_{\lambda\in\sigma(A)} |p(\lambda)|$.*

For normal A, we define for any function $f : \mathbb{C} \to \mathbb{C}$

$$f(A) := \sum_{\lambda\in\sigma(A)} f(\lambda)\, E^A(\lambda). \tag{M.27}$$

This definition agrees with the direct polynomial definition. For characteristic functions, one finds

$$\chi_S(A) = \sum_{\lambda\in\sigma(A)} \chi_S(\lambda)\, E^A(\lambda) = \sum_{\lambda\in S} E^A(\lambda) = E^A(S), \qquad S \subset \sigma(A). \tag{M.28}$$

It is also often useful to talk about joint spectrum and joint spectral measures.

Proposition M.16 (Joint spectral measures). *Let $A_1, \dots, A_n \in M_N(\mathbb{C})$ be selfadjoint (with spectral measures E^{A_k}, $k = 1, \dots, n$) and commuting, i.e. $A_k A_l = A_l A_k$ for all $l, k \in \{1, \dots, n\}$. Then*

$$E^{A_1,\dots,A_n}(\lambda_1, \dots, \lambda_n) := E^{A_1}(\lambda_1) \cdots E^{A_n}(\lambda_n), \qquad \lambda_1, \dots, \lambda_n \in \mathbb{R}, \tag{M.29}$$

are orthogonal projections, and

$$E^{A_1,\dots,A_n} : \mathfrak{B}(\mathbb{R}^n) \to M_N(\mathbb{C}), \qquad S \mapsto \sum_{(\lambda_1,\dots,\lambda_n)\in S} E^{A_1,\dots,A_n}(\lambda_1, \dots, \lambda_n) \tag{M.30}$$

is a projection-valued measure (the joint spectral measure of $A_1, \dots, A_n$*), with the property*

$$A_1^{m_1} \cdots A_n^{m_n} = \sum_{\lambda\in\mathbb{R}^n} \lambda_1^{m_1} \cdots \lambda_n^{m_n} E^{A_1,\dots,A_n}(\lambda_1, \dots, \lambda_n), \qquad m_1, \dots, m_n \in \mathbb{N}_0. \tag{M.31}$$

The support $\operatorname{supp} E^{A_1,\dots,A_n} \subset \mathbb{R}^n$ *of* $E^{A_1,\dots,A_n}$ *is called the* joint spectrum *of* $(A_1, \dots, A_n)$.

Proof. The $E^{A_1}(\lambda_1) \cdots E^{A_n}(\lambda_n)$ are orthogonal projections because the $E^{A_k}(\lambda_k)$ are commuting orthogonal projections. The range of $E^{A_1}(\lambda_1) \cdots E^{A_n}(\lambda_n)$ is the subspace $\ker(A_1 - \lambda_1) \cap \dots \cap \ker(A_n - \lambda_n)$ of $\mathbb{C}^N$, the space of joint eigenvectors ψ, i.e. $A_k\psi = \lambda_k\psi$, $k = 1, \dots, n$. The formula for $A_1^{m_1} \cdots A_n^{m_n}$ follows by inserting the spectral representations of the A_k and using that they commute. □

Note that the joint spectrum is a subset of $\mathbb{R}^n$. Its elements $\lambda = (\lambda_1, \dots, \lambda_n)$ record the eigenvalues λ_k of A_k belonging to joint eigenvectors, $k = 1, \dots, n$. Note that we can write this measure also as

$$E^{A_1,\dots,A_n} = \sum_{\lambda\in\sigma(A_1,\dots,A_n)} E^{A_1,\dots,A_n}(\lambda) \cdot d\delta_\lambda \tag{M.32}$$

The matrix trace

The matrix trace is the linear functional

$$\operatorname{Tr} : M_N(\mathbb{C}) \to \mathbb{C}, \qquad \operatorname{Tr}(A) := \sum_{k=1}^{N} A_{kk}. \tag{M.33}$$

Proposition M.17 (The matrix trace). *The matrix trace has the following properties:*

a) $\mathrm{Tr} : M_N(\mathbb{C})$ *is linear and positive (i.e.* $\mathrm{Tr}(A) \geq 0$ *for* $A \geq 0$*),*

b) $\mathrm{Tr} : M_N(\mathbb{C})$ *is faithful: If a positive matrix* $A \geq 0$ *satisfies* $\mathrm{Tr}(A) = 0$*, then* $A = 0$*.*

c) Tr *is invariant under cyclic permutations of its argument,*

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA), \qquad A, B \in M_N(\mathbb{C}). \tag{M.34}$$

d) If $\tau : M_N(\mathbb{C}) \to \mathbb{C}$ *is linear and satisfies* $\tau(AB) = \tau(BA)$ *for all* $A, B \in M_N(\mathbb{C})$*, then* τ *is a multiple of* Tr*.*

The cyclicity of Tr implies in particular that it is basis independent, namely

$$\mathrm{Tr}(A) = \sum_{k=1}^{N} \langle b_k, Ab_k \rangle \tag{M.35}$$

holds for every orthonormal basis (b_k) of $\mathbb{C}^N$.

The algebraic tensor product

We next recall some properties of algebraic tensor products. Proofs of the claims made here can be found in any text book on linear algebra.

Definition M.18 (Algebraic tensor product). Let V_1, V_2 be complex vector spaces. An (algebraic) *tensor product* of V_1 and V_2 is a complex vector space $T(V_1, V_2)$ together with a bilinear map $T : V_1 \times V_2 \to T(V_1, V_2)$ such that for every complex vector space W and every bilinear map $F : V_1 \times V_2 \to W$, there exists a unique linear map $\tilde{F} : T(V_1, V_2) \to W$ such that

$$\begin{array}{ccc} V_1 \times V_2 & \xrightarrow{T} & T(V_1, V_2) \\ & \searrow^{F} & \downarrow{\scriptstyle \tilde{F}} \\ & & W \end{array}$$

commutes.

Theorem M.19.

a) Every two vector spaces V_1, V_2 *have an algebraic tensor product. It is unique up to canonical isomorphism, and denoted* $V_1 \otimes V_2 := T(V_1, V_2)$*.*

b) $\dim(V_1 \otimes V_2) = \dim V_1 \cdot \dim V_2$*.*

c) If V_1, V_2 *are finite-dimensional and* $(b_j^1)_{j=1,\ldots,\dim V_1}$*,* $(b_j^2)_{j=1,\ldots,\dim V_2}$ *are bases of* V_1 *and* V_2*, respectively, then* $(T(b_j^1, b_l^2)_{j=1,\ldots,\dim V_1, l=1,\ldots,\dim V_2})$ *is a basis of* $T(V_1, V_2)$*.*

The bilinear map $T : V_1 \times V_2 \to V_1 \otimes V_2$ is denoted

$$V_1 \times V_2 \to V_1 \otimes V_2, \qquad (v_1, v_2) \mapsto v_1 \otimes v_2. \tag{M.36}$$

Note that given bases $(b_j^1)_{j=1,\dots,\dim V_1}, (b_j^2)_{j=1,\dots,\dim V_2}$ of V_1 and V_2, we have a tensor product basis $(b_j^1 \otimes b_l^2)_{j,l}$ of $V_1 \otimes V_2$. So every vector $\xi \in V_1 \otimes V_2$ may be written in the form

$$\xi = \sum_{j,l} c_{j,l}\, b_j^1 \otimes b_l^2.$$

However, vectors in $V_1 \otimes V_2$ are typically *not* of the form $\xi = v \otimes w$. Such tensor product vectors (or "simple tensors") exist and even contain bases, but form just a subset of $V_1 \otimes V_2$.

Corollary M.20. *Let V_1, V_2 be complex vector spaces and $A : V_1 \to V_1, A_2 : V_2 \to V_2$ linear. Then there exists a unique linear map*

$$A_1 \otimes A_2 : V_1 \otimes V_2 \to V_1 \otimes V_2 \tag{M.37}$$

such that

$$(A_1 \otimes A_2)(v_1 \otimes v_2) = (A_1 v_1) \otimes (A_2 v_2), \qquad v_1 \in V_1,\ v_2 \in V_2. \tag{M.38}$$

In particular,

$$(A_1 \otimes A_2)(B_1 \otimes B_2) = (A_1 B_1) \otimes (A_2 B_2) \tag{M.39}$$

for linear maps $A_1, B_1 : V_1 \to V_1, A_2, B_2 : V_2 \to V_2$.

In quantum mechanics, we are working with Hilbert spaces. In this setting, we are interested in how orthogonality and adjoints transfer to tensor products.

Proposition M.21 (Tensor products of finite-dimensional Hilbert spaces). *Let V_1, V_2 be finite-dimensional Hilbert spaces with scalar products $\langle\cdot,\cdot\rangle_1$ and $\langle\cdot,\cdot\rangle_2$.*

a) $\langle v_1 \otimes v_2, w_1 \otimes w_2\rangle := \langle v_1, w_1\rangle_1 \langle v_2, w_2\rangle_2$ is a scalar product on $V_1 \otimes V_2$ (so also $V_1 \otimes V_2$ is a finite-dimensional Hilbert space).

b) Given orthonormal bases $(b_j^1)_j, (b_l^2)_l$ of V_1, V_2, the tensor product basis $(b_j^1 \otimes b_l^2)_{j,l}$ is orthonormal w.r.t. the scalar product from a).

c) For linear maps $A_1 : V_1 \to V_1$ and $A_2 : V_2 \to V_2$, we have

$$(A_1 \otimes A_2)^* = A_1^* \otimes A_2^* \tag{M.40}$$

and

$$\mathrm{Tr}_{V_1 \otimes V_2}(A_1 \otimes A_2) = \mathrm{Tr}_{V_1}(A_1) \cdot \mathrm{Tr}_{V_2}(A_2). \tag{M.41}$$

We only comment on the last item: Picking orthonormal bases $(b_j^1)_j, (b_l^2)_l$ of V_1, V_2, the trace $\mathrm{Tr}_{V_1 \otimes V_2}(A_1 \otimes A_2)$ can be evaluated in the corresponding tensor product basis as

$$\begin{aligned}
\mathrm{Tr}_{V_1 \otimes V_2}(A_1 \otimes A_2) &= \sum_{j,l} \langle b_j^1 \otimes b_l^2, (A_1 \otimes A_2)(b_j^1 \otimes b_l^2)\rangle \\
&= \sum_{j,l} \langle b_j^1, A_1 b_j^2\rangle \langle b_l^2, A_2 b_l^2\rangle \\
&= \sum_{j} \langle b_j^1, A_1 b_j^2\rangle \sum_{l} \langle b_l^2, A_2 b_l^2\rangle \\
&= \mathrm{Tr}_{V_1}(A_1)\, \mathrm{Tr}_{V_2}(A_2).
\end{aligned}$$

M.3 Fourier analysis

In this section some facts about Fourier transforms are collected. For a detailed discussion including proofs, see, for example, [RS75] or [Hal13, App. A].

Given a function $f \in L^1(\mathbb{R}^d)$ (w.r.t. Lebesgue measure), we define its *Fourier transform* as

$$\tilde{f} : \mathbb{R}^d \to \mathbb{C}, \qquad \tilde{f}(p) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{-i\langle p,x\rangle} f(x)\, dx. \tag{M.42}$$

Since $f \in L^1(\mathbb{R}^d)$, this integral converges, i.e. $\tilde{f}$ is a well-defined function which is bounded by $|\tilde{f}(p)| \leq (2\pi)^{-d/2}\|f\|_1$. Here $\|f\|_1 = \int_{\mathbb{R}^d} |f(x)|dx$ denotes the L^1-norm.

The Riemann-Lebesgue "Lemma" clarifies more precisely what kind of functions the $\tilde{f}$ are. We write

$$C_0(\mathbb{R}^d) := \{f : \mathbb{R}^d \to \mathbb{C} \;:\; f \text{ continuous}, \quad \lim_{R\to\infty} \sup_{\|x\|\geq R} |f(x)| = 0\} \tag{M.43}$$

for the space of all continuous functions vanishing at infinity.

Theorem M.22 (Riemann-Lebesgue). *Let $f \in L^1(\mathbb{R}^d)$. Then $\tilde{f} \in C_0(\mathbb{R}^d)$.*

We may therefore view Fourier transform as a linear map $L^1(\mathbb{R}^d) \to C_0(\mathbb{R}^d)$, $f \mapsto \tilde{f}$. Because of the two different function spaces appearing here, it is often more convenient to consider the Fourier transform of Schwartz functions $f \in \mathcal{S}(\mathbb{R}^d)$ (a dense subspace of $L^1(\mathbb{R}^d)$).

Theorem M.23 (The Fourier transform on $\mathcal{S}(\mathbb{R}^d)$).

a) Let $f \in \mathcal{S}(\mathbb{R}^d)$. Then $\tilde{f} \in \mathcal{S}(\mathbb{R}^d)$.

b) The Fourier transform on $\mathcal{S}(\mathbb{R}^d)$, i.e. the map

$$\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d), \quad f \mapsto \tilde{f}, \tag{M.44}$$

is a linear bijection. Its inverse (the inverse Fourier transform*) is given by*

$$\begin{aligned} &\mathcal{F}^{-1} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d), \\ &(\mathcal{F}^{-1}f)(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{i\langle p,x\rangle} f(p)\, dp = \tilde{f}(-x). \end{aligned} \tag{M.45}$$

c) For any $k \in \{1, \dots, d\}$,

$$\mathcal{F}P_k = X_k\mathcal{F}, \qquad \mathcal{F}^{-1}X_k = P_k\mathcal{F}^{-1}, \tag{M.46}$$

where X_k, P_k are the Schrödinger position/momentum operators (Def. 3.4).

The exchange relations between the position/momentum operators and the Fourier trans-

form (M.46) can be memorized as follows. Formally,

$$\begin{aligned}(\mathcal{F}P_k f)(p) &= (2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{-i\langle p,x\rangle}(-i\partial_{x_k} f)(x)\,dx \\ &= (2\pi)^{-d/2}\int_{\mathbb{R}^d} (i\partial_{x_k} e^{-i\langle p,x\rangle})f(x)\,dx \qquad \text{(integration by parts)} \\ &= p_k\cdot(2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{-i\langle p,x\rangle}f(x)\,dx \\ &= (X_k\mathcal{F}f)(p)\end{aligned}$$

and

$$\begin{aligned}(\mathcal{F}^{-1}X_k f)(p) &= (2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{i\langle p,x\rangle}x_k f(x)\,dx \\ &= (2\pi)^{-d/2}\int_{\mathbb{R}^d} (-i\partial_{p_k} e^{i\langle p,x\rangle})f(x)\,dx \\ &= -i\partial_{p_k}\left((2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{i\langle p,x\rangle}f(x)\,dx\right) \qquad \text{(differentiation under the integral)} \\ &= (P_k\mathcal{F}^{-1}f)(p)\end{aligned}$$

imply (M.46). Using $f\in\mathcal{S}(\mathbb{R}^d)$, these steps can be justified.

Another important property of the Fourier transform is its interplay with the L^2-scalar product:

Proposition M.24 (Plancherel Theorem). *For $f,g\in\mathcal{S}(\mathbb{R}^d)$,*

$$\langle f,g\rangle_{L^2(\mathbb{R}^d)} = \langle\tilde f,\tilde g\rangle_{L^2(\mathbb{R}^d)}. \tag{M.47}$$

In particular,

$$\|f\|_2 = \|\tilde f\|_2. \tag{M.48}$$

As $\mathcal{S}(\mathbb{R}^d)$ is a dense subspace of $L^2(\mathbb{R}^d)$, the Plancherel Theorem allows us to extend the Fourier transform to a surjective isometry (a unitary operator) on $L^2(\mathbb{R}^d)$: There exists a unique unitary, usually also denoted $\mathcal{F}:L^2(\mathbb{R}^d)\to L^2(\mathbb{R}^d)$, such that it restricts to the previously defined Fourier transform on $\mathcal{S}(\mathbb{R}^d)$.

For $f\in L^2(\mathbb{R}^d)$, the Fourier integral $(2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{-i\langle p,x\rangle}f(x)\,dx$ might not converge. To compute $\mathcal{F}f$, one can proceed as follows: One considers the characteristic function χ_R of the ball $B_R\subset\mathbb{R}^d$ of radius $R>0$ and center 0. Then $f\chi_R\in L^1(\mathbb{R}^d)$, and $f\chi_R\to f$ in L^2-norm as $R\to\infty$. Hence

$$\mathcal{F}f = \lim_{R\to\infty}\mathcal{F}(\chi_R f),$$

where the limit holds in L^2-norm.

M.4 Functional analysis and spectral theory

In this section some important facts about operators on Hilbert spaces and in particular their spectral theory are collected.

Hilbert spaces

Definition M.25 (Hilbert spaces). A complex *Hilbert space* is a complex vector space $\mathcal{H}$ with a scalar product $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ that is complete in the topology defined by the norm $\|\psi\| := \langle \psi, \psi \rangle^{1/2}$, $\psi \in \mathcal{H}$.

Recall that in a Hilbert space, the *Cauchy-Schwarz Inequality* holds, namely

$$|\langle \psi, \varphi \rangle| \leq \|\psi\| \|\varphi\|, \qquad \psi, \varphi \in \mathcal{H}. \tag{M.49}$$

In particular, the scalar product determines the norm. In a Hilbert space, also the converse holds: The scalar product can be expressed by the norm, via the *polarization identity*

$$\langle \psi, \varphi \rangle = \frac{1}{4} \left(\|\psi + \varphi\|^2 - \|\psi - \varphi\|^2 - i\|\psi + i\varphi\|^2 + i\|\psi - i\varphi\|^2 \right). \tag{M.50}$$

This does not mean that any Banach space is a Hilbert space!

Two vectors ψ, φ in a Hilbert space $\mathcal{H}$ are called *orthogonal* if $\langle \psi, \varphi \rangle = 0$. This is also written as $\psi \perp \varphi$. The *orthogonal complement* of a subset $S \subset \mathcal{H}$ is defined as

$$S^\perp = \{ \xi \in \mathcal{H} \ : \ \xi \perp \psi \ \forall \psi \in S \}. \tag{M.51}$$

The orthogonal complement of a set is always a closed subspace of $\mathcal{H}$. A subspace $\mathcal{D} \subset \mathcal{H}$ is dense if and only if its orthogonal complement is zero, $\mathcal{D}^\perp = \{0\}$.

Given a closed subspace $\mathcal{K} \subset \mathcal{H}$, one has the orthogonal decomposition

$$\mathcal{H} = \mathcal{K} \oplus \mathcal{K}^\perp. \tag{M.52}$$

Example M.26 (Examples of Hilbert spaces).

a) $\mathcal{H} = \mathbb{C}^N$ with $\langle z, w \rangle = \sum_{k=1}^N \overline{z_k} w_k$,

b) $\mathcal{H} = \ell^2_{\mathbb{N}}$ with $\langle x, y \rangle = \sum_{n=1}^\infty \overline{x_n} y_n$,

c) $\mathcal{H} = L^2(X, \mu)$ (equivalence classes of square integrable functions) over a measure space (X, μ), with $\langle f, g \rangle = \int_X \overline{f(x)} g(x) \, d\mu(x)$.

The topological dual of a Hilbert space $\mathcal{H}$, i.e. the space of all continuous linear functionals $\mathcal{H} \to \mathbb{C}$, can be described explicitly:

Theorem M.27 (Riesz' Lemma). *For every $\xi \in \mathcal{H}$, the map*

$$\langle \xi | : \mathcal{H} \to \mathbb{C}, \qquad \psi \mapsto \langle \xi, \psi \rangle \tag{M.53}$$

is a continuous linear functional. Every continuous linear functional on $\mathcal{H}$ is of this form, with a unique $\xi \in \mathcal{H}$.

Definition M.28 (Orthonormal bases). An *orthonormal basis* (ONB) of a Hilbert space $\mathcal{H}$ is a set $(\psi_j)_{j\in J}$ indexed by some index set J that is orthonormal, namely

$$\langle \psi_j, \psi_k \rangle = \delta_{j,k}, \qquad j,k \in J, \tag{M.54}$$

and complete, namely

$$\operatorname{span}\{\psi_j \ : \ j \in J\}^\perp = \{0\}. \tag{M.55}$$

Every Hilbert space has an orthonormal basis (and in fact, many ONBs). The cardinality of the index set J is the same for all ONBs of a Hilbert space and called the *dimension* of $\mathcal{H}$. A Hilbert space is *separable* (contains a dense countable subset) if and only if its dimension is finite oder countably infinite.

Example M.29.

a) In $\mathcal{H} = \ell^2_{\mathbb{N}}$, the vectors e_n defined by $(e_n)_k = \delta_{nk}$ form an orthonormal basis $(e_n)_{n\in\mathbb{N}}$.

b) In $\mathcal{H} = L^2([0,1])$, the vectors e_n defined by $e_n(x) := e^{2\pi i n x}$, $n \in \mathbb{Z}$, form an orthonormal basis.

Lemma M.30 (Orthonormal bases). *Let $\mathcal{H}$ be a Hilbert space, $(e_j)_{j\in J}$ an orthonormal basis, and $\psi, \varphi \in \mathcal{H}$.*

a) $\psi = \sum_j \langle e_j, \psi\rangle e_j$, *with convergence in the norm of $\mathcal{H}$, independent of the order,*

b) $\|\psi\|^2 = \sum_j |\langle e_j, \psi\rangle|^2$,

c) $\langle \psi, \varphi\rangle = \sum_j \langle \psi, e_j\rangle\langle e_j, \varphi\rangle$ (*convergence independent of the order*)

We also recall some basic constructions with Hilbert spaces: Subspaces, direct sums, tensor products.

Definition M.31 (Closed subspaces, direct sums, tensor products). Let $\mathcal{H}, \mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces.

a) Let $\mathcal{K} \subset \mathcal{H}$ be a closed subspace, with orthogonal decomposition $\mathcal{H} = \mathcal{K} \oplus \mathcal{K}^\perp$ (M.52). Then the map

$$P_{\mathcal{K}} \ : \ \mathcal{H} = \mathcal{K} \oplus \mathcal{K}^\perp, \qquad k \oplus k^\perp \mapsto k \oplus 0 \tag{M.56}$$

is called the *orthogonal projection onto $\mathcal{K}$*.

b) The *direct sum* $\mathcal{H}_1 \oplus \mathcal{H}_2$ is the Hilbert space defined as the algebraic direct sum $\mathcal{H}_1 \oplus \mathcal{H}_2$ with the scalar product $\langle \psi_1 \oplus \psi_2, \xi_1 \oplus \xi_2\rangle := \langle \psi_1, \xi_1\rangle_{\mathcal{H}_1} + \langle \psi_2, \xi_2\rangle_{\mathcal{H}_2}$. (It is a Hilbert space.)

c) The (*Hilbert space*) *tensor product* of $\mathcal{H}_1$ and $\mathcal{H}_2$ is defined as follows. On the algebraic tensor product, denoted here $\mathcal{H}_1 \odot \mathcal{H}_2$, one defines a scalar product by

$$\langle \psi_1 \otimes \psi_2, \xi_1 \otimes \xi_2\rangle := \langle \psi_1, \xi_1\rangle_{\mathcal{H}_1} \cdot \langle \psi_2, \xi_2\rangle_{\mathcal{H}_2}, \tag{M.57}$$

and sesquilinear extension. Then the Hilbert space tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$

is defined as the completion of $\mathcal{H}_1 \odot \mathcal{H}_2$ w.r.t. the norm given by this scalar product.

Bounded operators on Hilbert spaces

Definition M.32 (Bounded operators). Let $\mathcal{H}$ be a Hilbert space. A *bounded operator* on $\mathcal{H}$ is a linear map $A : \mathcal{H} \to \mathcal{H}$ with finite *operator norm*

$$\|A\| := \sup_{\psi \in \mathcal{H}, \psi \neq 0} \frac{\|A\psi\|}{\|\psi\|} < \infty.$$

The set of all bounded operators on $\mathcal{H}$ is denoted $B(\mathcal{H})$.

Every operator $A \in B(\mathcal{H})$ has an *adjoint*, uniquely defined by

$$\langle \psi, A\varphi \rangle =: \langle A^*\psi, \varphi \rangle, \qquad \psi, \varphi \in \mathcal{H}. \tag{M.58}$$

The set $B(\mathcal{H})$ of bounded operators with the *-operation $A \mapsto A^*$ and the identity operator is a unital *-algebra. With the operator norm from Definition M.32 it is even a C^*-algebra, namely it is complete in operator norm and the *C^*-property*

$$\|A^*A\| = \|A\|^2, \qquad A \in B(\mathcal{H}), \tag{M.59}$$

holds.

Definition M.33 (Classes of bounded operators). $A \in B(\mathcal{H})$ is called

a) *selfadjoint* if $A = A^*$,
b) *isometry* if $A^*A = 1$,
c) *unitary* if A is invertible and $A^{-1} = A^*$,
d) *normal* if $AA^* = A^*A$,
e) *idempotent* if $A = A^2$,
f) *orthogonal projection* if $A = A^* = A^2$,
g) *partial isometry* if A^*A is an orthogonal projection,
h) *positive* if $A = B^*B$ for some $B \in B(\mathcal{H})$.

Besides the norm topology, there are several other interesting topologies on $B(\mathcal{H})$. Some of the most important ones are:

- The *strong operator topology* (SOT) is the initial topology defined by the seminorms $p_\psi(A) := \|A\psi\|$, $\psi \in \mathcal{H}$. That means in particular that a net (A_j) of bounded operators converges to $A \in B(\mathcal{H})$ if and only if $\lim_j \|A_j\psi - A\psi\| = 0$.
- The *weak operator topology* (WOT) is the initial topology defined by the seminorms $p_{\psi,\varphi}(A) := |\langle \psi, A\varphi \rangle|$, $\psi, \varphi \in \mathcal{H}$. That means in particular that a net (A_j) of bounded operators converges to $A \in B(\mathcal{H})$ if and only if $\lim_j \langle \psi, (A_j - A)\varphi \rangle = 0$.
- The *ultraweak topology* (or σ-weak topology) is the initial topology defined by the seminorms $p_T(A) := \mathrm{Tr}(TA)$, where T runs over the space of all trace class operators.

Unbounded Hilbert space operators

We now consider *unbounded operators*.

Basic definitions. Closed and closable operators

Definition M.34.

a) Let $\mathcal{H}$ be a Hilbert space. An *operator* is a pair $(\operatorname{dom} A, A)$ consisting of a subspace $\operatorname{dom} A \subset \mathcal{H}$, the *domain* of A, and a linear map $A : \operatorname{dom} A \to \mathcal{H}$.

b) The *range* of an operator $(\operatorname{dom} A, A)$ is $\operatorname{Ran} A := \{Av : v \in \operatorname{dom} A\}$, and the *kernel* is $\ker A := \{v \in \operatorname{dom} A : Av = 0\}$.

c) If $\ker A = \{0\}$, we define the *inverse* of A by $\operatorname{dom}(A^{-1}) := \operatorname{Ran} A$ and $A^{-1} : Av \mapsto v$. We call A *invertible* in this case.

d) An operator $(\operatorname{dom} A, A)$ is called *extension* of an operator $(\operatorname{dom} B, B)$ if $\operatorname{dom} A \supset \operatorname{dom} B$ and $A|_{\operatorname{dom} B} = B$. This is also written briefly as $B \subset A$.

e) Two operators $(\operatorname{dom} A, A)$ and $(\operatorname{dom} B, B)$ are *equal*, written as $A = B$, if $A \subset B$ and $B \subset A$, i.e. if $\operatorname{dom} A = \operatorname{dom} B$ and $Av = Bv$ for all $v \in \operatorname{dom} A$.

Of course any $A \in \mathcal{B}(\mathcal{H})$ is an operator with domain $\operatorname{dom} A = \mathcal{H}$. The *operator norm* of an operator is defined in essentially the same way as for $\mathcal{B}(\mathcal{H})$, namely

$$\|A\| := \sup_{v \in \operatorname{dom} A, v \neq 0} \frac{\|Av\|}{\|v\|} \in [0, +\infty]. \tag{M.60}$$

In case $\|A\| = \infty$, we say that A is *unbounded*. Often the term "unbounded operator" is also used to mean "not necessarily bounded operator", i.e. including the bounded case.

We will often write A instead of $(\operatorname{dom} A, A)$ in case the domain is clear from the context or not relevant. Nonetheless, the domain of an operator is a crucial part of the definition of A.
Remarks:

a) In almost all cases we will consider operators that are *densely defined*, i.e. with the domain being a dense subspace of $\mathcal{H}$.

b) With obvious modifications, one can define operators $A : \operatorname{dom} A \subset \mathcal{H}_1 \to \mathcal{H}_2$ between two different Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$. We then say that A is an operator from $\mathcal{H}_1$ to $\mathcal{H}_2$, meaning that it is a map from a subspace of $\mathcal{H}_1$ into $\mathcal{H}_2$.

c) In case an operator A is bounded, $\|A\| < \infty$, the map $A : \operatorname{dom} A \to \mathcal{H}$ is continuous. In this case A has a unique continuous extension to $\overline{\operatorname{dom} A}$. In the case of a dense domain, this extension is defined on the full Hilbert space and yields an element of $\mathcal{B}(\mathcal{H})$.

d) Every operator, bounded or not, has extensions to all of $\mathcal{H}$. The existence of these extensions can be seen by appealing to the axiom of choice and can usually not be specified concretely.

Whereas the set $\mathcal{B}(\mathcal{H})$ of all *bounded* operators is a C^*-algebra, the set of all unbounded operators has no good properties (we did not even give this set a name or symbol). By definition, the operator norm is not a norm on unbounded operators (it may be infinite). We have not yet introduced a *-operation, but will soon see that this is possible for densely defined operators. However, the basic algebraic operations of taking sums and products of two operators $(\operatorname{dom} A, A)$ and $(\operatorname{dom} B, B)$ are quite subtle in the unbounded case: They are naturally defined as

$$(\operatorname{dom} A, A) + (\operatorname{dom} B, B) := (\operatorname{dom} A \cap \operatorname{dom} B, (A + B)|_{\operatorname{dom} A \cap \operatorname{dom} B}), \tag{M.61}$$

$$(\operatorname{dom} A, A) \cdot (\operatorname{dom} B, B) := (\operatorname{dom} B \cap B^{-1}(\operatorname{dom} A), AB|_{\operatorname{dom} B \cap B^{-1}(\operatorname{dom} A)}). \tag{M.62}$$

We have an identity $1 = (\mathcal{H}, \mathrm{id}_{\mathcal{H}})$, and scalar multipliction $\lambda \cdot A := \lambda 1 \cdot A$ works as expected. Inserting the definitions, one also easily checks that the sum is commutative and associative, and the product of operators is associative. Furthermore, $(A + B)C = AC + BC$. However, $A(B + C) = AB + AC$ *fails* in general. Thus the set of all unbounded operators is not an algebra.

As a further warning, note that the sum of two densely defined operators A, B can have domain $\mathrm{dom}(A+B) = \{0\}$, and the set of all unbounded operators is not even a vector space.

For now, we recall the concept of a closed map from functional analysis [FA1,Def. 3.31], which will be helpful to select better behaved but still unbounded operators.

In the context of a direct sum $\mathcal{H} \oplus \mathcal{H}$, we write the vectors as either pairs (v, w) or sums $v \oplus w$ and equip it with the direct sum scalar product and corresponding norm $\|v \oplus w\|^2 = \|v\|^2 + \|w\|^2$. The following discussion works in the same way when considering two Banach spaces X, Y, but we restrict to the case where $X = Y = \mathcal{H}$ is a Hilbert space.

Definition M.35. Let $\mathcal{H}$ be a Hilbert space and $(\mathrm{dom}\, A, A)$ on operator on $\mathcal{H}$.

a) The *graph* of A is

$$\Gamma(A) := \{(v, Av) \ : \ v \in \mathrm{dom}\, A\} \subset \mathcal{H} \oplus \mathcal{H}. \tag{M.63}$$

b) An operator $(\mathrm{dom}\, A, A)$ is called *closed* if $\Gamma(A)$ is a closed subspace of $\mathcal{H} \oplus \mathcal{H}$.

c) An operator $(\mathrm{dom}\, A, A)$ is called *closable* if it has a closed extension.

Proposition M.36. *Let* $(\mathrm{dom}\, A, A)$ *be an operator on a Hilbert space* $\mathcal{H}$.

a) The following are equivalent:

(i) *A is closed.*

(ii) *If a sequence $(v_n)_n \subset \mathrm{dom}\, A$ converges to $v \in \mathcal{H}$ and $(Av_n)_n \subset \mathcal{H}$ converges to $w \in \mathcal{H}$, then $v \in \mathrm{dom}\, A$ and $Av = w$.*

(iii) $\mathrm{dom}\, A$ *is complete in the* graph norm *induced by the* graph scalar product $\langle v, w\rangle_A := \langle v, w\rangle + \langle Av, Aw\rangle$, $v, w \in \mathrm{dom}\, A$, *i.e.* $(\mathrm{dom}\, A, \langle \cdot\,, \cdot \rangle_A)$ *is a Hilbert space.*

b) Bounded operators are closed.

c) If $\mathrm{dom}\, A \subset \mathcal{H}$ *is closed, then* $(\mathrm{dom}\, A, A)$ *is closed if and only if it is bounded.*

Note that in terms of graphs, we have for any two operators A, B

$$A \subset B \iff \Gamma(A) \subset \Gamma(B). \tag{M.64}$$

With this, one can give a similar characterization of closable operators:

Proposition M.37. *Let* $(\mathrm{dom}\, A, A)$ *be an operator on a Hilbert space* $\mathcal{H}$. *The following are equivalent:*

(i) *A is closable.*

(ii) *If a sequence $(v_n)_n \subset \mathrm{dom}\, A$ converges to $0 \in \mathcal{H}$ and $(Av_n)_n \subset \mathcal{H}$ converges to $w \in \mathcal{H}$, then $w = 0$.*

(iii) *The closure $\overline{\Gamma(A)}$ is the graph of a (unique) operator.*

Definition M.38. Let A be closable. The unique operator $\overline{A}$ satisfying $\overline{\Gamma(A)} = \Gamma(\overline{A})$ is called the *closure* of A.

Note that $\operatorname{dom}\overline{A}$ consists of all vectors $v \in \mathcal{H}$ such that there exists a sequence $(v_n) \subset \operatorname{dom} A$ with $v_n \to v$ and such that $(Av_n)_n$ converges. The closure is then given by $\overline{A}v := \lim_n Av_n$. This is clearly the smallest closed extension of A.

Adjoints and (essential) selfadjointness

We now extend the concept of the adjoint $A \mapsto A^*$ to unbounded operators. Guided by the defining equation $\langle v, Aw\rangle = \langle A^*v, w\rangle$ (which now makes sense only for $w \in \operatorname{dom} A$), we see that A^*v will be uniquely defined only if $\operatorname{dom} A$ is dense.

Definition M.39. Let $(\operatorname{dom} A, A)$ be a densely defined operator. The *adjoint* $(\operatorname{dom} A^*, A^*)$ is defined as

$$\operatorname{dom} A^* := \{v \in \mathcal{H} \;:\; \exists c > 0 \;:\; |\langle v, Aw\rangle| \le c\|w\| \;\forall w \in \operatorname{dom} A\} \tag{M.65}$$

and $A^* : \operatorname{dom} A^* \to \mathcal{H}$, $A^*v := x_v$, where $x_v \in \mathcal{H}$ is the unique vector satisfying $\langle v, Aw\rangle = \langle x_v, w\rangle$.

The following definition will be crucial for the spectral theorem.

Definition M.40. Let $(\operatorname{dom} A, A)$ be densely defined.

a) A is called *symmetric* if $\langle v, Aw\rangle = \langle Av, w\rangle$ holds for all $v, w \in \operatorname{dom} A$. Equivalently $A \subset A^*$.

b) A is called *selfadjoint* if $A = A^*$ (i.e. $A \subset A^*$ and $A^* \subset A$).

Let us demonstrate that in the unbounded case, being symmetric does not imply being selfadjoint[18].

Example M.41. Consider $\mathcal{H} = L^2(\mathbb{R})$ and the multipliction operator $(Xf)(x) = xf(x)$ with domain $\operatorname{dom} X = C_c(\mathbb{R})$. Then $(\operatorname{dom} X, X)$ is clearly symmetric, but not selfadjoint. To see that, consider $f \in C_c(\mathbb{R})$ and $g(x) = (1+x^2)^{-1}$, so that $g \in L^2(\mathbb{R})$, but $g \notin \operatorname{dom} X$. Since $\int x^2|g(x)|^2dx < \infty$, the estimate

$$|\langle g, Xf\rangle| \le \int_{\mathbb{R}} |g(x)|x|f(x)|dx \le \left(\int_{\mathbb{R}} \frac{x^2}{(1+x^2)^2}dx\right)^{1/2} \cdot \|f\|_{L^2}$$

shows $g \in \operatorname{dom} X^*$. Hence $\operatorname{dom} X \neq \operatorname{dom} X^*$, so X is not selfadjoint.

A number of properties of adjoints that are familiar from bounded operators are generalized to the unbounded case in the following simple lemma.

[18] Note, however, if A is a symmetric operator with $\operatorname{dom} A = \mathcal{H}$, then it is bounded by the Hellinger-Toeplitz Theorem [Korollar 3.35, FA1], and hence selfadjoint.

Lemma M.42. *Let A be densely defined.*

a) A^* *is closed.*
b) $(\operatorname{Ran} A)^\perp = \ker(A^*)$.
c) If A^* *is densely defined, then* $A \subset A^{**} := (A^*)^*$.
d) If B *is another operator with* $A \subset B$, *then* $B^* \subset A^*$.
e) $(\lambda A)^* = \overline{\lambda} A^*$.
f) If B *is another operator s.t.* $A + B$ *is densely defined, then* $(A + B)^* \supset A^* + B^*$.

A simple but important corollary is that because adjoints of densely defined operators are always closed, symmetric operators ($A \subset A^*$) have closed extensions.

Corollary M.43. *Symmetric operators are closable and selfadjoint operators are closed.*

We now come to the main result about closable operators and adjoints.

Theorem M.44. *A densely defined operator A is closable if and only if* $\operatorname{dom} A^*$ *is dense. In this case* $A^{**} = \overline{A}$ *and* $(\overline{A})^* = A^*$.

We note that the adjoint A^* of a densely defined operator needs not to be densely defined.

In general, it is difficult to determine whether a symmetric operator ($A \subset A^*$) is selfadjoint ($A = A^*$). Since selfadjointness will be crucial for the spectral theorem, we need to adress this question.

The idea is to consider selfadjoint extensions of symmetric operators, possibly without explicitly determining the domain of the extension. If A is symmetric and B an extension, we have $A \subset A^*$ and $A \subset B \Rightarrow B^* \subset A^*$, but it is not clear whether B can be chosen such that $B = B^*$, i.e. selfadjoint. There exist symmetric operators that have no selfadjoint extension. On the other hand, a symmetric operator might have several different selfadjoint extensions (see below). The best situation arises when we have a unique selfadjoint extension.

Definition M.45. A symmetric operator A is called *essentially selfadjoint* if $\overline{A}$ is selfadjoint.

Lemma M.46. *An essentially selfadjoint operator A has a unique selfadjoint extension, namely* $\overline{A}$.

The following theorem is the basic criterion for checking (essential) selfadjointness.

Theorem M.47. *Let A be a symmetric operator. Then the following are equivalent:*

a) A is selfadjoint.
b) A is closed and $\ker(A^* \pm i) = \{0\}$.
c) $\operatorname{Ran}(A \pm i) = \mathcal{H}$.

b) and c) are meant to hold for both *signs* $\pm$.

Corollary M.48. *Let A be a symmetric operator. Then the following are equivalent:*

a) A is essentially selfadjoint.
b) $\ker(A^* \pm i) = \{0\}$.

c) $\mathrm{Ran}(A \pm i)$ *are dense.*

b) and c) are meant to hold for both *signs* $\pm$.

Spectra of unbounded operators

We will now develop spectral theory for unbounded (closed) operators.

Definition M.49. Let A be a closed operator. The *resolvent set* of A is

$$\rho(A) := \{\lambda \in \mathbb{C} \,:\, (A-\lambda) : \mathrm{dom}\, A \to \mathcal{H} \text{ is bijective and } \|(A-\lambda)^{-1}\| < \infty\},$$

and the *spectrum* of A is

$$\sigma(A) := \mathbb{C} \backslash \rho(A). \tag{M.66}$$

The *resolvent* of A at $\lambda \in \rho(A)$ is $R_\lambda(A) := (A-\lambda)^{-1} : \mathcal{H} \to \mathrm{dom}\, A$.

This definition applies in particular to bounded operators defined on all of $\mathcal{H}$, but also works more generally for unbounded closed operators. One can show that for non-closed operators, this definition of spectrum always yields $\sigma(A) = \mathbb{C}$ and is therefore meaningless. One can also show that the boundedness of the resolvent, $\|(A-\lambda)^{-1}\| < \infty$, $\lambda \in \rho(A)$, is automatically fulfilled (use the closed graph theorem).

Proposition M.50. *Let A be a symmetric closed operator. Then $\sigma(A)$ is either a subset of $\mathbb{R}$, or one of the three sets $\overline{\mathbb{C}_+}$ (closed upper halfplane in $\mathbb{C}$), $\overline{\mathbb{C}_-}$ (closed lower halfplane in $\mathbb{C}$), or $\mathbb{C}$.*

The operator is selfadjoint if and only if $\sigma(A) \subset \mathbb{R}$.

This result explains why selfadjoint operators are distinguished within the set of all symmetric operators. For a closed symmetric operator A that is not selfadjoint, the spectrum contains hardly any information about A, in contrast to the selfadjoint case.

Selfadjoint operators are furthermore characterized by the fact that they admit a spectral decomposition. The main tool for understanding these are projection-valued measures.

Definition M.51. A *projection-valued measure* on a mesurable space (X, Σ) is a map

$$E : \Sigma \to \mathcal{P}(\mathcal{H}) \tag{M.67}$$

satisfying

a) $E(\emptyset) = 0$ and $E(X) = 1$,

b) if J is a countable index set and $S_j \in \Sigma$, $j \in J$, are pairwise disjoint measurable sets, then

$$E\Big(\bigcup_j S_j\Big) = \sum_j E(S_j) \tag{M.68}$$

with SOT-convergence of the right hand side.

Observe that a projection-valued measure E on $\mathcal{H}$ and a vector $v \in \mathcal{H}$ define a *finite* measure

$$E_v : \Sigma \to [0, \infty), \qquad S \mapsto E_v(S) := \|E(S)v\|^2 \leq \|v\|^2. \tag{M.69}$$

We also note that for a projection-valued measure E, one has $E(S)E(S') = E(S')E(S)$ for all measurable sets S, S'. In the following, we will mostly consider projection-valued measures on the Borel σ-algebra of $X = \mathbb{R}, X = \mathbb{C}$, or $X = \mathbb{R}^n$ and briefly refer to these as "projection-valued measures on $\mathbb{R}, \mathbb{C}, \mathbb{R}^n$".

One can integrate w.r.t. projection-valued measures:

Proposition M.52. *Let E be a projection-valued measure on a measurable space (X, Σ). Then*

$$I_E : \mathfrak{B}(X) \to \mathcal{B}(\mathcal{H}), \qquad I_E(f) := \int_X f(x)\, dE(x), \tag{M.70}$$

is a homomorphism of unital C^*-algebras. It satisfies

$$\langle v, I_E(f)v\rangle = \int f(x)\, dE_v(x), \tag{M.71}$$

$$\|I_E(f)v\|^2 = \int_X |f(x)|^2\, dE_v(x) = \|f\|^2_{L^2(X,E_v)}, \qquad v \in \mathcal{H}. \tag{M.72}$$

If $(f_n) \subset \mathfrak{B}(X)$ is a uniformly bounded sequence that converges E-almost everywhere to a function $f : X \to \mathbb{C}$, then $I_E(f_n) \to I_E(f)$ in SOT.

Note that as a homomorphism, I_E is automatically contractive, i.e. $\|I_E(f)\| \leq \|f\|_\infty$. It is however usually not injective because we did not quotient by zero sets.

As a first step towards the spectral theorem, we will consider a given projection-valued measure E and extend the map I_E from $\mathfrak{B}(X)$ to arbitrary measurable functions. For unbounded f, we do not expect $I_E(f)$ to be bounded, so that we have to take care of domains.

The setting that we are using is the following: (X, Σ) is a measurable space and $E : \Sigma \to \mathcal{P}(\mathcal{H})$ a projection-valued measure. We write $\mathcal{M}(X)$ the set of all measurable functions $f : X \to \mathbb{C}$.

Definition and Proposition M.53. Let E be a projection-valued measure on a measurable space (X, Σ) and $f \in \mathcal{M}(X)$ a measurable function.

a) The set

$$\operatorname{dom} I_E(f) := \{v \in \mathcal{H} \ : \ \int_X |f|^2\, dE_v < \infty\} \tag{M.73}$$

is a dense linear subspace of $\mathcal{H}$.

b) A vector $v \in \mathcal{H}$ lies in $\operatorname{dom} I_E(f)$ if and only if $(I_E(f\chi_n)v)_n$ converges in $\mathcal{H}$ if and only if $(I_E(f\chi_n)v)_n$ is bounded in $\mathcal{H}$, where χ_n is the characteristic function of $S_n := \{|f| \leq n\}$.

c) The assignment

$$I_E(f) : \operatorname{dom} I_E(f) \to \mathcal{H}, \qquad I_E(f)v := \lim_{n\to\infty} I_E(f\chi_n)v, \tag{M.74}$$

defines an operator $I_E(f)$ which satisfies

$$\langle v, I_E(f)v\rangle = \int f(x)\, dE_v(x), \tag{M.75}$$

$$\|I_E(f)v\| = \|f\|_{L^2(X,E_v)}, \qquad v \in \operatorname{dom} I_E(f). \tag{M.76}$$

The operator constructed in this proposition is usually denoted

$$\int f(x)\, dE(x) := I_E(f). \tag{M.77}$$

The following lemma contains further information about how to work with the operators $I_E(f)$.

Lemma M.54. *Let (X,Σ) be a measurable space, $E : \Sigma \to \mathcal{P}(\mathcal{H})$ a projection-valued measure, and $f \in \mathcal{M}(X)$.*

a) If f is bounded, then $\operatorname{dom} I_E(f) = \mathcal{H}$.

b) If f is essentially bounded (i.e. if f is bounded on the complement of an E-zero set), then $I_E(f)$ is bounded[a].

c) Let $S \in \Sigma$. Then $E(S)\operatorname{dom} I_E(f) \subset \operatorname{dom} I_E(f) \subset \operatorname{dom} I_E(\chi_S f)$, and

$$I_E(f)E(S)v = E(S)I_E(f)v = I_E(\chi_S f)v, \qquad v \in \operatorname{dom} I_E(f). \tag{M.78}$$

[a] In this case, one can also show $\|I_E(f)\|_{\mathcal{B}(\mathcal{H})} = \|f\|_\infty$.

We have now extended integration of measurable functions against projection-valued measures to unbounded measurable functions. The next proposition shows that the map $f \mapsto I_E(f)$ still behaves like a functional calculus with minor modifications due to the unboundedness of the operators involved.

Proposition M.55. *Let (X,Σ) be a measurable space, $E : \Sigma \to \mathcal{P}(\mathcal{H})$ a projection-valued measure, and $f, g \in \mathcal{M}(X)$.*

a) $I_E(\overline{f}) = I_E(f)^*$. In particular, $I_E(f)$ is selfadjoint for real-valued f.

b) $I_E(\lambda f + \mu g) = \overline{\lambda I_E(f) + \mu I_E(g)}$ for $\lambda, \mu \in \mathbb{C}$.

c) $I_E(fg) = \overline{I_E(f)I_E(g)}$ and $\operatorname{dom}(I_E(f)I_E(g)) = \operatorname{dom} I_E(g) \cap \operatorname{dom} I_E(fg)$.

d) $I_E(f)$ is closed and normal in the sense that $I_E(f)^* I_E(f) = I_E(\bar{f}f) = I_E(f)I_E(f)^*$.

Summarizing the developments so far, we have seen that a projection-valued measure gives rise to the operators $I_E(f)$ which are selfadjoint for real f, including in particular $A := I_E(\mathrm{id}) = \int \lambda\, dE(\lambda)$. These operators share the homomorphism properties of the measurable functional calculus, properly generalized to the unbounded setting.

The spectral theorem proceeds in the opposite direction, and gets a measure from a selfadjoint operator.

Theorem M.56 (The spectral theorem). *For any selfadjoint operator A there exists a unique projection-valued measure E^A on $\mathbb{R}$ such that $A = I_{E^A}(\mathrm{id})$, i.e.*

$$A = \int \lambda \, dE^A(\lambda), \qquad \mathrm{dom}\, A = \{v \in \mathcal{H} \ : \ \int \lambda^2 dE^A_v(\lambda) < \infty\}. \tag{M.79}$$

The support of this measure (the complement of the union of all open sets $O \subset \mathbb{R}$ with measure zero, $E(O) = 0$) is the spectrum $\sigma(A)$ of A.

Unitary one-parameter groups

In this section we will discuss Stone's Theorem about unitary one-parameter groups which emphasizes the importance of selfadjoint operators and has applications in many fields.

Definition M.57. A *unitary one-parameter group* U is a homomorphism from the additive group $(\mathbb{R}, +)$ into the group $\mathcal{U}(\mathcal{H})$ of unitary operators on a Hilbert space $\mathcal{H}$. That is, for each $t \in \mathbb{R}$ there is a unitary $U(t) \in \mathcal{U}(\mathcal{H})$ and

$$U(t)U(s) = U(t+s), \qquad t, s \in \mathbb{R}. \tag{M.80}$$

Note that since group homomorphisms respect neutral and inverse elements, a unitary one-parameter group U automatically satisfies also $U(0) = 1$ and $U(-t) = U(t)^{-1} = U(t)^*$, $t \in \mathbb{R}$.

We will mostly be interested in strongly continuous unitary one-parameter groups, namely those for which $\mathbb{R} \ni t \mapsto U(t) \in B(\mathcal{H})$ is continuous with the strong operator topology on $B(\mathcal{H})$. This means precisely that for any $\psi \in \mathcal{H}$, the map

$$\mathbb{R} \ni t \mapsto \psi(t) := U(t)\psi \in \mathcal{H} \tag{M.81}$$

is continuous in the norm topology of $\mathcal{H}$

Definition M.58. Let U be a unitary one-parameter group. A vector $\psi \in \mathcal{H}$ is called continuous (or continuously differentiable, or smooth) (w.r.t. U) if (M.81) is a continuous (or continuously differentiable, or smooth) function (w.r.t. the norm topology of $\mathcal{H}$).

Remarks:

a) A unitary one-parameter group is strongly continuous if and only if every $\psi \in \mathcal{H}$ is continuous w.r.t. U if and only if

$$\lim_{t \to 0} \|U(t)\psi - \psi\| = 0, \qquad \psi \in \mathcal{H}. \tag{M.82}$$

The necessity of this is clear, as it coincides with continuity at $t = 0$, and the sufficiency follows from $\|U(t)\psi - U(s)\psi\| = \|U(-s)U(t)\psi - \psi\| = \|U(t-s)\psi - \psi\|$, using the unitarity and the group properties.
One can easily show that it is enough to verify (M.82) for all ψ in a dense subspace of $\mathcal{H}$ to conclude the strong continuity of U (exercise).

b) One does not consider *weakly* continuous unitary one-parameter groups because these are automatically strongly continuous (easy exercise).

c) If $t \mapsto U(t)\psi$ is differentiable at $t = 0$, with derivative $\psi'(0)$, then it is differentiable everywhere, with $\psi'(t) = U(t)\psi'(0)$:

$$\begin{aligned}\left\|\frac{1}{h}\left(U(t+h)\psi - U(t)\psi\right) - U(t)\psi'(0)\right\| &= \left\|U(t)\left(\frac{1}{h}\left(U(h)\psi - \psi\right) - \psi'(0)\right)\right\| \\ &= \left\|\frac{1}{h}\left(U(h)\psi - \psi\right) - \psi'(0)\right\| = 0.\end{aligned}$$

In particular, ψ is automatically continuously differentiable in this case.

The following structure theorem about strongly continuous unitary one-parameter groups generalizes the last example above to general (not necessarily bounded) selfadjoint generators.

Theorem M.59 (Stone's Theorem).

a) *Let* $(\operatorname{dom} A, A)$ *be a selfadjoint operator on a Hilbert space* $\mathcal{H}$*. Then* $U(t) := e^{itA} = \int e^{it\lambda} dE^A(\lambda)$ *is a strongly continuous unitary one-parameter group. Moreover, a vector* $\psi \in \mathcal{H}$ *lies in* $\operatorname{dom} A$ *if and only if* ψ *is differentiable w.r.t.* U*. In this case,* $\psi'(0) = iA\psi$*.*

b) *Every strongly continuous unitary one-parameter group* $U(t)$ *on* $\mathcal{H}$ *is of the form described in a), with some selfadjoint operator* A *(called the generator of* $U(t)$*) uniquely determined by* $U(t)$*,* $t \in \mathbb{R}$*.*

List of Symbols

$\{\cdot,\cdot\}$	canonical Poisson bracket on $C^\infty(\mathbb{R}^{2d})$	13
$[\cdot,\cdot]$	commutator	43
$\mathcal{A}_{\mathrm{cl}}$	$\mathcal{A}_{\mathrm{cl}} = C^\infty(M)$, classical observable algebra of Ham. system (M,H) (Def. 1.13)	12
α^H	Hamiltonian dynamics on $C^\infty(M)$ generated by Hamiltonian H (Thm. 1.16)	13
$\mathrm{Aut}\,\mathcal{A}$	automorphism group of unital *-algebra $\mathcal{A}$	58
$\mathbb{b}_\omega$	Bell correlation in state ω	91
$C_0(\mathbb{R}^d)$	continuous functions vanishing at infinity (M.43)	179
$C_c^\infty(\mathbb{R}^d)$	smooth functions of compact support	106
CCR	CCR algebra over symplectic space $(\mathbb{R}^{2d},\sigma)$ (Def. 3.1)	106
$C^\infty_{\mathbb{R}}(U)$	smooth functions $U \to \mathbb{R}$	7
$C^\infty(U)$	smooth functions $U \to \mathbb{C}$	7
Cov_ω	covariance of state ω (Def. 1.33)	27, 45
Corr_ω	correlation of state ω (Def. 1.33)	27, 45
$C_{\mathbb{R}}(U)$	continuous functions $U \to \mathbb{R}$	7
$C(U)$	continuous functions $U \to \mathbb{C}$	7
$\mathcal{D}_N$	space of density matrices in $M_N(\mathbb{C})$	47
$\mathrm{Diff}(\mathcal{O})$	diffeomorphism group of $\mathcal{O}$	6
F	force field $F \in \mathcal{X}(U)$	4
ϕ_H	Hamiltonian flow, flow of of Hamiltonian vector field of H (Def. 1.12)	11
$\phi_{\mathcal{V}}$	flow of vector field $\mathcal{V}$ (1.15)	6
$\mathcal{F}$	Fourier transform (M.42)	179
H	Hamiltonian function (Def. 1.8) or operator	9
$\hbar$	Planck's constant	42
L	angular momentum observable (classical: Def. 1.25)	19
$L^2(\mathbb{R}^d)$	square integrable functions (w.r.t. Lebesgue measure)	106
ℓ	Lebesgue measure	49
m	mass	3
M	phase space $M = U \times \mathbb{R}^d$ (1.5)	4
(M,H)	classical Ham. system with phase space M and Ham. function H (Def. 1.13)	12
$M_N(\mathbb{C})$	algebra of complex $(N \times N)$-matrices	43
ω	state (expectation value functional) (Def. 2.5)	44
p	momentum $p = m\dot{x}$, momentum space variable	3
P	momentum observable (classical: (1.39))	12
$P_*\mu$	probability measure of position P under phase space measure μ (Def. 1.28)	23
$\mathrm{Pol}(\mathbb{R}^d)$	polynomial functions $\mathbb{R}^d \to \mathbb{C}$	106
$P_{\mathrm{Ran}\,\rho}$	support projection of density matrix ρ	49
$\mathrm{PU}(N)$	projective unitary group	60
$\mathcal{P}(X)$	set of Borel probability measures on X (Def. M.1)	170
$\mathcal{P}_c(X)$	set of compactly support Borel probability measures on X (Def. M.1)	170
ρ	density matrix (Def. 2.10)	47
$\varphi_*\mu$	pushforward measure of μ by φ (Def. M.5)	171
$\rho\mu$	measure with density ρ relative to μ (Def. M.7)	172
$\sigma(f)$	$\sigma(f) = \overline{f(M)}$ spectrum of classical observable $f \in C^\infty(M)$	20

$S(\mathcal{A})$	state space of unital *-algebra $\mathcal{A}$ (Def. 2.5)	44
$\mathcal{S}(\mathbb{R}^d)$	Schwartz functions	106
σ_k	Pauli Matrix (Example 2.38)	74
Σ	canonical symplectic matrix on $\mathbb{R}^{2d}$ (1.33)	11
$S_{\text{pure}}(\mathcal{A})$	set of pure states of $\mathcal{A}$ (Def. 2.6)	44
SU(2)	special unitary group	77
$\mathbb{T}$	unit circle in complex plane	60
Tr	matrix trace on $M_N(\mathbb{C})$	176
$U(N)$	unitary group	58
V	potential (Def. 1.5)	7
$\mathcal{V}$	general vector field $\mathcal{V} \in \mathcal{X}(\mathcal{O})$	5
$\mathcal{V}_F$	vector field defined by force F (1.8)	4
$\{V \leq V_0\}$	sublevel set of potential V (1.30)	9
W	Weyl Quantization (Def. 3.9)	111
x	position, configuration space variable	3
X	position observable (classical: (1.39))	12
$\dot{x}$	velocity, time derivative of curve $x : I \to \mathbb{R}^d$	3
$X_*\mu$	probability measure of position X under phase space measure μ (Def. 1.28)	23
$\mathcal{X}(U)$	space of smooth vector fields on an open subset $U \subset \mathbb{R}^d$ (1.6)	4
$\ddot{x}$	acceleration, second time derivative of curve $x : I \to \mathbb{R}^d$	3
ξ	$\xi = (x, p)$, phase space variable	3

Index

References

[AM87] R. Abraham and J. E. Marsden. *Foundations of Mechanics*. 2. Addison-Wesley Publishing Company, 1987.

[Arn89] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer, 1989.

[BLM96] P. Busch, P. J. Lahti, and P. Mittelstaedt. *The Quantum Theory of Measurement*. Springer, 1996.

[BW00] H. Bursztyn and S. Waldmann. "Algebraic Rieffel Induction, Formal Morita Equivalence, and Applications to Deformation Quantization". *J. Geom. Phys.* 37.4 (2000), 307–364.

[Chi99] C. Chicone. *Ordinary Differential Equations with Applications*. Vol. 34. Texts in Applied Mathematics. Springer, 1999.

[Els18] J. Elstrodt. *Maß- und Integrationstheorie*. 8th ed. Springer, 2018.

[GP90] A. Galindo and P. Pascual. *Quantum Mechanics I*. Springer, 1990.

[Hal03] B. Hall. *Lie Groups, Lie Algebras, and Representations. An Elementary Introduction*. Springer, 2003.

[Hal13] B. Hall. *Quantum Theory for Mathematicians*. Springer, 2013.

[Heu95] H. Heuser. *Gewöhnliche Differentialgleichungen*. Teubner, 1995.

[Hor+09] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. "Quantum entanglement". *Rev.Mod.Phys.* 81 (Apr. 2009), 865–942. DOI: 10.1103/RevModPhys.81.865.

[Lan17] K. Landsman. *Foundations of Quantum Theory. From Classical Concepts to Operator Algebras.* Fundamental Theories of Physics. Springer, 2017.

[RS72] M. Reed and B. Simon. *Methods of Modern Mathematical Physics I - Functional Analysis*. Academic Press, 1972.

[RS75] M. Reed and B. Simon. *Methods of Modern Mathematical Physics II - Fourier Analysis*. Academic Press, 1975.

[Sch12] K. Schmüdgen. *Unbounded Self-Adjoint Operators on Hilbert Space*. Springer, 2012.

[Tak08] L. A. Takhtajan. *Quantum Mechanics for Mathematicians*. AMS, 2008.

[Tes09] G. Teschl. *Mathematical Methods of Quantum Mechanics. With Applications to Schrödinger Operators.* Amer. Math. Soc., 2009.

[Var06] V. S. Varadarajan. *Geometry of quantum theory*. 2nd ed. New York, NY: Springer, Dec. 2006.

[Wal07] S. Waldmann. *Poisson-Geometrie und Deformationsquantisierung*. Springer, 2007.

[WW01] R. F. Werner and M. M. Wolf. *Bell inequalities and Entanglement*. 2001.