

# Functional analysis and quantum mechanics: an introduction for physicists

Kedar S. Ranade\*

Received 4 May 2015, revised 4 May 2015, accepted 8 June 2015

Published online 20 August 2015

We give an introduction to certain topics from functional analysis which are relevant for physics in general and in particular for quantum mechanics. Starting from some examples, we discuss the theory of Hilbert spaces, spectral theory of unbounded operators, distributions and their applications and present some facts from operator algebras. We do not give proofs, but present examples and analogies from physics which should be useful to get a feeling for the topics considered.

## 1 Introduction

It is well-known that physics and mathematics are closely interconnected sciences. New results in physics lead to new branches of mathematics, and new concepts from mathematics are used to describe physical situations and to predict new physical results. In the study of physics quantum mechanics plays a fundamental role for almost all fields of modern physics: from quantum optics, quantum information to solid-state physics, high-energy physics, particle physics etc. A similar role is enjoyed in mathematics by functional analysis in the theory of differential equations, numerical methods and elsewhere.

While every physicist knows quantum mechanics and every mathematician knows functional analysis, they often do not know the connection between these two fields—both of which were founded at the beginning of the 20<sup>th</sup> century. While every student of physics attends lectures on higher mathematics, such as analysis of one or several variables, linear algebra, differential equations and complex analysis, it is not so common that he gets to know functional analysis. Though it is possible to understand the basics of quantum mechanics with pure linear algebra, a deeper understanding is gained by knowing at least the very structure of functional analysis.

It is the intention of this tutorial to introduce basic concepts of functional analysis to physicists and to point out their influence on quantum mechanics. We aim at mathematical rigour in terminology, but we will leave out proofs. Our approach is not necessarily consistent with a mathematics textbook building up from bottom to top by proving every theorem from some basic axioms. Rather, we strive at an intuitive understanding from concepts and analogies known either from basic linear algebra or used in basic quantum mechanics.

Prerequisites for reading this tutorial are a knowledge of textbook quantum mechanics (such as wave mechanics and Dirac notation) including some basic quantum optics (harmonic oscillator, ladder operators, coherent states etc.), but excluding quantum field theory; from mathematics, knowledge of analysis and linear algebra is required.<sup>1</sup>

## 2 Hilbert space theory

The formalism of quantum mechanics uses the concept of a *Hilbert space*. In this section we will elaborate on the theory and give a full classification of Hilbert spaces, i. e. we list (in some sense) all possible Hilbert spaces. The mathematics covered here can be found in several textbooks on functional analysis, e. g. [1–3].

<sup>1</sup> This manuscript is partially based on a series of lectures entitled *Mathematische Aspekte der Quantenmechanik* presented (in German) at the Institut für Quantenphysik, Universität Ulm, from April to August 2014.

Institut für Quantenphysik, Universität Ulm, and Center for Integrated Quantum Science and Technology (IQ<sup>ST</sup>), Albert-Einstein-Allee 11, D-89081 Ulm, Deutschland, Germany

\* Corresponding author: E-mail: Kedar.Ranade@uni-ulm.de, Phone: +49/731/50-22783

## 2.1 Introductory examples

We start this tutorial by presenting four examples, which show that the naïve use of linear algebra fails in certain situations.<sup>2</sup>

### 2.1.1 Commutators and traces

It is a simple exercise in linear algebra to show that there holds  $\text{Tr } AB = \text{Tr } BA$  for any two matrices  $A$  and  $B$  (prove it by coordinate representation), which may be rewritten as a trace of a commutator:  $\text{Tr } [A, B] = 0$ . In quantum mechanics, one requires, by analogy to the Poisson bracket of classical mechanics,  $[\hat{x}, \hat{p}] = i\hbar \mathbb{1}_{\mathcal{H}}$  for position and momentum operators  $\hat{x}$  and  $\hat{p}$ , respectively (known as *canonical quantisation*). Taking the trace on both sides of this equation yields  $0 = i\hbar \dim \mathcal{H}$ .<sup>3</sup> *So does quantum mechanics really exist?*

### 2.1.2 Distributions

Consider a particle in a box (square-well potential) with infinitely high walls:  $V(x) = 0$ , if  $|x| \leq a$  and  $V(x) = \infty$  otherwise. Let  $\Psi(x) = \frac{\sqrt{15}}{4a^{5/2}}(a^2 - x^2)$  be the wavefunction in the interior part which shall vanish outside of the box. If we want to calculate the variance of the energy  $\Delta \hat{H} = \langle \hat{H}^2 \rangle - \langle \hat{H} \rangle^2$ , we need the expectation value of  $\hat{H}^2$ . From  $\hat{H}^2 \Psi = \frac{\hbar^4}{4m^2} \frac{\partial^4 \Psi}{\partial x^4} = 0$ , we have  $\langle \hat{H}^2 \rangle = \int_{x=-a}^a \Psi^*(x) \hat{H}^2 \Psi(x) dx = 0$ . Since  $\langle \hat{H}^2 \rangle$  is strictly positive, the variance  $\Delta \hat{H}$  is negative, which obviously is impossible. *What is wrong here?*

### 2.1.3 Eigenvalues of hermitian operators I

Consider a radially-symmetric potential  $V(\vec{r})$  in three space-dimensions. Using the ansatz  $\Psi(\vec{r}) = R(r) Y_{lm}(\vartheta, \varphi)$  in spherical coordinates we get the spherical harmonics  $Y_{lm}$  and by substituting  $R(r) = \frac{u(r)}{r}$  we get a radial Schrödinger equation for  $u(r)$  by adding a centrifugal potential  $\frac{\hbar^2 l(l+1)}{2mr^2}$  to  $V(r)$ . We thus have a radial position operator  $\hat{r}$  and a corresponding momentum  $\hat{p}_r = \frac{\hbar}{i} \frac{\partial}{\partial r}$ , which fulfil the canonical commutator relation. Now,  $\hat{p}_r$  is hermitian, but  $\psi(r) = e^{\frac{i}{\hbar} \alpha r}$  is a normalisable eigen-

function on  $[0; \infty)$  with a possibly complex eigenvalue  $\alpha$ , provided that  $\text{Im } \alpha > 0$ . *How is this possible?*

### 2.1.4 Eigenvalues of hermitian operators II

In the previous example, one may argue that the radius cannot be negative and some part of the eigenfunction is cut off. Now we give a more dramatic example. First we consider an abstract hermitian operator  $A$  with eigenvector  $v$  and eigenvalue  $\lambda$ :  $Av = \lambda v$ . Hermiticity of  $A$  is defined by  $\langle Ax|y \rangle = \langle x|Ay \rangle$  for all  $x, y \in \mathcal{H}$ , and textbook quantum mechanics (cf. e. g. Schwabl [10]) tells us that eigenvalues of hermitian operators are real by the following simple argument: we calculate

$$\lambda^* \langle v|v \rangle = \langle \lambda v|v \rangle = \langle Av|v \rangle = \langle v|Av \rangle = \langle v|\lambda v \rangle = \lambda \langle v|v \rangle, \quad (1)$$

so  $\langle v|v \rangle \neq 0$  yields  $\lambda = \lambda^*$ , and this implies  $\lambda \in \mathbb{R}$ .

Now take a spinless pointlike one-dimensional particle in position space, which is described by a square-integrable function on  $\mathbb{R}$ . Consider the hermitian operator  $\hat{A} = \hat{x}^3 \hat{p} + \hat{p} \hat{x}^3$  with the usual  $\hat{x} = x$  and  $\hat{p} = \frac{\hbar}{i} \frac{\partial}{\partial x}$  and take the function (plotted in the diagram)

$$f(x) = \frac{1}{\sqrt{2}} |x|^{-3/2} e^{-\frac{1}{4x^2}}. \quad (2)$$

Setting  $f(0) := 0$ , the function  $f$  is everywhere defined on  $\mathbb{R}$  and “smooth”.<sup>4</sup> Further—unlike plane waves, delta functions and the like—it is square-integrable and normalised:  $\int_{x \in \mathbb{R}} |f(x)|^2 dx = \int_{x=0}^{+\infty} x^{-3} e^{-\frac{1}{2x^2}} dx = [e^{-\frac{1}{2x^2}}]_{x=0}^{+\infty} = 1$ . For the derivative of  $g(x) := e^{-\frac{1}{4x^2}}$  we find  $g'(x) = (-\frac{1}{4}x^{-2})' \cdot g(x) = \frac{1}{2}x^{-3}g(x)$ , thus for  $x > 0$ , there holds

$$\hat{x}^3 \hat{p} x^{-3/2} g(x) = \frac{\hbar}{i} x^3 \left( -\frac{3}{2} x^{-5/2} + x^{-3/2} \frac{x^{-3}}{2} \right) g(x) \quad \text{and} \quad (3)$$

$$\begin{aligned} \hat{p} \hat{x}^3 x^{-3/2} g(x) &= \frac{\hbar}{i} \frac{d}{dx} [x^{3/2} g(x)] \\ &= \frac{\hbar}{i} \left( \frac{3}{2} x^{1/2} + x^{3/2} \frac{x^{-3}}{2} \right) g(x). \end{aligned} \quad (4)$$

As  $f$  is even, and each power of  $\hat{x}$  and  $\hat{p}$  changes symmetries once, there holds  $\hat{A}f = \frac{\hbar}{i} f$  for the function as a whole. Altogether,  $\hat{A}$  is a hermitian operator with an eigenfunction  $f$ , but its eigenvalue is not a real number. *Where is the error?*

<sup>2</sup> For these and other related examples see the articles by Gieres [4] and Bonneau *et al.* [5], which are very much recommended to the reader.

<sup>3</sup> For a real commutator (without the imaginary prefactor  $i\hbar$ ), one can use the ladder operators in a harmonic oscillator with  $[\hat{a}, \hat{a}^\dagger] = 1$  in a similar fashion.

<sup>4</sup> It can be arbitrarily often differentiated. It should not bother us that as a complex function there is an essential singularity in the origin, since we only consider real functions here.

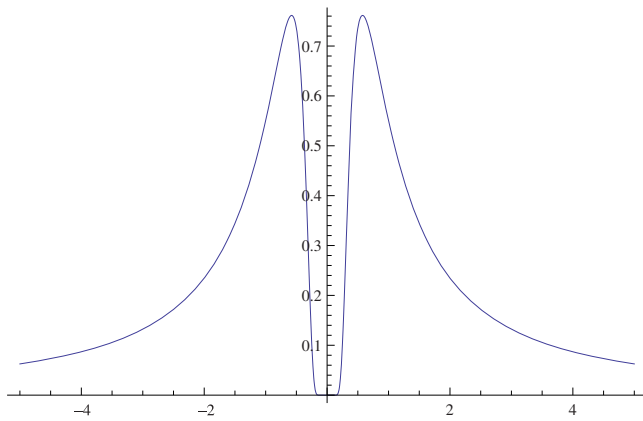


Figure 1 A wavefunction with imaginary eigenvalue.

## 2.2 Linear algebra and functional analysis

In the previous examples we have used concepts familiar from linear algebra, which work well in finite-dimensional systems. But very often quantum-mechanical systems need infinite-dimensional Hilbert spaces (e. g. for a particle in space). Let us summarise:

1. We often use terminology and methods of linear algebra, such as vector spaces, bases, scalar products, orthogonality, matrices and their diagonalisation, operators etc.
2. Most quantum systems—notable exceptions are spin systems—have infinite dimension: a free particle, a particle in a box, a particle in a harmonic oscillator, the Hydrogen atom etc.
3. A mathematically rigorous description thus must make use of functional analysis, in some sense “linear algebra in infinite-dimensional vector spaces”.
4. Many aspects of linear algebra extend to functional analysis, but not all, and even if they do, several theorems and statements are much more complicated than is expected at first sight.

At this point we cannot show how to resolve the problems posed in detail, for which we refer to the references already mentioned, but we can give some brief explanation: In the first example the trace is simply not defined on all operators on a Hilbert space, but only on a subset called the trace-class operators. In the second example we get by differentiating delta distributions at  $-a$  and  $a$ , which we cannot ignore in our calculation. The third example shows, that on infinite-dimensional spaces we have to consider domains of definition. In the fourth example, we presented the applications of  $\hat{x}^3 \hat{p}$  and  $\hat{p} \hat{x}^3$  on the function separately; these map  $f$  to  $\frac{e^{-1/4x^2}}{2\sqrt{2}} \frac{\hbar}{i} (x^{-3/2} \mp 3x^{1/2})$ , respectively, which are not square-integrable, but their

sum is, i. e. in the course of calculation we secretly leave the Hilbert space for some time, which is not allowed.

## 2.3 Hilbert spaces

We start with the definition of a Hilbert space. We will use Dirac's bra-ket notation when we find it appropriate.

**Definition 1** (Hilbert space). A pre-Hilbert space is a vector space over some field  $K$  equipped with a scalar product (or inner product), i. e. a function  $\langle \cdot | \cdot \rangle : V \times V \rightarrow K$  with the following three properties:

1. *sesquilinearity*:  $\langle x | \lambda v + \mu w \rangle = \lambda \langle x | v \rangle + \mu \langle x | w \rangle$  and  $\langle \lambda v + \mu w | x \rangle = \lambda^* \langle v | x \rangle + \mu^* \langle w | x \rangle$ ,
2. *anti-symmetry*:  $\langle w | v \rangle = \langle v | w \rangle^*$  and
3. *positive definiteness*:  $\langle v | v \rangle > 0$  for  $v \neq 0$ .

A Hilbert space is a pre-Hilbert space which is complete (in the sense of metric spaces, i. e. every Cauchy sequence converges with a limit in the space itself). A real or complex Hilbert space is a Hilbert space, where  $K = \mathbb{R}$  or  $K = \mathbb{C}$ , respectively.<sup>5</sup>

Though completeness is not explicitly used in physics, it is necessary for certain mathematical theorems. From now on, we shall only consider complex Hilbert spaces, and  $\mathcal{H}$  shall always denote such a complex Hilbert space. Hilbert spaces are essentially defined by their bases.

**Definition 2** (Bases of Hilbert spaces). A basis (more precisely, an orthonormal basis) of  $\mathcal{H}$  is a set  $B$  of vectors in  $\mathcal{H}$ , such that there holds (i)  $\langle v | w \rangle = \delta_{v,w}$  for all  $v, w \in B$  and (ii)  $x = \sum_{v \in B} x(v) |v\rangle$  for all  $x \in \mathcal{H}$ .

A Hilbert space which is very important for quantum mechanics is the space  $L^2(\mathbb{R})$  of square-integrable functions.<sup>6</sup> Note that the uncountable position and momentum bases are not true bases, since neither plane waves nor delta functions are square-integrable functions.

It is very easy to give a formal classification of Hilbert spaces in purely mathematical terms. However, it will be

<sup>5</sup> Note that the notation in mathematics and physics differs: the complex conjugate of  $z \in \mathbb{C}$  is denoted  $\bar{z}$  in mathematics and  $z^*$  in physics, the adjoint  $A^*$  and  $A^\dagger$ , respectively. The scalar product is usually linear in the left component in mathematics and in the right component in physics. In older mathematics literature and sometimes in physics Hilbert spaces are occasionally required to be infinite-dimensional and/or separable; cf. e. g. von Neumann [11] or Scheck [12].

<sup>6</sup> This is strictly speaking not true; a precise definition will follow, when we discuss measure and integration theory.

seen that it is not the different Hilbert spaces which distinguishes different physical systems, but the operators on these Hilbert spaces.

**Theorem 1** (Classification of Hilbert spaces). *For any possible cardinality (i. e. “number of elements”) of a set there exists up to an isomorphism precisely one Hilbert space; in particular, two Hilbert spaces are isomorphic, if and only if their dimensions coincide.*<sup>7</sup>

The isomorphism concept can be understood as “indistinguishable by inner structure”, i. e. by Hilbert space operations only; in the well-known setting of square-integrable functions we may always think of all functions, some of which lie inside Hilbert space and others outside (are “non-normalisable”), but in a pure Hilbert-space setting such considerations are not allowed. For a positive integer  $d$ , the Hilbert space  $\mathbb{C}^d$  of column vectors with  $d$  complex entries is essentially the only Hilbert space with  $\dim \mathcal{H} = d$ . A Hilbert space is called *separable*, if its dimension is countable, i. e. less or equal of the cardinality of the integers.<sup>8</sup> Thus, up to an isomorphism there exists precisely one infinite-dimensional separable Hilbert space.

Consider for example a single particle in one space dimension: the Hilbert space is the set of square-integrable functions  $L^2(\mathbb{R})$ . Now change to the Fock basis, the eigenstates of the harmonic oscillator  $\{|n\rangle \mid n \in \mathbb{N}_0\}$ , which is a countable and infinite basis. But obviously we have by this an isomorphism to the Hilbert space  $\ell^2(\mathbb{N}_0)$  of square-integrable sequences. Note that this is still a basis regardless of the potential (though not necessarily an eigenbasis), so that  $L^2(\mathbb{R})$  is separable. Further, finite or countable unions and cartesian products of countable sets are still countable; thus direct sums and tensor products of finitely and countably many separable Hilbert spaces are separable.

Although inseparable Hilbert spaces are not so common in physics, it is very easy to construct them. To construct a Hilbert space for an arbitrary given dimension, choose an arbitrary index set  $I$  of that cardinality and let

$$\ell^2(I) := \left\{ (x_i)_{i \in I} \in \mathbb{C}^I \mid \sum_{i \in I} |x_i|^2 < \infty \right\}. \quad (5)$$

The sum of possibly uncountably many terms is to be understood in the following sense of *summability*: it is defined only if at most countably many  $x_i$  are non-

zero (otherwise divergent) and moreover, any series constructed out of these non-zero elements converges absolutely. The dimension of  $\ell^2(I)$  then is  $|I|$ , the scalar product  $\langle x|y \rangle := \sum_{i \in I} x_i^* y_i$ . For example, the set of reals  $\mathbb{R}$  is uncountable, so that  $\ell^2(\mathbb{R})$  is inseparable.

To conclude this section: In quantum mechanics we essentially deal with a single Hilbert space, the separable, infinite-dimensional Hilbert space—though it is not the Hilbert space itself, but rather the structure of operators on this space which distinguishes different physical situations.

### 3 Spectral theory

Spectral theory most prominently deals with the generalisation of the spectral theorem from linear algebra—the diagonalisation of hermitian (or normal) matrices—to infinite-dimensional spaces. To understand these things, a basic knowledge of measure and integration theory is necessary, which will be introduced in the first two subsections. Then we can finally state the precise meaning of the Hilbert space  $L^2(\mathbb{R})$ .

#### 3.1 Measure theory

Measure theory deals with the problem of associating a measure to certain sets.<sup>9</sup> Measures can be thought of the length of a distance, the area of a surface or the volume of some space. Thus, given a set  $\Omega$  (think of  $\mathbb{R}^n$ ), one could envision a function  $\mu : \mathfrak{P}(\Omega) \rightarrow [0; \infty]$  (where  $\mathfrak{P}(\Omega)$  is the power set, the set of all subsets of  $\Omega$ ) which maps to each set its “volume”—which is non-negative. However, there are certain pitfalls, which are best illustrated by the following paradox (in mathematics, this is related to the axiom of choice).

**Theorem 2** (Banach-Tarski paradox). *Let  $A, B \in \mathbb{R}^n$  be arbitrary sets with non-empty interior. Then, there exist countably many sets  $A_1, A_2, \dots$  and rigid motions<sup>10</sup>  $m_1, m_2, \dots$ , such that  $A = \bigcup_{n \in \mathbb{N}} A_n$  and  $B = \bigcup_{n \in \mathbb{N}} m_n(A_n)$ , where both decompositions are disjoint, i. e.  $A_i \cap A_j = \emptyset$  and  $m_i(A_i) \cap m_j(A_j) = \emptyset$  for  $i \neq j$ . For  $A, B \in \mathbb{R}^n$  bounded and  $n \geq 3$ , this is even possible with finitely many sets and rigid motions.*

<sup>7</sup> The proof is rather simple, it essentially reads *Map one basis bijectively onto another.*

<sup>8</sup> This notion of separability is completely unrelated to the same term from entanglement theory.

<sup>9</sup> Measure theory is not to be confused with the theory of measurements in quantum physics.

<sup>10</sup> These are bijective mappings  $m_i$  on  $\mathbb{R}^n$  which preserve Euclidean distances, i. e.  $d(m_i(x), m_i(y)) = d(x, y)$ .



It is actually possible to cut a sphere in just five parts, which can be reassembled to two spheres both of the same size as the original sphere. To avoid such seemingly absurd effects, one restricts the measure function  $\mu$  to a class of *measurable* sets  $\Sigma$  and calls  $(\Omega, \Sigma)$  a *measurable space*, if  $\Sigma \subseteq \mathfrak{P}(\Omega)$  forms a  $\sigma$ -algebra. A  $\sigma$ -algebra must fulfil three conditions: (i) it contains the basic set:  $\Omega \in \Sigma$ , (ii) it contains every complement:  $A \in \Sigma \Rightarrow \Omega \setminus A \in \Sigma$  and (iii) it contains countable(!) unions:  $A_1, A_2, \dots \in \Sigma \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in \Sigma$ .

A *measure* now is a function  $\mu : \Sigma \rightarrow [0; \infty]$ , which maps to each measurable set a non-negative number or infinity and fulfils the following axioms: (i)  $\mu(\emptyset) = 0$  and (ii) for countably many disjoint sets  $(A_n)_{n \in \mathbb{N}}$  there holds  $\sigma$ -additivity:  $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ . The triple  $(\Omega, \Sigma, \mu)$  is called a *measure space*.

Examples of measures are:

- the *counting measure*  $|\cdot|$  on arbitrary sets, i. e. the cardinality (number of elements) of a set;
- the *Lebesgue measure*  $\lambda$  on  $\mathbb{R}^n$  is defined by  $\lambda([a_1; b_1] \times \dots \times [a_n; b_n]) = \prod_{i=1}^n (b_i - a_i)$  and extending it (as a so-called regular measure) to other Lebesgue-measurable sets by exhaustion.

The Lebesgue measure is very important, and the set of measurable sets is very large; it is rather difficult to find a set which is not measurable (indeed, without invoking the axiom of choice, it is not possible to construct a non-measurable set). Important are *null sets*, i. e. sets of measure zero and their subsets; a property is said to hold *almost everywhere*, if it holds everywhere with the possible exception of a nullset. Finally, note that every countable set is a Lebesgue nullset, but not vice versa (e. g. the Cantor set).

Measures for which there hold  $\mu(\Omega) = 1$  are called *probability measures* and are the foundation of Kolmogoroff's probability theory. There are some generalisations of the notion of a measure to be informally mentioned here: for a *signed measure*, the measure of a set may be negative (in  $\mathbb{R}$ ), and for a *complex measure*, the measure of a set may be complex (in  $\mathbb{C}$ ). More important in quantum physics, in particular in quantum information, are *operator-valued measures*, where the measure of a set is not a number, but an operator on a Hilbert space:

- a *positive operator-valued measure* (POVM) maps a set to a positive operator, such that  $\mu(\Omega) = \mathbb{1}$  (this may be thought of as a generalised probability measure);
- a *spectral measure* is a POVM with the additional property that  $E_1 \cap E_2 = \emptyset$  implies that  $\mu(E_1)$  and  $\mu(E_2)$

are orthogonal (these measures will be used in the spectral theorem below).

### 3.2 Integration theory

Usually while learning mathematics in school one starts with the Riemann integral of a function. The preimage  $[a, b]$  of a bounded function to be integrated, e. g.  $f : [a, b] \rightarrow \mathbb{R}$  is partitioned into smaller and smaller intervals, and the limit of upper and lower sums, if they both exist and coincide, is called the Riemann integral of  $f$ . We shall now introduce the Lebesgue integral.

The guiding principle of the Lebesgue integral is to partition the image instead of the preimage space. For a measurable set  $M \subseteq \Omega$  the *characteristic function* (or indicator function)  $\chi_M$  is defined by  $\chi_M(x) = 1$ , if  $x \in M$ , and zero otherwise. The integral for such a function is then defined as  $\int_{\Omega} \chi_M(x) d\mu(x) := \mu(M)$  and for a step function  $f(x) = \sum_{i=1}^n a_i \chi_{A_i}(x)$  with positive  $a_i$  accordingly by  $\int_{\Omega} f(x) d\mu(x) := \sum_{i=1}^n a_i \mu(A_i)$ . For a general non-negative function  $f$  the integral is defined by approximation from below as the supremum of all non-negative step functions majorised by  $f$ . General real-valued functions are split into positive and negative part by  $f = f_+ - f_-$  with  $f_{\pm}(x) := \max\{\pm f(x), 0\}$ , and the integral is given by  $\int f := \int f_+ - \int f_-$ , if both parts exist separately; complex-valued functions are further split into real and imaginary part.

**Theorem 3** (Riemann-integrable functions). *A bounded function  $f : [a; b] \rightarrow \mathbb{R}$  is integrable in the sense of Riemann, if and only if the set of discontinuities has vanishing Lebesgue measure.*

An example of a function, which is integrable according to Lebesgue, but not in the sense of Riemann, is the Dirichlet function  $\chi_{\mathbb{Q}}$  on  $\mathbb{R}$ . However, there may exist improper integrals in the Riemann sense for functions which do not possess a Lebesgue integral, e. g.  $\text{sinc}(x) = \frac{\sin x}{x}$  on  $\mathbb{R}$ ; here, positive and negative part diverge, but the limit  $x \rightarrow \infty$  is possible (compare the conditionally convergent series  $a_n := \frac{(-1)^n}{n}$ ).

**Definition 3** ( $L^p$  spaces). *For  $p \in [1; \infty)$  and a measurable functions  $f$ , define  $\|f\|_p := [\int |f(x)|^p d\mu]^{1/p}$ . The set  $\mathcal{L}^p(\Omega, \Sigma, \mu)$  is the set of functions  $f : \Omega \rightarrow \mathbb{C}$ , for which  $\|f\|_p$  is finite. If  $\mathcal{N}(\Omega, \Sigma, \mu)$  denotes the set of all functions which are zero almost everywhere (in the sense mentioned above), we can identify two functions  $f$  and  $g$ , if  $f - g \in \mathcal{N}(\Omega, \Sigma, \mu)$ . The set of equivalence classes of functions which are identical almost everywhere is the set  $L^p(\Omega, \Sigma, \mu) := \mathcal{L}^p(\Omega, \Sigma, \mu) / \mathcal{N}(\Omega, \Sigma, \mu)$ .*

Instead of always writing  $L^p(\Omega, \Sigma, \mu)$ , one uses shorthand notations like  $L^p(\Omega)$ . Note that the Hilbert space  $L^2(\Omega)$  is different from the set of square-integrable functions  $\mathcal{L}^2(\Omega)$ . In particular, since a set with one element  $\{x\}$  has Lebesgue measure zero, for a function  $f \in L^p(\Omega)$ ,  $f(x)$  is strictly speaking not defined. Considering that in physics we only deal with integrals of  $L^2$  functions, it is also reasonable not to distinguish between functions which integrated over every subset give the same value.

### 3.3 Operators and their domains

After introducing the Lebesgue integral, we can come back to spectral theory. In the examples we presented a hermitian operator with an imaginary eigenvalue. On the other hand, one requires physical quantities like energy etc. to be real. The crucial point is that the terms “hermitian” and “self-adjoint” are different for unbounded operators, while they coincide for matrices and, more general, bounded operators. Only for the self-adjoint operators the spectral theorem holds.

A linear mapping  $A: \mathcal{H} \rightarrow \mathcal{H}$  on a Hilbert space can either be bounded or unbounded. Such mapping is bounded, if the operator norm

$$\|A\| := \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} \quad (6)$$

exists, i. e. it is finite. (For example, the operator norm of a hermitian matrix is the maximum of the absolute values of the eigenvalues.).

It can be shown that linear operators are bounded, if and only if they are continuous. In particular, all linear operators on finite-dimensional Hilbert spaces—essentially matrices—are bounded and thus continuous. Examples of unbounded operators from physics are position and momentum operator and many Hamiltonians (free particle, harmonic oscillator etc.).

Several theorems of linear algebra (in particular, the spectral theorem) can be extended to the set of bounded operators (sometimes to a subset known as compact operators) and in the mathematical theory one only later discusses unbounded operators. We shall not present proofs here and thus deal only with the general case.

We consider an unbounded linear operator  $A$ . In general, such an operator is not defined on the Hilbert space, but only on a subset, its *domain*  $D(A)$ . For example, the position operator in positions representation  $(\hat{x}\psi)(x) = x\psi(x)$  may (by multiplication) throw some  $L^2$  function with asymptotics  $\psi(x) \sim \frac{1}{|x|}$  for  $|x| \gg 1$  out

of Hilbert space, i. e.  $\hat{x}\psi \notin \mathcal{H}$ . At least such cases must be excluded by restricting to an appropriate domain of definition  $D(A)$ .

An operator  $A: D(A) \rightarrow \mathcal{H}$  with domain of definition  $D(A) \subseteq \mathcal{H}$ —which always shall be a dense subset of  $\mathcal{H}$ —is called *hermitian* (in mathematics the term *symmetric* is more common), if

$$\langle w|Av \rangle = \langle Aw|v \rangle \quad \text{for all } v, w \in D(A). \quad (7)$$

Due to the Hellinger-Toeplitz theorem,  $D(A) = \mathcal{H}$  is in this case only possible, if  $A$  is bounded. The operators from quantum mechanics ( $\hat{x}$ ,  $\hat{p}$ ,  $\hat{H}$  etc.) are mostly unbounded, and their domains are dense subsets of  $\mathcal{H}$ . The *adjoint operator* of  $\hat{A}$  has (precisely) the domain of definition

$$D(A^\dagger) := \{w \in \mathcal{H} \mid v \mapsto \langle w|Av \rangle \text{ is continuous on } D(A)\}, \quad (8)$$

and is on this domain defined as follows: For  $w \in D(A^\dagger)$ , the map  $v \mapsto \langle w|Av \rangle$  can uniquely be extended to  $\mathcal{H}$  and by the Riesz-Fréchet theorem (see below), there exists a  $z \in \mathcal{H}$ , such that  $\langle z|v \rangle = \langle w|Av \rangle$ ; we then define  $A^\dagger w := z$ . An operator is *self-adjoint* (rarely also *hypermaximal hermitian*), if  $A = A^\dagger$ , where in particular the domains of definition coincide, i. e.  $D(A) = D(A^\dagger)$ ; only for these operators the spectral theorem holds.

The (algebraic) dual of a vector space  $E$  over some field  $K$  is defined as the space of linear functionals  $E \rightarrow K$ ; for a normed space the (topological) dual  $E'$  of  $E$  consists of the functionals which are additionally required to be continuous.

**Theorem 4** (Riesz-Fréchet theorem). *Let  $\mathcal{H}$  be a Hilbert space and consider some  $\varphi \in \mathcal{H}$ . The mapping  $\psi \mapsto \langle \varphi|\psi \rangle$  is in  $\mathcal{H}'$ , and, on the other hand, every element of  $\mathcal{H}'$  can be written in this form by some  $\varphi \in \mathcal{H}$ .*

In physics, one denotes elements of  $\mathcal{H}$  as “ket” vectors  $|\psi\rangle$  and elements of the dual  $\mathcal{H}'$  as “bra” vectors  $\langle\varphi|$ ; the stated theorem is the mathematical reason for Dirac’s formalism.

### 3.4 Multiplication operators

Given an operator  $A$ , it is not always easy to determine  $D(A^\dagger)$ . An example, where the domain of definition of the adjoint is straightforward to specify, is the case of multiplication operators. Let  $(\Omega, \Sigma, \mu)$  be a measure space, e. g.  $\Omega = \mathbb{R}$  with the usual Lebesgue measure, and let

$f : \Omega \rightarrow \mathbb{R}$  be a measurable function. The Hilbert space again is  $L^2(\Omega) := L^2(\Omega, \Sigma, \mu)$ .

We can define a multiplication operator  $A_f : D(A_f) \rightarrow \mathcal{H}$  by  $(A_f g)(x) := f(x) \cdot g(x)$  on the domain  $D(A_f) := \{g \in L^2(\Omega) \mid f \cdot g \in L^2(\Omega)\}$ . Obviously, this is the largest possible domain of definition for  $A_f$ , since we require the image of  $A_f$  to lie in  $L^2(\Omega)$ . It can be shown that this operator is selfadjoint with  $D(A_f^\dagger) = D(A_f)$ , and the spectral theorem will tell us, that every selfadjoint operator can be written in this fashion. One example for such an operator is the position operator  $\hat{x}$ .

### 3.5 Extensions of operators

Sometimes, for non-selfadjoint operators, there exist selfadjoint extensions. An operator  $B$  is called an extension of another operator  $A$ , if  $D(A) \subseteq D(B)$  and  $A = B|_{D(A)}$  ( $B$  restricted to the domain of  $A$ ); this is usually denoted as  $A \subseteq B$ . The question of whether selfadjoint extensions exist, is answered by the theory of defect indices. Let  $z_\pm$  be arbitrary complex numbers with positive and negative imaginary part, respectively (one usually takes  $z_\pm = \pm i$ ). The defect indices are—independent of the choice of  $z_\pm$ —given by  $n_\pm := \dim \{\psi \in D(A^\dagger) \mid A^\dagger \psi = z_\pm \psi\} \in \mathbb{N}_0$ . An operator is selfadjoint, if and only if  $n_+ = n_- = 0$ ; there exist selfadjoint extensions, if and only if  $n_+ = n_- > 0$ , but not otherwise.

A semibounded operator is an operator which fulfils  $\langle v | A v \rangle \geq C \|v\|^2$  (and therefore is hermitian) for all  $v \in D(A)$  (or, similarly “ $\leq$ ”). These operators possess selfadjoint extensions with the same constant  $C$  (Friedrichs extension). Note that Hamiltonian operators are usually semibounded because there should exist a ground-state energy.

### 3.6 Spectral theorem

For a matrix  $A$ , the spectrum is the set of eigenvalues, and  $\lambda \in \mathbb{C}$  is an eigenvalue of  $A$ , if and only if  $\det(\lambda \mathbb{1} - A) = 0$ , i. e., if  $\lambda \mathbb{1} - A$  is not invertible. For operators on infinite-dimensional Hilbert spaces, there are different classes of “non-invertibility” defining different parts of the spectrum. We shall start with the complement of the spectrum, the resolvent set. For an operator  $A : D(A) \rightarrow \mathcal{H}$  the resolvent set is defined by

$$\rho(A) := \{\lambda \in \mathbb{C} \mid \lambda \mathbb{1} - A : D(A) \rightarrow \mathcal{H} \text{ is bijective and } (\lambda \mathbb{1} - A)^{-1} \text{ is bounded}\}. \quad (9)$$

On this set, the resolvent (mapping) is  $R_A(\lambda) := (\lambda \mathbb{1} - A)^{-1}$ . The spectrum  $\sigma(A) := \mathbb{C} \setminus \rho(A)$  is the complement of the resolvent set; it is closed and can be separated into three disjoint parts:

1. the point spectrum:  $\lambda \mathbb{1} - A$  is non-injective (i. e. there are eigenvectors);
2. the continuous spectrum:  $\lambda \mathbb{1} - A$  is injective, but not surjective, and possesses a dense image;
3. the residual spectrum:  $\lambda \mathbb{1} - A$  is injective, but not surjective and its image is not dense.

The residual spectrum of a selfadjoint operator is empty, but e. g. zero is in the residual spectrum of the raising operator  $\hat{a}^+$ . The spectrum of the lowering operator  $\hat{a}$  is  $\mathbb{C}$  and a pure point spectrum (cf. coherent states).

The principal property of selfadjoint operators is that for these the spectral theorem holds, the “diagonalisation” of operators. There are several forms of the spectral theorem, but we shall discuss only two of them. Let  $A$  be a selfadjoint operator on an appropriate domain  $D(A)$ :<sup>11</sup>

1. There exists a measure space  $(\Omega, \Sigma, \mu)$ , a measurable function  $f : \Omega \rightarrow \mathbb{R}$  and a unitary operator  $U : \mathcal{H} \rightarrow L^2(\mu)$ , such that (i)  $x \in D(A) \Leftrightarrow f \cdot Ux \in L^2(\mu)$  and (ii)  $UAU^\dagger \varphi = f \cdot \varphi$  for  $\varphi \in L^2(\Omega)$  with some  $f \cdot \varphi \in L^2(\Omega)$ .
2. There exists a unique spectral measure  $E : \Sigma \rightarrow \mathcal{B}(\mathcal{H})$  (i.e. the bounded operators on  $\mathcal{H}$ ) such that  $\langle y | Ax \rangle = \int_{\lambda \in \mathbb{R}} \lambda \, d\langle y | E_\lambda x \rangle$  for  $x \in D(A)$  and  $y \in \mathcal{H}$ ; here, a spectral measure is a mapping from subsets of  $\mathbb{R}$  (the so-called Borel- $\sigma$ -algebra), where  $E_A$  are orthogonal projections,  $E_\emptyset = 0$ ,  $E_{\mathbb{R}} = \mathbb{1}$  and for pairwise disjoint  $A_1, A_2, \dots \in \Sigma$  there holds  $\sum_{i=1}^\infty E_{A_i} = E_{\bigcup_{i=1}^\infty A_i}$  pointwise.

The first version informally says that every selfadjoint operator can be understood as multiplication operator on some Hilbert space. The second version allows us to apply a measurable function  $f : \mathbb{R} \rightarrow \mathbb{R}$  to  $A$ : in the spectral integral replace  $\lambda$  by  $f(\lambda)$ , i. e.  $\int_{\lambda \in \mathbb{R}} f(\lambda) \, d\langle y | E_\lambda x \rangle$ .

## 4 Distributions

In classical physics concepts like point charges  $\rho(\vec{r}) = q \cdot \delta(\vec{r})$  in electrodynamics, line and area charges, currents etc. are commonly used. In quantum mechanics the eigenstates of the position operator  $\hat{x}$  (in position

<sup>11</sup> This and related statements usually also hold for normal operators, i. e. operators  $N$ , for which  $NN^\dagger = N^\dagger N$ , but where the spectrum does not need to be real.

representation) are said to be delta functions, because  $\hat{x}\delta(x) = x\delta(x)$ , those of the momentum operator  $\hat{p} = \frac{\hbar}{i} \frac{\partial}{\partial x}$  plane waves  $e^{ikx}$ . In both cases, these eigenfunctions are non-normalisable or not even true functions and thus lie outside of the Hilbert space  $\mathcal{H} = L^2(\mathbb{R})$ . Nevertheless, one regularly writes for position and momentum the equalities

$$\int_{x \in \mathbb{R}} |x\rangle\langle x| dx = \int_{p \in \mathbb{R}} |p\rangle\langle p| dp = \mathbb{1}_{\mathcal{H}} \quad (10)$$

and

$$\int_{x \in \mathbb{R}} x |x\rangle\langle x| dx = \hat{x}, \quad \int_{p \in \mathbb{R}} p |p\rangle\langle p| dp = \hat{p}, \quad (11)$$

which do work well in calculations. Similarly, for a potential well  $V(\vec{r})$ , one usually distinguishes between bound discrete and free continuous eigenstates and often writes the identities

$$\int |E\rangle\langle E| dE = \mathbb{1}_{\mathcal{H}} \quad \text{and} \quad \int E |E\rangle\langle E| dE = \hat{H}. \quad (12)$$

But how to mathematically understand these things? In the 1930s Dirac postulated a “delta function”  $\delta$  with the property that  $\delta(x) = 0$  for  $x \neq 0$ , but  $\int \delta(x) dx = 1$ ; for some other function  $f$ , this implies  $\int f(x)\delta(x) dx = f(0)$ . It is obvious that strictly speaking no such function  $\delta$  can exist, but the concept is nevertheless very useful in physics. The mathematical theory of such objects as the “delta function” is the theory of *distributions* (or *generalised functions*), which was developed already in the 1940s by S. L. Sobolev and Laurent Schwartz.

#### 4.1 Introduction

The basic idea of distributions in mathematics is, not to consider the “delta function” and similar objects as functions but as functionals—objects mapping vectors to scalars, as e. g. the delta function mapping the function  $f$  to a complex number  $f(0)$ ; in physics one would say that a delta function is defined only inside an integral. Mathematically, one can give a one-sentence definition of distributions: they are the *continuous linear functionals on the space of test functions*. Before explaining these terms in detail, we shall rephrase the Riesz-Fréchet theorem in a slightly different fashion.

**Theorem 5** (Riesz-Fréchet theorem). *Let  $\mathcal{H}$  be a Hilbert space and let  $\mathcal{H}'$  be its dual, i. e. the set of continuous linear functionals  $\mathcal{H} \rightarrow \mathbb{C}$ . For every  $\psi \in \mathcal{H}$ , the functional  $T_\psi : \mathcal{H} \rightarrow \mathbb{C}$  with  $T_\psi(\varphi) = \langle \psi | \varphi \rangle$  is an element of  $\mathcal{H}'$ , and,*

*vice versa, for every  $T \in \mathcal{H}'$  there exists some  $\psi$ , such that  $T = T_\psi$ .*

To each element of a Hilbert space there is associated a unique element of its dual and vice versa, i. e. there is some type of “symmetry” between  $\mathcal{H}$  and  $\mathcal{H}'$ ; they are isomorphic,  $\mathcal{H} \cong \mathcal{H}'$ , and “of the same size”.

In case of the Hilbert space  $L^2(\mathbb{R})$ , the explicit form of the functional  $T_f$  is  $\varphi \mapsto \int f(x)^* \varphi(x) dx$ . Now, if we choose a subset  $\Phi \subseteq \mathcal{H}$ —limiting the elements  $\varphi$  allowed in the domain of definition—, the number of admissible functionals  $\Phi'$  increases. The space  $\Phi$  is in general not a Hilbert space any more, and not all functionals can be written in the form of an integral. By an appropriately limited choice of  $\Phi$  (the *test functions*), we get a very large set  $\Phi'$  (the distributions, which “test” the test functions). A second aspect is that the functionals in  $\Phi'$  need to be continuous; (sequential) continuity of  $f$  is the property that  $x \rightarrow x_0$  in  $\Phi$  implies  $f(x) \rightarrow f(x_0)$  in  $\mathbb{C}$ . If one would like to have many continuous functionals, there must be not too many convergent series in  $\Phi$ . This will be achieved by an appropriate notion of continuity.

Before going into details, we shall give an overview on various mathematical structures in functional analysis in the following in Table 1; usually a set with some function has a specific name (such as Hilbert space).

For the details we must refer to the usual textbooks, but we give some notes here:

1. A structure in an upper row induces a structure in a lower below; there sometimes is the possibility to go from a lower row to a higher one (parallelogram equality, metrisability).
2. Scalar products and norms need a vector space in order to be defined, metrics and topologies can be defined on sets in general.
3. Weakening the axioms of a scalar product yields notions of bilinear and sesquilinear forms; without definiteness norms and metrics become semi-norms and semi-metrics. (Note that the term metric is used in a different fashion in differential geometry.)
4. Completeness is a notion in metric spaces (because of Cauchy sequences). In topological spaces other notions of convergence (net or filter convergence, which are equivalent) can be used.
5. Characteristic notions for Hilbert spaces are orthogonality, for normed spaces duality (algebraic, topological), for metric spaces sequences, Cauchy sequences etc.



Table 1 Mathematical structures.

	Function		Name of the space	Name, if complete
vector space $V$	scalar product	$\langle \cdot   \cdot \rangle : V \times V \rightarrow \mathbb{K}$	pre-Hilbert space	Hilbert space
	norm	$\  \cdot \  : V \rightarrow \mathbb{R}_0^+$	normed space	Banach space
general set $X$	metric	$d : X \times X \rightarrow \mathbb{R}_0^+$	metric space	complete metric space
	topology	system $\tau$ of “open sets”	topological space	–

## 4.2 Notions of convergence

In a very general sense, continuity in mathematics is a property of topological spaces (“Pre-images of open sets are open”), but in metric spaces the sequential continuity is more intuitive. In normed spaces, continuity is induced by the norm: a sequence  $(x_n)_{n \in \mathbb{N}}$  converges to  $x$ , if  $\|x_n - x\| \xrightarrow{n \rightarrow \infty} 0$ . In the theory of distributions, we consider so-called *topological vector spaces*, i. e. vector spaces equipped with a topology, such that the vector operations are continuous, or in particular *locally convex spaces* which are characterised by a family of seminorms  $(p_i)_{i \in I}$ : a sequence  $(x_n)_{n \in \mathbb{N}}$  converges to  $x$ , if for all  $i \in I$  there holds  $p_i(x_n - x) \xrightarrow{n \rightarrow \infty} 0$ .

As an example consider the vector space  $C[a; b]$  of continuous functions on an interval  $[a; b]$ . There are two well-known notions of convergence here: *pointwise* and *uniform* convergence. The latter is defined by the norm

$$\|f\|_\infty := \max \{ |f(x)| \mid x \in [a; b] \}; \quad (13)$$

a sequence of functions  $(f_n)_{n \in \mathbb{N}}$  converges to  $f$  uniformly, if and only if  $\|f_n - f\|_\infty$  converges to zero in  $\mathbb{R}$ . Pointwise convergence cannot be described by a single norm; therefore, consider for every  $x \in [a; b]$  the seminorm  $p_x(f) := |f(x)|$ : the sequence converges to  $f$  pointwise, if every (of the uncountably many) sequences  $p_x(f_n)$  converges to  $p_x(f)$  in  $\mathbb{R}$ .

## 4.3 Test functions and distributions

We shall start by defining a certain class of functions, the so-called *test functions*. To this aim, consider an open set  $\Omega \subseteq \mathbb{R}^n$ . The *support* of a function  $f$  is the closure of the set, where  $f$  does not vanish:  $\text{supp } f := \{x \in \Omega \mid f(x) \neq 0\}$ . Further, define by  $C^k(\Omega)$  the set of functions which are  $k$  times differentiable with a continuous derivative (where  $k$  may be infinite). The test functions now are the

set of  $C^\infty$ -functions with compact support in  $\Omega$ :

$$\begin{aligned} D(\Omega) &:= C_c^\infty(\Omega) \\ &= \{f : \Omega \rightarrow \mathbb{C} \mid f \in C^\infty(\Omega), \text{supp } f \text{ compact}\}. \end{aligned} \quad (14)$$

Compact support here is the property that the function vanishes outside of a certain bounded set.<sup>12</sup> An example of such a functions, which may be adapted in various ways, is  $\varphi(x) := \exp(\frac{1}{x^2-1})$ , if  $|x| < 1$ , and  $\varphi(x) = 0$  otherwise.

The notion of continuity on  $D(\Omega)$  is derived from the *topology of the strict inductive limit*, which is rather complicated. However, it is possible to describe sequences converging to zero in this topology: a sequence  $(\varphi_n)_{n \in \mathbb{N}}$  of test functions does so, if

1. there is a single compact set  $K \subseteq \Omega$ , such that  $\text{supp } \varphi_n \subseteq K$  holds for all  $n \in \mathbb{N}$ , and
2. for all  $m \in \mathbb{N}$  and all  $\alpha \in \mathbb{N}_0^m$  the sequence  $D^\alpha \varphi_n$  converges uniformly to zero.<sup>13</sup>

The *distributions*  $D'(\Omega)$  now are the continuous linear functionals on  $D(\Omega)$ . Examples are the delta distribution:  $\delta(\varphi) := \varphi(0)$ . Many functions can be interpreted as distributions. Let  $L_{loc}^1(\Omega)$  be the set of *locally integrable* functions  $f : \Omega \rightarrow \mathbb{C}$ , i. e. for every point in  $\Omega$  there exists a neighbourhood on which  $f$  is integrable (i. e. essentially excluding non-integrable poles). For each  $f \in L_{loc}^1(\Omega)$  the associated *regular distribution* is defined by  $T_f(\varphi) := \int_{x \in \Omega} f(x) \varphi(x) dx$ .

On  $D'(\Omega)$  one uses the *weak\*-convergence*: a sequence  $(T_n)_{n \in \mathbb{N}}$  of distributions converges to zero, if  $T_n(\varphi)$  does

<sup>12</sup> I. e., the functions must be exactly zero, not just approaching zero. In complex analysis, this is not possible for holomorphic functions  $\mathbb{C} \rightarrow \mathbb{C}$  because of the identity theorem, except for  $\varphi \equiv 0$ .

<sup>13</sup>  $D^\alpha$  is a shorthand notation for derivatives: for a function  $f : \mathbb{R}^n \rightarrow \mathbb{C}$  higher-order and higher-dimensional derivatives are written in by a *multi-index*  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ : there holds  $D^\alpha f := \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \dots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f$ , and  $|\alpha| := \alpha_1 + \dots + \alpha_n$  is called the order of  $\alpha$ .

so for all  $\varphi \in D(\Omega)$ ; the set  $D'(\Omega)$  is then complete (in the sense of sequences).

#### 4.4 Structure of distributions

The set of distributions  $D'(\Omega)$  obviously is a vector space. Moreover, every distribution can be multiplied with a  $C^\infty$ -function  $f$ :  $(fT)(\varphi) := T(f\varphi)$ ,  $\varphi \in D(\Omega)$ , but two distributions cannot be multiplied.<sup>14</sup> It is also possible to define the convolution of two distributions, and distributions can be approximated by  $C^\infty$ -functions (e. g. the delta distribution by Gaussians).

Distributions can be differentiated arbitrarily: in analogy to the product rule for a  $C^1(\mathbb{R})$  functions,  $\int f'\varphi = -\int f\varphi'$  (the boundary terms vanish), one defines the derivative of a distribution  $T$  by  $T'(\varphi) := -T(\varphi')$ . Higher-order and higher-dimensional derivatives are accordingly defined by  $(D^\alpha T)(\varphi) := (-1)^{|\alpha|} T(D^\alpha \varphi)$ .

#### 4.5 Sobolev spaces

Given a differentiable function  $f$  and any  $\varphi \in D(\Omega)$ , there holds that  $\int f\varphi' = -\int f'\varphi$  (the boundary terms are zero). Even if  $f$  is non-differentiable, there may exist some function  $g$ , such that  $\int f\varphi' = -\int g\varphi$  for all  $\varphi \in D(\Omega)$ . In contrast to the common (strong) derivative,  $g$  is called the *weak derivative* of  $f$ . The derivative in the sense of distributions is still weaker. Altogether there are three notions of derivatives: the common (strong), the weak and the distributional derivative.

**Example 1** (Derivatives). *The absolute-value function  $|\cdot| : \mathbb{R} \rightarrow \mathbb{R}$  is continuous for  $x = 0$ , but not differentiable. Nevertheless, there is a weak derivative, the sign function  $\text{sgn}$ , which is  $+1$  for  $x > 0$ ,  $0$  for  $x = 0$  and  $-1$  for  $x < 0$ . This function does not possess a weak derivative, but  $2\delta$  is its derivative as a distribution.*

The space of  $m$ -times weakly differentiable  $L^p$ -functions (i. e. all  $D^\alpha f$  exists for all  $\alpha$  with  $|\alpha| \leq m$ ), where all derivatives again lie in  $L^p$ , is called the *Sobolev space*  $W^{m,p}(\Omega)$  (or  $H^{m,p}(\Omega)$ ), and  $p = 2$  is often omitted. An interesting connection of weak and strong derivative is given by Sobolev's lemma (or embedding theorem): Let  $\Omega \subseteq \mathbb{R}^n$  be an open set,  $k \in \mathbb{N}_0$  and  $m > k + \frac{n}{p}$ . Then, there holds  $W^{m,p}(\Omega) \subseteq C^k(\Omega)$ .

#### 4.6 Tempered distributions and the Fourier transform

For test functions (or more generally for  $L^1$  and  $L^2$  functions) on  $\mathbb{R}^n$  one can define the Fourier transform. To define the Fourier transform for distributions, note that for  $L^2(\mathbb{R}^n)$  functions, there holds

$$\int (\mathcal{F}f)(x)g(x) dx = \int f(x)(\mathcal{F}g)(x) dx, \quad (15)$$

and one can, in the same spirit as in the case of derivatives, try to define the Fourier transform by acting on the test function. However,  $\mathcal{F}\varphi$  is never a test function, except for  $\varphi \equiv 0$ . Therefore, it is necessary to enlarge the space of test functions, which results in a smaller dual. The space which one considers is known as the space of *rapidly decreasing functions* (or SCHWARTZ functions):

$$S(\mathbb{R}^n) := \left\{ f : \mathbb{R}^n \rightarrow \mathbb{C} \mid f \in C^\infty(\mathbb{R}^n), \right. \\ \left. (\forall \alpha, \beta \in \mathbb{N}_0) \left( \lim_{n \rightarrow \infty} x^\alpha D^\beta f(x) = 0 \right) \right\}. \quad (16)$$

These are the functions which are arbitrarily often differentiable and go to zero faster than every polynomial, and their derivatives should do the same. Here the standard example is a Gaussian  $\varphi(x) = \exp(-\gamma x)$ ,  $\gamma > 0$ . The important property with respect to Fourier transforms is that  $\mathcal{F}(S(\mathbb{R}^n)) \subseteq S(\mathbb{R}^n)$ . There holds  $D(\mathbb{R}^n) \subseteq S(\mathbb{R}^n)$  and thus  $S'(\mathbb{R}^n) \subseteq D'(\mathbb{R}^n)$  – we shall not mention the topology here. The space  $S'(\mathbb{R}^n)$  is called the space of *tempered distributions*.

The Fourier transform now is defined by  $(\mathcal{F}T)(\varphi) := T(\mathcal{F}\varphi)$ . There holds e. g.  $(\mathcal{F}D^{(\alpha)}T)(\varphi) = (i/\hbar)^{|\alpha|} x^\alpha \mathcal{F}T$ ,  $\mathcal{F}\delta_a = \exp[-\frac{i}{\hbar}ax]/(2\pi\hbar)^{n/2}$  and  $\mathcal{F}1 = (2\pi\hbar)^{n/2}\delta_0$  (in  $\mathbb{R}^n$ ), where  $\delta_a$  is the delta function at some point  $a \in \mathbb{R}^n$ , i. e.  $\delta_a(x) = \delta(x - a)$ . For a polynomial  $f(x) = \sum_{k=1}^n a_k x^k$  there follows  $(\mathcal{F}f)(p) = \sum_{k=1}^n a_k (\frac{\hbar}{i} \frac{\partial}{\partial p})^k \delta$ . There is a regularisation of distributions: for  $f \in S'(\mathbb{R}^n)$  there exists a function  $g$  and some  $\alpha \in \mathbb{N}_0^n$ , such that  $f = D^\alpha g$ .

#### 4.7 Applications

Having discussed the guiding ideas in the definition of distributions, we shall present some applications of the theory of distributions, which are relevant to physics.

##### 4.7.1 Differential operators

Consider a linear partial differential operator  $L : S(\mathbb{R}^n) \rightarrow S(\mathbb{R}^n)$  of order  $m \in \mathbb{N}$  given by  $T \mapsto \sum_{|\alpha| \leq m} c_\alpha \frac{\partial}{\partial \alpha} T$  with

<sup>14</sup> In physics, however, this is done in some instances.

$c_\alpha \in C^\infty(\mathbb{R}^n)$  being smooth functions and the linear partial differential equation  $Lu = f$  with  $f \in D'(\mathbb{R}^n)$ . We look for a *fundamental solution* for  $L$ , i. e. a distribution  $E$ , for which there holds  $LE = \delta$ . If this  $E$  is known,  $u := f * E$  solves the original equation, because of  $Lu = L(f * E) = f * LE = f * \delta = f$ ; note that this is essentially the method of Green's functions. In particular, the delta function is the identity with respect to the convolution.

The standard example is the Poisson equation from electrostatics:  $\Delta \Phi(\vec{r}) = -4\pi\rho(\vec{r})$ . A fundamental solution of the Laplacian  $\Delta$  is  $E = -\frac{1}{4\pi|\vec{r}|}$  and  $\Phi(\vec{r}) = \int \frac{\rho(\vec{r}')}{|\vec{r}-\vec{r}'|} d\vec{r}'$ . (Note that for convolutions there holds  $T_{f_1} * T_{f_2} = T_{f_1 * f_2}$ ) The basic idea is to use the Fourier transform:  $\Delta E = \delta$  becomes  $(-x^2 - y^2 - z^2) \cdot \mathfrak{F}E = -\vec{r}^2 \cdot \mathfrak{F}E = 1$  with  $r = \sqrt{x^2 + y^2 + z^2}$ , so  $\mathfrak{F}E = -1/\vec{r}^2$ ; the inverse transform of  $-1/\vec{r}^2$  is  $-1/(4\pi r)$  which gives the desired result.

#### 4.7.2 Rigged Hilbert spaces

At the end of this chapter, we would like to mention that there indeed exists a possibility to mathematically explain generalised eigenfunctions of position and momentum and of other operators, although this is rather complicated. Note first, that in the continuous spectrum of a selfadjoint (or more generally normal) operator there are no true eigenfunctions, but *approximate eigenfunctions*.

**Theorem 6** (Approximate eigenvectors). *Let  $T$  be a bounded selfadjoint operator on  $\mathcal{H}$  and  $\lambda \in \sigma(T)$ . Then, there exists a sequence  $(x_n)_{n \in \mathbb{N}}$  in  $\mathcal{H}$  with  $\|x_n\| = 1$  for all  $n \in \mathbb{N}$ , such that  $\lim_{n \rightarrow \infty} (\lambda \mathbb{1} - T)x_n = 0$ .*

For example, in case of the position operator consider Gaussians of decreasing width approach a delta function (up to normalisation).

For a generalised version of the spectral theorem one must choose a space  $\Phi$  on which there is defined a scalar product and which additionally is nuclear (a property which cannot be explained here); possible choices for  $\Phi$  are the test functions or the Schwartz functions. The completion of  $\Phi$  is a Hilbert space  $\mathcal{H}$  and dual of  $\Phi$  is  $\Phi'$ . With the appropriate embeddings, the triplet  $\Phi \subseteq \mathcal{H} \subseteq \Phi'$  is called *rigged Hilbert space* or *GELFAND triplet*. There holds the Gelfand-Kostyuchenko theorem [1, 8, 9], which we cannot give in rigorous form here. But informally spoken, it states that for an operator  $T$  under certain assumptions, there exist enough generalised eigenvectors  $|x_\lambda\rangle$ , such that we may write  $\mathbb{1} = \int_\lambda |x_\lambda\rangle\langle x_\lambda| d\mu_\lambda$  and  $T = \int_\lambda \lambda |x_\lambda\rangle\langle x_\lambda| d\mu_\lambda$ . Note that these operators cannot act on the whole Hilbert space, but only on the subset  $\Phi$ .

## 5 Further topics

In this final section of this tutorial, we discuss some further topics from functional analysis and also from algebra. However, a fully rigorous treatment of these topics requires in general much more terminology and notation, which would render it too long for this introductory text. Nevertheless, we want to include these topics in order to show that mathematics has many aspects relevant to physics, which are somewhat less known to physicists.

### 5.1 Operator algebras

Operator algebras in some sense make abstract the notions of operators on a Hilbert space. The theory of operator algebras is relevant to physics, since there is quite a lot of mathematics literature dealing with topics from quantum mechanics, but in general, difficult to read for physicists.

The term *operator* is as before a linear mapping on a Hilbert space. The term *algebra* appears in two notions: on the one hand as the name of a mathematical field of the study of algebraic structures such as groups, rings, fields etc., on the other hand—as we use it here—as one of these algebraic structures, which can informally be described as a “vector space with multiplication”. Axiomatically, it is a quadruple  $(V, K, +, \cdot)$  consisting of a vector space  $V$  over a scalar field  $K$ , on which there is additionally defined a multiplication of two vectors  $\cdot : V \rightarrow V$  which fulfils the following conditions (for all  $a, b, c \in V$  and  $\alpha \in K$ ); as is customary, we suppress the multiplication symbol:

1. associativity:  $(ab)c = a(bc)$ ;<sup>15</sup>
2. distributivity:  $a(b+c) = ab+ac$  and  $(a+b)c = ac+bc$ ;
3. scalar multiplicativity:  $\alpha(ab) = (\alpha a)b = a(\alpha b)$ .

The algebra is *commutative*, if there holds  $ab = ba$ , and it is called *unital* (or algebra with identity) if there is a neutral element  $e \in V$ , such that there holds  $ae = ea = a$ .

The principal examples of such algebras are  $n \times n$  matrices  $M_n(K)$  over a field  $K$ . Moreover, the bounded linear operators on a Hilbert space  $\mathcal{H}$ , written as  $\mathcal{B}(\mathcal{H})$  (or  $L(\mathcal{H})$ ) form such an algebra as well as their subalgebras,

<sup>15</sup> Note that in Lie theory a Lie algebra with Lie bracket  $[\cdot, \cdot] : V \times V \rightarrow V$  is in general non-associative, and algebras in our sense are called associative algebras there. (The Lie bracket can be thought of an abstract version of the commutator of operators.)

like block matrices. Commutative algebras are e. g. the set of diagonal matrices or continuous functions  $C[a, b]$  on an interval or more generally on a compact set. An algebra which does not possess an identity is the continuous functions on an interval vanishing at the boundary  $C_0[a, b]$ .

An *involution* is a function  $\dagger : A \rightarrow A$  on an algebra  $A$  fulfilling

1. involutivity:  $x^{\dagger\dagger} = x$ ;
2. anti-linearity:  $(\alpha x + \beta y)^{\dagger} = \alpha^* x^{\dagger} + \beta^* y^{\dagger}$ ;
3. anti-multiplicativity  $(xy)^{\dagger} = y^{\dagger} x^{\dagger}$ .

An involutive algebra is an algebra with an involution. In mathematics, usually a star is used instead of a dagger ( $A^*$  instead of  $A^{\dagger}$ ), and such algebras are also called  $*$ -algebra; this is the star in  $C^*$ -algebra. The standard example of involutive algebras are complex square matrices or bounded operators with their adjoints, i. e. the transposed and complex conjugated matrix. The guiding idea is to make terminology from spectral theory abstract, i. e. to characterise them by their inner properties.

The topic of operator algebras is connected to functional analysis by *normed algebras*, i. e. algebras with a norm. A *Banach algebra* (similarly to a Banach space) is a normed algebra, where  $\|xy\| \leq \|x\| \|y\|$  for all  $x, y \in A$ ; this condition implies continuity of the multiplication. If there holds the (at first sight somewhat odd-looking) condition  $\|x^{\dagger}x\| = \|x\|^2$ , the algebra is called  $C^*$ -algebra. (Using the operator norm this holds e. g. for square matrices.)

## 5.2 Notions of convergence

In the preceding sections we already have mentioned the possibility to use different notions of convergence. Here we should consider in more detail some notions of convergence one usually encounters in mathematics literature. Let  $(x_n)_{n \in \mathbb{N}}$  be a sequence and  $x$  be an element in some Hilbert space  $\mathcal{H}$ :

1. The sequence  $(x_n)_{n \in \mathbb{N}}$  converges *strongly* or *in the norm* to  $x$ , if  $\lim_{n \rightarrow \infty} \|x_n - x\| = 0$  in  $\mathbb{R}$ .
2. It converges *weakly* to  $x$ , if for every  $y \in \mathcal{H}$  the sequence  $\langle y | x_n - x \rangle$  converges to zero in  $\mathbb{C}$ .

In case of doubt, one usually takes strong convergence; more generally, weak convergence can be defined on normed vector spaces and their dual. Strong convergence implies weak, but in infinite-dimensional spaces not vice versa: consider an infinite-dimensional Hilbert space, e. g. Fock space  $\{|0\rangle, |1\rangle, |2\rangle, \dots\}$ , and a sequence  $x_n = |n\rangle$ . This does not converge strongly (e. g. it is not a

Cauchy sequence), but weakly: for every bra-vector  $\langle \psi |$ ,  $\langle \psi | n \rangle$  converges to zero.

Given a sequence not of vectors but of operators, there are even more notions of convergence; we shall consider only the three most important. Let  $(A_n)_{n \in \mathbb{N}}$  be a sequence in  $\mathcal{B}(\mathcal{H})$ , the bounded operators, and  $A \in \mathcal{B}(\mathcal{H})$  be an operator.

1. The sequence  $(A_n)_{n \in \mathbb{N}}$  converges *in the norm* (or *uniformly*) to  $A$ , if  $\lim_{n \rightarrow \infty} \|A_n - A\| = 0$  in  $\mathbb{R}$ . Here,  $\|A\|$  is the as usual operator norm of eq. (6).
2. It converges *strongly* (more precisely: *with respect to the strong operator topology*) to  $x$ , if for every  $x \in \mathcal{H}$ , the sequence  $(Ax_n)_{n \in \mathbb{N}}$  in  $\mathcal{H}$  converges strongly to  $Ax \in \mathcal{H}$ .
3. It converges *weakly* (more precisely: *with respect to the weak operator topology*) to  $x$ , if for every  $x \in \mathcal{H}$ , the sequence  $(Ax_n)_{n \in \mathbb{N}}$  in  $\mathcal{H}$  converges weakly to  $Ax \in \mathcal{H}$ .

Convergence in the norm implies strong, and this implies weak convergence. To give counterexamples for the opposite, consider  $A_n := |n\rangle\langle n|$  and  $B_n := |n\rangle\langle 0|$ . Both sequences do not converge in the norm, but weakly to the null operator; the first one does so also strongly, the latter does not. In the mathematical-physics literature there often appear notions for which we shall give just the definitions, in order that the reader may grasp the idea, if he finds them somewhere (usually it is always good to think of square matrices):

1. A  $C^*$ -algebra is a a norm-closed involutive subalgebra of  $\mathcal{B}(\mathcal{H})$ ;
2. a von-Neumann algebra is a strongly (or equivalently weakly) closed involutive subalgebra of  $\mathcal{B}(\mathcal{H})$ .

It can be shown that every  $C^*$ -algebra in this sense fulfils the  $C^*$ -property from above. The finite-dimensional  $C^*$ -algebras—which always contain the identity—are easy to classify; they are the direct sums of simple  $C^*$ -algebras, which again are the full matrix algebras  $M_n(\mathbb{C})$  of complex  $n \times n$ -matrices.

The *commutant* of a set  $M \subseteq \mathcal{B}(\mathcal{H})$  is the subalgebra  $M' := \{A \in \mathcal{B}(\mathcal{H}) \mid (\forall B \in M)(AB = BA)\}$  of all elements in  $\mathcal{B}(\mathcal{H})$  commuting with every element in  $M$  (the notation should not be confused with the dual); for example, as an instance of Schur's lemma, if  $A = M_n(\mathbb{C})$ , then  $A' = \mathbb{C}\mathbb{1}$ . The bicommutant theorem says that  $M$  is a von-Neumann algebra, if and only if  $M = M''$ . von-Neumann algebras can be decomposed into so-called *factors* (of types  $I_n$ ,  $I_\infty$ ,  $II_1$ ,  $II_\infty$  and  $III$ ), where a factor is defined by  $M \cap M' = \mathbb{C}\mathbb{1}$ .

Coming back to quantum mechanics: the postulates of quantum mechanics say that there is a Hamiltonian operator  $\hat{H}$ , which generates the dynamics of a state



vector by the Schrödinger equation  $i\hbar \frac{d}{dt}|\Psi\rangle = \hat{H}|\Psi\rangle$ . If  $\hat{H}$  is time-independent, the solution is the unitary time evolution  $|\Psi(t)\rangle = e^{-i\hat{H}t/\hbar}|\Psi(0)\rangle$ . So for every time  $t$  there exist a unitary operator  $U(t)$  mapping  $|\Psi(0)\rangle$  to  $|\Psi(t)\rangle$ . In mathematical terms, the function  $U(t) = e^{-i\hat{H}t/\hbar}$  is a unitary-operator-valued function of the (time) parameter  $t \in \mathbb{R}$  which fulfils the following conditions:

1. Semigroup property:  $U(t+s) = U(t)U(s)$  for  $t, s \geq 0$  and
2. Strong continuity: for  $t \rightarrow t_0$ , there holds  $U(t)|\Psi\rangle \rightarrow U(t_0)|\Psi\rangle$  (it suffices to show this for  $t = 0$ ).

On the other hand, given such a function  $U$ , the generator  $\hat{H}$  can be recovered through formal derivation; Stone's theorem tells us that this is indeed mathematically possible. The one-parameter group is not necessarily norm-continuous: take the harmonic oscillator Hamiltonian  $\hat{H} = \hbar\omega(|n\rangle\langle n| + 1/2)$ , where  $U(t) = \sum_{n=0}^{\infty} e^{-\frac{i\hbar\omega(n+1/2)t}{\hbar}}|n\rangle\langle n|$ . The  $n + 1/2$  term in the exponent oscillates arbitrarily fast and therefore is not continuous (e. g. at  $t = 0$ ).

### 5.3 Polar decomposition

A complex number  $z = a + bi \in \mathbb{C}$  may be decomposed into polar representation: there holds  $z = re^{i\varphi}$  for  $r = |z| \in \mathbb{R}_0^+$  and  $\varphi \in [0; 2\pi)$ . This is generalised to the *polar decomposition* of an operator. The *absolute value* of an operators  $A \in \mathcal{B}(\mathcal{H})$  is defined as  $|A| := \sqrt{A^\dagger A}$ , and this operator is always positive semidefinite. In the finite-dimensional case  $A$  is a matrix, and  $\langle x|A^\dagger Ax\rangle = \langle Ax|Ax\rangle \geq 0$ , so we can diagonalise it and take the square root; by the spectral theorem, we can do the same for all bounded operators  $\mathcal{B}(\mathcal{H})$ .

The polar decomposition now states that there exist partial isometries  $U$  and  $V$  in  $\mathcal{B}(\mathcal{H})$  (i. e.  $U^\dagger U$  and  $UU^\dagger$  are projections on  $\mathcal{H}$ ), such that  $A = U|A| = |A|V$ . If the Hilbert space is finite-dimensional,  $U$  and  $V$  can be chosen to be unitary.

### 5.4 Classes of operators

An operator is called *compact*, if the image of a compact set is relatively compact, i. e., the closure of the image is compact; on a Hilbert space they are the uniform limit of finite-rank operators. The compact operators in some sense behave like matrices, e. g. the set of non-zero eigenvalues is discrete, so that we can define the Schatten- $p$ -norms as follows: the singular values of  $A$  are

the eigenvalues of  $|A|$  and their  $\ell^p$  norm is the norm of the operator

$$\|A\|_p := (\text{Tr } |A|^p)^{1/p}. \quad (17)$$

Those operators, for which the norm is finite, form the Schatten classes  $S_p$ . Special cases are the *trace class* operators  $S_1$  (such as density operators), Hilbert-Schmidt operators  $S_2$  and the (non-compact) bounded operators  $S_\infty = \mathcal{B}(\mathcal{H})$ . As is usually the case for  $p$ -norms, the 2-norm is induced by a scalar product,  $\langle A|B\rangle := \text{Tr } A^\dagger B$ , and Hölder's inequality,  $\|AB\|_r \leq \|A\|_p \cdot \|B\|_q$  with  $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$  and  $r, p, q \in [1; \infty]$ , holds.

### 5.5 Commutators of position and momentum

Concepts and methods from the theory of operators may be used to give a mathematical formulation of quantum mechanics. This shall be presented here in a short and simplified way. In general, one starts from the Heisenberg commutator relations  $[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}$ . Now consider two cases of these operators in one mode:

- a free particle: Hilbert space  $L^2(\mathbb{R})$  and
- a particle in a box: Hilbert space  $L^2[a; b]$ .

The spectrum of  $\hat{x}$  and  $\hat{p}$  is both  $\mathbb{R}$  in the first case, but  $[a; b]$  and discrete in the second case. This shows that the Heisenberg commutator relations do not specify the spectrum uniquely. On the other hand, we can (formally) use the Baker-Campbell-Hausdorff identity to get the commutator relations in the Weyl form:

$$e^{it\hat{x}_i} e^{is\hat{p}_j} = e^{-ist\delta_{ij}} e^{is\hat{p}_j} e^{it\hat{x}_i} \quad \text{for } t, s \in \mathbb{R}. \quad (18)$$

Note that this is a special case of the displacement operators  $\hat{D}(\alpha)$  for  $\alpha \in \mathbb{C}$ , usually used in discussing the harmonic oscillator and coherent states. Further, the operators appearing in this equation are bounded, thus easier to handle.

The commutators used in the Weyl form are not the usual commutators of operators  $[A, B] = AB - BA$  (in mathematics also known as Lie bracket in terms of Lie algebras) but a related concept, the group-theoretic commutator  $ABA^{-1}B^{-1}$ , which can be defined on any group and gives identity, if  $AB = BA$ . The following theorem says that these operators are essentially unique [13].

**Theorem 7** (Stone-von Neumann theorem). *Let  $A_i$  and  $B_i$  for  $i \in \{1, \dots, n\}$  be operators on a Hilbert space  $\mathcal{H}$ , such that for all  $i, j \in \{1, \dots, n\}$  the WEYL commutator*

relations  $e^{itA_i}e^{isA_j} = e^{isA_j}e^{itA_i}$  and  $e^{itB_i}e^{isB_j} = e^{isB_j}e^{itB_i}$  and  $e^{itA_i}e^{isB_j} = e^{-ist\delta_{ij}}e^{isB_j}e^{itA_i}$

are fulfilled. If the operators act irreducibly<sup>16</sup> on  $\mathcal{H}$ , then there exists a unitary mapping  $U: \mathcal{H} \rightarrow L^2(\mathbb{R}^n)$ , unique up to a phase  $e^{i\varphi}$ , such that  $Ue^{itA_i}U^\dagger = e^{it\hat{x}_i}$  and  $Ue^{itB_j}U^\dagger = e^{it\hat{p}_j}$  where  $\hat{x}_i = x_i$  and  $\hat{p}_j = \frac{\hbar}{i} \frac{\partial}{\partial x_j}$  are the usual position and momentum operators on  $\mathbb{R}^n$ .

The intuitive reason behind this is that the Heisenberg commutators exhibit the local properties (at every point  $x$ ), but not the global ones. The Weyl commutators take into account global properties; remember that the exponentiated operators are translations of momentum and position, respectively, and thus connect different points—in the case of the particle in a box, one would “hit the wall” if the shift distance is too large. (This is related to the theory of Lie groups and Lie algebras.)

## 5.6 Mathematical structure of quantum mechanics

Thinking of the postulates of quantum mechanics (Hilbert space, state vector or density matrix etc.), one may ask what precise reason there is to use these postulates. There is a mathematical modelling of quantum mechanics, which could explain this. Starting from the assumption that quantum mechanics is a theory of what is measurable, we must first state the measurable quantities (observables). A state of a system then is an object (a function) which maps to every observable its expectation value.

The crucial assumption now is to describe observables by elements of an operator algebra with certain commutation relations. A state is a functional on the algebra of these operators, which should be continuous, linear, positive and normalised. A mathematical theorem, the Gelfand, Neumark and Segal (GNS) representation, then tells us that, if these operators form a  $C^*$ -Algebra and we have such a state, these operators may be represented by a subalgebra of the bounded linear operators  $\mathcal{A} = \mathcal{B}(\mathcal{H})$  on some Hilbert space  $\mathcal{H}$ . A state in mathematics is a functional  $\varphi: \mathcal{A} \rightarrow \mathbb{C}$  which is positive (maps positive operators to non-negative numbers) and which fulfils  $\|\varphi\| = \varphi(\mathbb{1}) = 1$  (the name of such a functional even in mathematics is “state”). Comparing with physics, a state is a density operator  $\rho$ ; this operator generates a functional by  $\varphi(A) := \text{Tr } \rho A$ , which is a state in

the mathematical sense. However, not all mathematical states can be described this way, and the notion there is somewhat more general, but under some continuity assumptions, the converse also holds.

The states moreover form a convex and compact set. Within these sets there exist extremal points, i. e. points which cannot be written as true convex combinations of other states. Formally: a state is extremal, if from  $\varphi = \lambda\varphi_1 + \mu\varphi_2$  with  $\lambda, \mu > 0$  and  $\lambda + \mu = 1$ , it follows that  $\varphi_1 = \varphi_2 = \varphi$ . Extremal points in the states space are the pure states, other states are mixed states. According to the Krein-Milman theorem, a convex and compact set (such as the states) is the closed convex span of its extremal states. For a two-dimensional Hilbert spaces this is very well illustrated by the Bloch sphere. The infinitely many points on the surface correspond to the pure states, the convex hull (the interior) to the mixed states. (One can compare the example of a probability distribution with finitely many outcomes: there the extremal points are the probability distributions with one certain outcome.).

## 5.7 More examples

It is beyond the scope of this introduction to present more than an overview of all the applications of modern mathematics, i. e. still mathematics of the 20th century, to physics. But at the end of this introduction we would like to give some examples, where theorems from mathematics have direct application in physics:

- The *Schmidt decomposition* was derived by Erhard Schmidt in his 1906 doctoral thesis [14] and reappeared in Schrödinger's papers on entanglement in 1936 [15].
- The idea of realising quantum operations as unitary operators on larger Hilbert spaces and Kraus operators [16] is related to Stinespring's dilation theorem from 1955 [17].
- The quantum de-Finetti theorem was derived by Størmer in 1969 [18], then by Hudson and Moody [19] and in physics by Caves, Fuchs and Schack [20].
- The proof of entropic uncertainty relations for Rényi entropies by Maassen and Uffink [21] essentially relies on the Riesz-Thorin interpolation theorem found in several textbooks.

## 6 Summary

It could not have been the aim of this introduction to give a complete overview of functional analysis or its use in

<sup>16</sup> That is, if only  $\{0\}$  and  $\mathcal{H}$  are invariant closed subspaces of all  $e^{itA_i}$  and  $e^{itB_i}$ . If this is not the case, the theorem can be adapted accordingly.

physics and quantum mechanics, but to give the reader examples of mathematical thinking in a language appropriate for physicists. In this spirit we discussed some topics of functional analysis which are relevant for physics. It is clear that there is a vast amount of mathematical knowledge relevant to quantum mechanics and physics in general, but which is unfortunately not so widely known by physicists.

**Key words.** Functional analysis, quantum mechanics, tutorial.

## References

- [1] D. Werner, *Funktionalanalysis* (Springer 5<sup>th</sup> ed., 2005).
- [2] M. Reed and B. Simon, *Methods of modern mathematical physics*, Vol. I: Functional Analysis (Revised and enlarged edition 1980), Vol. II: Fourier Analysis, Self-Adjointness (1975).
- [3] J. B. Conway, *A Course in Functional Analysis* (Springer 2<sup>nd</sup> ed., 1990).
- [4] F. Gieres, *Reports on Progress in Physics*, **63**, 1893–1931 (2000).
- [5] G. Bonneau, J. Faraut, and G. Valent, *American Journal of Physics* **69**, 322–331 (2001).
- [6] J. Elstrodt, *Maß- und Integrationstheorie* (Springer 4<sup>th</sup> ed., 2005).
- [7] I. M. Gelfand and A. G. Kostyuchenko, *Dokl. Akad. Nauk SSSR* **103**, 349–352 (1955), translated in Izrail M. Gelfand, *Collected papers*, Vol. I (Springer 1987), 505–509.
- [8] I. M. Gelfand and N. J. Wilenkin, *Verallgemeinerte Funktionen (Distributionen)* Vol. IV (VEB Deutscher Verlag der Wissenschaften, 1964).
- [9] G. G. Gould, *J. London Math. Soc.* **43**, 745–754 (1968).
- [10] F. Schwabl, *Quantenmechanik* (Springer, 6<sup>th</sup> ed., 2002).
- [11] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer, 1932).
- [12] F. Scheck, *Theoretische Physik 2* (Springer 2<sup>nd</sup> ed., 2006).
- [13] B. C. Hall, *Quantum Theory for Mathematicians* (Springer, 2013).
- [14] E. Schmidt, *Math. Ann.* **63**, 433–476 (1907).
- [15] E. Schrödinger, *Proc. Cambridge Phil. Soc.* **31**, 555–563 (1935), and 32, 446–452 (1936).
- [16] K. Kraus, *States, effects, and operations* (Springer, 1983).
- [17] W. F. Stinespring, *Proc. Amer. Math. Soc.* **6**, 211–216 (1955).
- [18] E. Størmer, *J. Funct. Anal.* **3**, 48–68 (1969).
- [19] R. L. Hudson and G. R. Moody, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **33**, 343–351 (1976).
- [20] C. M. Caves, C. A. Fuchs, and R. Schack, *J. Math. Phys.* **43**, 4537–4559 (2002).
- [21] H. Maassen and J. B. M. Uffink, *Phys. Rev. Lett.* **60**, 1103–1106 (1988).