Schrödinger Operators

Gerald Teschl

Gerald Teschl
Institut für Mathematik
Nordbergstraße 15
Universität Wien
1090 Wien, Austria

E-mail address: Gerald.Teschl@univie.ac.at
URL: http://www.mat.univie.ac.at/~gerald/

1991 Mathematics subject classification. 81-01, 81Qxx

Abstract. This manuscript provides a brief introduction to Schrödinger operators. We start with the mathematical foundations of quantum mechanics and prove the spectral theorem. Then we consider quantum dynamics and establish Stone's and the RAGE theorem.

After a detailed study of the free Schrödinger operator we investigate self-adjointness of atomic Schrödinger operators and their essential spectrum, in particular the HVZ theorem. Finally we have a look at scattering theory and prove asymptotic completeness in the short range case.

Keywords and phrases. Schrödinger operators, quantum mechanics, spectral theory.

Typeset by $\mathcal{AMS}$-LaTeX and Makeindex.
Version: October 19, 2004

Contents

Preface

The present manuscript was written for my course *Schrödinger Operators* given at the University of Vienna in Winter 1999 and Summer 2002. It is supposed to give a brief but rather self contained introduction to the field of Schrödinger operators with emphasis on applications in quantum mechanics. I assume some previous experience with Hilbert spaces and bounded linear operators which should be covered in any basic course on functional analysis. However, all necessary results are reviewed in the first chapter. The material presented is highly selective and many important and interesting topics are not touched.

It is available from

`http://www.mat.univie.ac.at/~gerald/ftp/book-schroe/`

Acknowledgments

I'd like to thank Volker Enß for making his lecture notes available to me and Maria Hoffmann-Ostenhof, Harald Rindler, and Karl Unterkofler for pointing out errors in previous versions.

Gerald Teschl

Vienna, Austria
February, 2002

Part 1

Mathematical Foundations of Quantum Mechanics

Chapter 0

Preliminaries

I assume that the reader has some basic familiarity with measure theory and functional analysis. For convenience, all facts needed are stated in this chapter (without proofs). If you feel comfortable with terms like *Lebesgue integral*, *dominated convergence*, *Banach space*, or *bounded linear operator*, you can skip this entire chapter. However, you might want to at least browse through it to refresh your memory. Good references are [**2**] or [**14**].

0.1. Borel measures

Recall that a **σ-algebra** Σ is a collection of subsets of a given set M such that

- $M \in \Sigma$,
- Σ is closed under countable unions,
- Σ is closed under complements.

It follows from the de Morgan rules that Σ is closed under countable intersections as well. The intersection of any family of σ-algebras $\{\Sigma_\alpha\}$ is again a σ-algebra and for any collection S of subsets there is a unique smallest σ-algebra containing S (namely the intersection of all σ-algebra containing S). It is called the σ-algebra generated by S.

The **Borel σ-algebra** of $\mathbb{R}^n$ is defined to be the σ-algebra generated by all open (respectively all closed) sets and is denoted by $\mathfrak{B}^n$. We will abbreviate $\mathfrak{B} = \mathfrak{B}^1$ and call sets $\Omega \in \mathfrak{B}^n$ simply **Borel sets**.

A **measure** μ is a map $\mu : \Sigma \to [0, \infty]$ such that

- $\mu(\emptyset) = 0$,

- $\mu(\bigcup_{j=1}^{\infty} \Omega_j) = \sum_{j=1}^{\infty} \mu(\Omega_j)$ if $\Omega_j \cap \Omega_k = \emptyset$ for all j, k.

It is called **σ-finite** if there is a countable cover $\{\Omega_j\}_{j=1}^{\infty}$ of M with $\mu(\Omega_j) < \infty$ for all j. We will assume all measures to be σ-finite. It is called **finite** if $\mu(M) < \infty$.

A set $\Omega \in \Sigma$ is called a **support** for μ if $\mu(M \backslash \Omega) = 0$. A property is said to hold **μ-almost everywhere** (a.e.) if the set for which it holds is a support for μ or, equivalently, if the set where it does not hold has measure zero.

A measure on the Borel σ-algebra is called a **Borel measure** if $\mu(C) < \infty$ for any compact set C. Borel measures on $\mathfrak{B}^n$ automatically satisfy the following regularity property

$$\mu(\Omega) = \sup_{C \subseteq \Omega, C \text{ compact}} \mu(C) = \inf_{\Omega \subseteq O, O \text{ open}} \mu(O), \tag{0.1}$$

which shows that μ is uniquely determined by its value on compact respectively open sets. (This is not true in general if $\mathbb{R}^n$ is replaced by an arbitrary topological space.) We will only consider the case of Borel measures on $\mathfrak{B}^n$.

To every Borel measure on $\mathfrak{B}$ we can assign its distribution function

$$\mu(\lambda) = \begin{cases} \mu((x, 0]), & x < 0 \\ 0, & \lambda = 0 \\ \mu((0, x]), & x > 0 \end{cases} \tag{0.2}$$

which is right continuous and non-decreasing. Conversely, given a right continuous non-decreasing function $\mu : \mathbb{R} \to \mathbb{R}$ we can set

$$\mu(\Omega) = \begin{cases} \mu(b) - \mu(a), & \Omega = (a, b] \\ \mu(b) - \mu(a-), & \Omega = [a, b] \\ \mu(b-) - \mu(a), & \Omega = (a, b) \\ \mu(b-) - \mu(a-), & \Omega = [a, b) \end{cases}, \tag{0.3}$$

where $\mu(a-) = \lim_{\varepsilon \downarrow 0} \mu(a - \varepsilon)$. Then μ gives rise to a unique Borel measure.

Theorem 0.1. *For every right continuous non-decreasing function $\mu : \mathbb{R} \to \mathbb{R}$ there exists a unique Borel measure μ which extends (0.3). Two different functions generate the same measure if and only if they differ by a constant.*

Example. Suppose $\mu(\lambda) = 0$ for $\lambda < 0$ and $\mu(\lambda) = 1$ for $\lambda \geq 0$. Then we obtain the so-called **Dirac measure** at 0, which is given by $\mu(\Omega) = 1$ if $0 \in \Omega$ and $\mu(\Omega) = 0$ if $0 \notin \Omega$. ◇

Example. Suppose $\mu(\lambda) = \lambda$, then the associated measure is the ordinary **Lebesgue measure** on $\mathbb{R}$. We will abbreviate the Lebesgue measure of a Borel set Ω by $|\Omega|$. ◇

A function $f : M \to \mathbb{R}$ is called **measurable** if $f^{-1}(\Omega) \in \Sigma$ for every $\Omega \in \mathfrak{B}$. A complex-valued function is called measurable if both its real and imaginary parts are. Clearly it suffices to check this condition for every set Ω in a collection of sets which generate $\mathfrak{B}$, say for all open intervals. If Σ is the Borel σ-algebra, we will call a measurable function also **Borel function**. Note that, in particular, any continuous function is Borel.

Moreover, sometimes it is also convenient to allow $\pm\infty$ as possible values for f, that is, functions $f : M \to \overline{\mathbb{R}}$, $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$. In this case $\Omega \subseteq \overline{\mathbb{R}}$ is called Borel if $\Omega \cap \mathbb{R}$ is.

The set of all measurable functions forms an algebra.

Lemma 0.2. *Suppose f, g are measurable functions. Then the sum $f + g$ and the product fg is measurable.*

Moreover, the set of all measurable functions is closed under all important limiting operations.

Lemma 0.3. *Suppose f_n is a sequence of measurable functions, then*

$$\inf_{n\in\mathbb{N}} f_n, \quad \sup_{n\in\mathbb{N}} f_n, \quad \liminf_{n\to\infty} f_n, \quad \limsup_{n\to\infty} f_n \tag{0.4}$$

are measurable as well.

It follows that if f and g are measurable functions, so are $\min(f, g)$, $\max(f, g)$, $|f| = \max(f, -f)$, $f^{\pm} = \max(\pm f, 0)$.

0.2. Integration

Now we can define the integral for measurable functions as follows. A measurable function $s : M \to \mathbb{R}$ is called **simple** if its range is finite, that is, if $s = \sum_{j=1}^{n} c_j \chi_{\Omega_j}$, $\Omega_j \in \Sigma$. Here χ_Ω is the **characteristic function** of Ω, that is, $\chi_\Omega(\lambda) = 1$ if $\lambda \in \Omega$ and $\chi_\Omega(\lambda) = 0$ else.

For a positive simple function we define its **integral** as

$$\int s(\lambda) d\mu(\lambda) = \sum_{j=1}^{n} c_j \mu(\Omega_j). \tag{0.5}$$

It can be extended to the set of all positive measurable functions by approximating f by a nondecreasing sequence s_n of simple functions and set

$$\int f(\lambda) d\mu(\lambda) = \lim_{n\to\infty} \int_{\mathbb{R}} s_n(\lambda) d\mu(\lambda). \tag{0.6}$$

If the integral is finite for both the positive and negative part $f^{\pm}$ of an arbitrary measurable function f, we call f **integrable** and set

$$\int f(\lambda) d\mu(\lambda) = \int_{\mathbb{R}} f^{+}(\lambda) d\mu(\lambda) - \int f^{-}(\lambda) d\mu(\lambda). \tag{0.7}$$

The set of all integrable functions is denoted by $L^1(M)$. Similarly, we handle the case where f is complex-valued by calling f integrable if both the real and imaginary part are and setting

$$\int f(\lambda)d\mu(\lambda) = \int_{\mathbb{R}} \mathrm{Re}(f(\lambda))d\mu(\lambda) + \mathrm{i}\int \mathrm{Im}(f(\lambda))d\mu(\lambda). \tag{0.8}$$

Note that f is integrable if and only if $|f|$ is.

It is custom to set

$$\int_{\Omega} f(\lambda)d\mu(\lambda) = \int \chi_{\Omega}(\lambda)f(\lambda)d\mu(\lambda). \tag{0.9}$$

The integral has the following properties

Theorem 0.4. *Suppose f and g are integrable functions, then so is any linear combination and*

$$\int (\alpha f(\lambda) + \beta g(\lambda))d\mu(\lambda) = \alpha \int f(\lambda)d\mu(\lambda) + \beta \int g(\lambda)d\mu(\lambda). \tag{0.10}$$

Moreover,

$$|\int f(\lambda)d\mu(\lambda)| \le \int |f(\lambda)|d\mu(\lambda) \tag{0.11}$$

and

$$|\int_{\Omega} f(\lambda)d\mu(\lambda)| \le \sup_{\lambda\in\Omega} |f(\lambda)|\mu(\Omega). \tag{0.12}$$

If f, g are real-valued and satisfy $f \le g$, we have

$$\int f(\lambda)d\mu(\lambda) \le \int g(\lambda)d\mu(\lambda). \tag{0.13}$$

In particular, by $|f + g| \le |f| + |g|$ we infer

$$\int |f(\lambda) + g(\lambda)|d\mu(\lambda) \le \int |f(\lambda)|d\mu(\lambda) + \int |g(\lambda)|d\mu(\lambda). \tag{0.14}$$

In addition, our integral is well behaved with respect to limiting operations. The most important results in this respect are

Theorem 0.5 (monotone convergence)**.** *Let f_n be a monotone non-decreasing sequence of positive measurable functions and set $f = \lim_{n\to\infty} f_n$. If $\int f_n d\mu \le C$ for some finite constant C, then f is integrable and $\lim_{n\to\infty} \int |f - f_n|d\mu = 0$.*

Theorem 0.6 (dominated convergence)**.** *Let f_n be a convergent sequence of measurable functions and set $f = \lim_{n\to\infty} f_n$. Suppose there is an integrable function g such that $|f_n| \le g$. Then f is integrable and $\lim_{n\to\infty} \int |f - f_n|d\mu = 0$.*

Finally, we note that if there are two measures μ_1 and μ_2 on Σ_1 and Σ_2, respectively, then there is a unique measure, the **product measure**, $\mu_1 \times \mu_2$ on the σ-algebra generated by $\Sigma_1 \times \Sigma_2$ which satisfies

$$\mu_1 \times \mu_2(\Omega_1 \times \Omega_2) = \mu_1(\Omega_1)\mu_2(\Omega_2), \quad \Omega_j \in \Sigma_j,\ j = 1, 2. \tag{0.15}$$

Moreover, it can be shown that if f is a measurable function on $M_1 \times M_2$, then $f(., \lambda_2)$ is measurable on M_1 for every λ_2 and $f(\lambda_1, .)$ is measurable on M_2 for every λ_1. Moreover, if these functions are integrable, then $\int f(., \lambda_2) d\mu_2(\lambda_2)$ and $\int f(\lambda_1, .) d\mu_1(\lambda_1)$ are measurable as well.

Theorem 0.7 (Fubini)**.** *Let f be a measurable function on $M_1 \times M_2$ and let μ_1, μ_2 be measures on M_1, M_2, respectively.*

Then

$$\int \left(\int |f(\lambda_1, \lambda_2)| d\mu_1(\lambda_1) \right) d\mu_2(\lambda_2) < \infty \tag{0.16}$$

if and only if

$$\int \left(\int |f(\lambda_1, \lambda_2)| d\mu_2(\lambda_2) \right) d\mu_1(\lambda_1) < \infty \tag{0.17}$$

and if one (and thus both) of these integrals is finite, then

$$\begin{aligned} \iint f(\lambda_1, \lambda_2) d\mu_1 \times \mu_2(\lambda_1, \lambda_2) &= \int \left(\int f(\lambda_1, \lambda_2) d\mu_1(\lambda_1) \right) d\mu_2(\lambda_2) \\ &= \int \left(\int f(\lambda_1, \lambda_2) d\mu_2(\lambda_2) \right) d\mu_1(\lambda_1). \end{aligned} \tag{0.18}$$

0.3. The decomposition of measures

The results in this section are somewhat more special and will not be needed until Section 3.2.

Let μ, ν be two measures on a measure space (M, Σ). They are called **mutually singular** if they are supported on disjoint sets. That is, there is a measurable set Ω such that $\mu(\Omega) = 0$ and $\nu(M \backslash \Omega) = 0$. The measure ν is called **absolutely continuous** with respect to μ if $\mu(\Omega) = 0$ implies $\nu(\Omega) = 0$.

The two main results read

Theorem 0.8 (Radon-Nikodym)**.** *Let μ, ν be two measures on a measure space (M, Σ). Then ν is absolutely continuous with respect to μ if and only if there is a positive measurable function f such that*

$$\nu(\Omega) = \int_\Omega f(\lambda) d\mu(\lambda) \tag{0.19}$$

for every $\Omega \in \Sigma$. The function f is determined uniquely a.e. with respect to μ and is called the **Radon-Nikodym derivative** $\frac{d\nu}{d\mu}$ *of ν with respect to μ.*

Theorem 0.9 (Lebesgue decomposition). *Let μ, ν be two measures on a measure space (M, Σ). Then ν can be uniquely decomposed as $\nu = \nu_{ac} + \nu_{sing}$, where ν_{ac} and ν_{sing} are mutually singular and ν_{ac} is absolutely continuous with respect to μ.*

In the case of a Borel measure μ on $\mathfrak{B}$ the situation is as follows:

We call

$$(D\mu)(\lambda) = \lim_{\varepsilon \downarrow 0} \frac{\mu(I_\varepsilon)}{|I_\varepsilon|}, \tag{0.20}$$

the Radon-Nikodym derivative of μ at λ provided the above limit exists for any sequence of intervals I_ε, which contain λ and have length $|I_\varepsilon| = \varepsilon$.

Note that $(D\mu)(\lambda)$ exists if and only if $\mu(\lambda)$ (as defined in (0.2)) is differentiable at λ and $(D\mu)(\lambda) = \mu'(\lambda)$ in this case.

Theorem 0.10. *The derivative $D\mu$ of μ has the following properties:*

- *$D\mu$ exists a.e. with respect to Lebesgue measure.*
- *$D\mu \in L^1(\mathbb{R})$.*
- *$D\mu$ is the Radon-Nikodym derivative of the absolutely continuous part of μ with respect to Lebesgue measure, that is,*

$$\mu_{ac}(\Omega) = \int_\Omega (D\mu)(\lambda) d\lambda.$$

In particular, μ is singular with respect to Lebesgue measure if and only if $(D\mu)(\lambda) = 0$ a.e. with respect to Lebesgue measure.

To find supports for the absolutely and singularly continuous part, it is useful to consider

$$(\overline{D}\mu)(\lambda) = \limsup_{\varepsilon \downarrow 0} \frac{\mu(I_\varepsilon)}{|I_\varepsilon|} \quad \text{and} \quad (\underline{D}\mu)(\lambda) = \liminf_{\varepsilon \downarrow 0} \frac{\mu(I_\varepsilon)}{|I_\varepsilon|}. \tag{0.21}$$

Then

Theorem 0.11. *The set $\{\lambda | (\underline{D}\mu)(\lambda) = \infty\}$ is a support for the singularly and $\{\lambda | (\overline{D}\mu)(\lambda) < \infty\}$ is a support for the absolutely continuous part.*

0.4. Banach spaces

A **normed linear space** X is a vector space X over $\mathbb{C}$ (or $\mathbb{R}$) with a real-valued function (the **norm**) $\|.\|$ such that

- $\|\psi\| \geq 0$ for all $\psi \in X$ and $\|\psi\| = 0$ if and only if $\psi = 0$,
- $\|z\,\psi\| = |z|\,\|\psi\|$ for all $z \in \mathbb{C}$ and $\psi \in X$, and
- $\|\psi + \varphi\| \leq \|\psi\| + \|\varphi\|$ for all $\psi, \varphi \in X$ **(triangle inequality).**

Clearly, the norm is continuous.

The norm gives rise to a metric

$$d(\psi, \varphi) = \|\psi - \varphi\| \tag{0.22}$$

and hence concepts like convergence, Cauchy sequence, and completeness are well-defined for a normed linear space. In particular, a complete normed linear space is called a **Banach space**.

For a given set of vectors $\{\psi_j\}_{j\in J}$ the **span** $\mathrm{span}\{\psi_j\}_{j\in J}$ is the set of all finite linear combinations of the vectors ψ_j. The set $\{\psi_j\}_{j\in J}$ is called **linearly independent** if any finite subset is and it is called **total** if its span is dense. A Banach space is called **separable** if it contains a countable dense set or, equivalently, if it contains a countable total set.

Example. The set $C(K)$ of all continuous functions on a compact interval K together with the sup norm

$$\|f\|_\infty = \sup_{x\in K} |f(x)| \tag{0.23}$$

is a Banach space. (This follows since the uniform limit of continuous functions is again continuous.) It is even separable since the set of all polynomials is dense by the Stone-Weierstrass theorem. ⋄

A linear map A between two Banach spaces X and Y will be called a **(linear) operator**

$$A : \mathfrak{D}(A) \subseteq X \to Y. \tag{0.24}$$

The linear subspace $\mathfrak{D}(A)$ on which A is defined, is called the **domain** of A and is usually required to be dense. The operator A is called **bounded** if the following operator norm

$$\|A\| = \sup_{\|\psi\|_X=1} \|A\psi\|_Y \tag{0.25}$$

is finite.

Theorem 0.12. *An operator A is bounded if and only if it is continuous.*

Moreover, if A is bounded and densely defined, it is no restriction to assume that it is defined on all of X.

Theorem 0.13. *If A is bounded and $\mathfrak{D}(A)$ is dense, there is a unique extension of A to X, which has the same norm.*

Proof. Since a bounded operator maps Cauchy sequences to Cauchy sequences, this extension is clearly given by

$$A\psi = \lim_{n\to\infty} A\psi_n, \qquad \psi_n \in \mathfrak{D}(A), \quad \psi \in X. \tag{0.26}$$

To show that this definition is independent of the sequence $\psi_n \to \psi$, let $\phi_n \to \psi$ be a second sequence and observe

$$\|A\psi_n - A\phi_n\| = \|A(\psi_n - \phi_n)\| \le \|A\|\|\psi_n - \phi_n\| \to 0. \tag{0.27}$$

To prove that this extension is linear and has the same norm is left as an exercise. □

The set of all bounded linear operators from X to Y is denoted by $\mathfrak{L}(X,Y)$ and, together with the operator norm (0.25), it is again a Banach space. If $X = Y$ we write $\mathfrak{L}(X,X) = \mathfrak{L}(X)$.

An operator in $\mathfrak{L}(X,\mathbb{C})$ is called a **bounded linear functional** and the space $X^* = \mathfrak{L}(X,\mathbb{C})$ is called the dual space of X. A sequence ψ_n is said to **converge weakly** $\psi_n \rightharpoonup \psi$ if $\ell(\psi_n) \to \ell(\psi)$ for every $\ell \in X^*$.

The following important result will be needed later on.

Theorem 0.14 (uniform boundedness principle)**.** *Let X and Y be two Banach spaces and $\mathcal{F}$ a family of operators in $\mathfrak{L}(X,Y)$. If for each $x \in X$ the set $\{\|Ax\|_Y \,|\, A \in \mathcal{F}\}$ is bounded, then so is $\{\|A\| \,|\, A \in \mathcal{F}\}$.*

In particular, every weakly convergent sequence is bounded.

0.5. Lebesgue spaces

Let (M,μ) be a measure space and $p \ge 1$. Two measurable functions $f : M \to \mathbb{C}$ are considered equivalent if they only differ on a set of measure zero. The set of all equivalence classes together with the norm

$$\|f\|_p = \left(\int_M |f(x)|^p d\mu(x)\right)^{1/p} \tag{0.28}$$

forms a normed space denoted by $L^p(M,d\mu)$. If $d\mu$ is the Lebesgue measure on $\mathbb{R}^n$ we simply write $L^p(M)$.

Theorem 0.15. *Set space $L^p(M)$ is a Banach space (i.e. complete).*

Even though the elements of $L^p(M,d\mu)$ are strictly speaking equivalence classes of functions, we will still call them functions for notational convenience. However, note that for $f \in L^p(M,d\mu)$ the value $f(x)$ is not well defined (unless there is a continuous representative).

Theorem 0.16. *If $M \subseteq \mathbb{R}^n$ and μ is a Borel measure, then the set $C_0^\infty(M)$ of all smooth functions with compact support is dense in $L^p(M,d\mu)$, $1 \le p < \infty$.*

The set of all equivalence classes (as in the previous example) together with the norm

$$\|f\|_\infty = \inf\{C \,|\, |f(x)| \le C \text{ a.e. } x \in M\} \tag{0.29}$$

is also a Banach space denoted by $L^\infty(M, d\mu)$.

Theorem 0.17 (Hölder's inequality)**.** *Let* (M, μ) *be a measure space and let* $1 \le p, q, r \le \infty$ *such that*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{r}. \tag{0.30}$$

Suppose $f \in L^p(M, d\mu)$ *and* $q \in L^q(M, d\mu)$*, then* $f\,g \in L^r(M, d\mu)$ *and*

$$\|f\,g\|_r \le \|f\|_p \|g\|_q. \tag{0.31}$$

Example. Let $p, q > 1$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Then every element $f \in L^p(M, d\mu)$ gives rise to a bounded linear functional on $L^q(M, d\mu)$ via

$$g \mapsto \int_M f(x) g(x) d\mu(x). \tag{0.32}$$

One can even show $L^p(M, d\mu)^* \simeq L^q(M, d\mu)$ if $1 \le p < \infty$. $\diamond$

Chapter 1

Hilbert spaces

The phase space in classical mechanics is the Euclidean space $\mathbb{R}^{2n}$ (for the n position and n momentum coordinates). In quantum mechanics the phase space is always a Hilbert space $\mathfrak{H}$. Hence the geometry of Hilbert spaces stands at the outset of our investigations.

1.1. Hilbert spaces

Suppose $\mathfrak{H}$ is a vector space. A map $\langle ., .. \rangle : \mathfrak{H} \times \mathfrak{H} \to \mathbb{C}$ is called skew linear form if it is conjugate linear in the first and linear in the second argument, that is,

$$\begin{array}{rcl} \langle z_1\psi_1 + z_2\psi_2, \varphi\rangle & = & z_1^*\langle \psi_1, \varphi\rangle + z_2^*\langle \psi_2, \varphi\rangle \\ \langle \psi, z_1\varphi_1 + z_2\varphi_2\rangle & = & z_1\langle \psi, \varphi_1\rangle + z_2\langle \psi, \varphi_2\rangle \end{array}, \quad z_1, z_2 \in \mathbb{C}. \tag{1.1}$$

A skew linear form satisfying the requirements

- $\langle \psi, \psi\rangle > 0$ for $\psi \neq 0$ and
- $\langle \psi, \varphi\rangle = \langle \varphi, \psi\rangle^*$

is called **inner product** or **scalar product**. Associated with every scalar product is a norm

$$\|\psi\| = \sqrt{\langle \psi, \psi\rangle} \tag{1.2}$$

(we will prove in a moment that this is indeed a norm). The pair $(\mathfrak{H}, \langle ., .. \rangle)$ is called **inner product space**. If $\mathfrak{H}$ is complete with respect to the above norm, it is called a **Hilbert space**. It is no restriction to assume that $\mathfrak{H}$ is complete since one can easily replace it by its completion.

Example. The space $L^2(M, d\mu)$ is a Hilbert space with scalar product given by

$$\langle f, g \rangle = \int_M f(x)^* g(x) d\mu(x). \qquad (1.3)$$

Similarly, the set of all square summable sequences $\ell^2(\mathbb{N})$ is a Hilbert space with scalar product

$$\langle f, g \rangle = \sum_{j\in\mathbb{N}} f_j^* g_j. \qquad (1.4)$$

(Note that the second example is a special case of the first one; take $M = \mathbb{R}$ and μ is a sum of Dirac measures.) ⋄

A vector $\psi \in \mathfrak{H}$ is called **normalized** or **unit vector** if $\|\psi\| = 1$. Two vectors $\psi, \varphi \in \mathfrak{H}$ are called **orthogonal** or **perpendicular** ($\psi \perp \varphi$) if $\langle \psi, \varphi \rangle = 0$ and **parallel** if one is a multiple of the other. If ψ and φ are orthogonal we have the **Pythagorean theorem**:

$$\|\psi + \varphi\|^2 = \|\psi\|^2 + \|\varphi\|^2, \qquad \psi \perp \varphi, \qquad (1.5)$$

which is straightforward to check.

Suppose φ is a unit vector, then the projection of ψ in the direction of φ is given by

$$\psi_\| = \langle \varphi, \psi \rangle \varphi \qquad (1.6)$$

and $\psi_\perp$ defined via

$$\psi_\perp = \psi - \langle \varphi, \psi \rangle \varphi \qquad (1.7)$$

is perpendicular to φ since $\langle \varphi, \psi_\perp \rangle = \langle \varphi, \psi - \langle \varphi, \psi \rangle \varphi \rangle = \langle \varphi, \psi \rangle - \langle \varphi, \psi \rangle \langle \varphi, \varphi \rangle = 0$. Taking any other vector parallel to φ it is easy to see

$$\|\psi - c\varphi\|^2 = \|\psi_\perp\|^2 + |c - \langle \varphi, \psi \rangle|^2 \qquad (1.8)$$

and hence $\psi_\| = \langle \varphi, \psi \rangle \varphi$ is the unique vector parallel to φ which is closest to ψ.

As a first consequence we obtain the **Cauchy-Schwarz inequality**:

$$|\langle \psi, \varphi \rangle| \le \|\psi\| \, \|\varphi\| \qquad (1.9)$$

with equality if and only if ψ and φ are parallel. In fact, it suffices to prove the case $\|\varphi\| = 1$. But then the claim follows from $\|\psi\|^2 = |\langle \varphi, \psi \rangle|^2 + \|\psi_\perp\|^2$.

Note that the Cauchy-Schwarz inequality implies that the scalar product is continuous in both variables.

As another consequence we infer that the map $\|.\|$ is indeed a norm. Only the triangle inequality is nontrivial. It follows from the Schwarz inequality since

$$\|\psi + \varphi\|^2 = \|\psi\|^2 + \langle \psi, \varphi \rangle + \langle \varphi, \psi \rangle + \|\varphi\|^2 \le (\|\psi\| + \|\varphi\|)^2. \qquad (1.10)$$

These results can also be generalized to more than one vector. A set of vectors $\{\varphi_j\}$ is called **orthonormal set** if $\langle\varphi_j,\varphi_k\rangle = 0$ for $j \neq k$ and $\langle\varphi_j,\varphi_j\rangle = 1$.

Lemma 1.1. *Suppose $\{\varphi_j\}_{j=0}^n$ is an orthonormal set. Then every $\psi \in \mathfrak{H}$ can be written as*

$$\psi = \psi_\parallel + \psi_\perp, \qquad \psi_\parallel = \sum_{j=0}^n \langle\varphi_j,\psi\rangle\varphi_j, \tag{1.11}$$

where $\psi_\parallel$ and $\psi_\perp$ are orthogonal. Moreover, $\langle\varphi_j,\psi_\perp\rangle = 0$ for all $1 \leq j \leq n$. In particular,

$$\|\psi\|^2 = \sum_{j=0}^n |\langle\varphi_j,\psi\rangle|^2 + \|\psi_\perp\|^2. \tag{1.12}$$

Moreover, every $\hat{\psi}$ in the span of $\{\varphi_j\}_{j=0}^n$ satisfies

$$\|\psi - \hat{\psi}\| \geq \|\psi_\perp\| \tag{1.13}$$

with equality holding if and only if $\hat{\psi} = \psi_\parallel$. In other words, $\psi_\parallel$ is uniquely characterized as the vector in the span of $\{\varphi_j\}_{j=0}^n$ being closest to ψ.

Proof. A straightforward calculation shows $\langle\varphi_j,\psi - \psi_\parallel\rangle = 0$ and hence $\psi_\parallel$ and $\psi_\perp = \psi - \psi_\parallel$ are orthogonal. The formula for the norm follows by applying (1.5) iteratively.

Now, fix a vector

$$\hat{\psi} = \sum_{j=0}^n c_j\varphi_j. \tag{1.14}$$

in the span of $\{\varphi_j\}_{j=0}^n$. Then one computes

$$\begin{aligned} \|\psi - \hat{\psi}\|^2 &= \|\psi_\parallel + \psi_\perp - \hat{\psi}\|^2 = \|\psi_\perp\|^2 + \|\psi_\parallel - \hat{\psi}\|^2 \\ &= \|\psi_\perp\|^2 + \sum_{j=0}^n |c_j - \langle\varphi_j,\psi\rangle|^2 \end{aligned} \tag{1.15}$$

from which the last claim follows. □

From (1.12) we obtain **Bessel's inequality**

$$\sum_{j=0}^n |\langle\varphi_j,\psi\rangle|^2 \leq \|\psi\|^2 \tag{1.16}$$

with equality holding if and only if ψ lies in the span of $\{\varphi_j\}_{j=0}^n$.

Note that a scalar product can be recovered from its norm by virtue of the **polarization identity**

$$\langle\varphi,\psi\rangle = \frac{1}{4}\left(\|\varphi+\psi\|^2 - \|\varphi-\psi\|^2 + \mathrm{i}\|\varphi - \mathrm{i}\psi\|^2 - \mathrm{i}\|\varphi + \mathrm{i}\psi\|^2\right), \tag{1.17}$$

which can be readily verified. However, note that this formula cannot be used to define a scalar product from a norm. It can be shown that this only works if the **parallelogram law**

$$\|\varphi + \psi\|^2 + \|\varphi - \psi\|^2 = 2\|\varphi\|^2 + 2\|\psi\|^2 \tag{1.18}$$

holds (which is easy to verify for inner product spaces).

A bijective operator $U \in \mathfrak{L}(\mathfrak{H}_1, \mathfrak{H}_2)$ is called **unitary** if U preserves scalar products:

$$\langle U\varphi, U\psi\rangle_2 = \langle \varphi, \psi\rangle_1, \qquad \varphi, \psi \in \mathfrak{H}_1. \tag{1.19}$$

By the polarization identity this is the case if and only if U preserves norms: $\|U\psi\|_2 = \|\psi\|_1$ for all $\psi \in \mathfrak{H}_1$. The two Hilbert space $\mathfrak{H}_1$ and $\mathfrak{H}_2$ are called **unitarily equivalent** in this case.

1.2. Orthonormal bases

Of course, since we cannot assume $\mathfrak{H}$ to be a finite dimensional vector space, we need to generalize Lemma 1.1 to arbitrary orthonormal sets $\{\varphi_j\}_{j\in J}$. We start by assuming that J is countable. Then Bessel's inequality (1.16) shows that

$$\sum_{j\in J} |\langle \varphi_j, \psi\rangle|^2 \tag{1.20}$$

converges absolutely. Moreover, for any finite subset $K \subset J$ we have

$$\|\sum_{j\in K} \langle \varphi_j, \psi\rangle \varphi_j\|^2 = \sum_{j\in K} |\langle \varphi_j, \psi\rangle|^2 \tag{1.21}$$

by the Pythagorean theorem and thus $\sum_{j\in J} \langle \varphi_j, \psi\rangle \varphi_j$ is Cauchy if and only if $\sum_{j\in J} |\langle \varphi_j, \psi\rangle|^2$ is. Now let J be arbitrary. Again, Bessel's inequality shows that for any given $\varepsilon > 0$ there are at most finitely many j for which $|\langle \varphi_j, \psi\rangle| \geq \varepsilon$. Hence there are at most countably many j for which $|\langle \varphi_j, \psi\rangle| > 0$. Thus it follows that

$$\sum_{j\in J} |\langle \varphi_j, \psi\rangle|^2 \tag{1.22}$$

is well-defined and so is

$$\sum_{j\in J} \langle \varphi_j, \psi\rangle \varphi_j. \tag{1.23}$$

In particular, by continuity of the scalar product we see that Lemma 1.1 holds for arbitrary orthonormal sets without modifications.

Theorem 1.2. *Suppose $\{\varphi_j\}_{j\in J}$ is an orthonormal set. Then every $\psi \in \mathfrak{H}$ can be written as*

$$\psi = \psi_\parallel + \psi_\perp, \qquad \psi_\parallel = \sum_{j\in J} \langle \varphi_j, \psi\rangle \varphi_j, \tag{1.24}$$

where $\psi_\parallel$ *and* $\psi_\perp$ *are orthogonal. Moreover,* $\langle\varphi_j, \psi_\perp\rangle = 0$ *for all* $j \in J$*. In particular,*

$$\|\psi\|^2 = \sum_{j\in J} |\langle\varphi_j, \psi\rangle|^2 + \|\psi_\perp\|^2. \tag{1.25}$$

Moreover, every $\hat\psi$ *in the span of* $\{\varphi_j\}_{j\in J}$ *satisfies*

$$\|\psi - \hat\psi\| \geq \|\psi_\perp\| \tag{1.26}$$

with equality holding if and only if $\hat\psi = \psi_\parallel$*. In other words,* $\psi_\parallel$ *is uniquely characterized as the vector in the span of* $\{\varphi_j\}_{j\in J}$ *being closest to* ψ*.*

Note that from Bessel's inequality (which of course still holds) it follows that the map $\psi \to \psi_\parallel$ is continuous.

The collection of all orthonormal sets in $\mathfrak{H}$ can be partially ordered by inclusion. Moreover, any linearly ordered chain has an upper bound (the union of all sets in the chain). Hence Zorn's lemma implies the existence of a maximal element, that is, an orthonormal set which is not a proper subset of any other orthonormal set. Such an orthonormal set is called an **orthonormal basis** due to following result:

Theorem 1.3. *For an orthonormal set* $\{\varphi_j\}_{j\in J}$ *the following conditions are equivalent:*
(i) $\{\varphi_j\}_{j\in J}$ *is an orthonormal basis.*
(ii) For every vector $\psi \in \mathfrak{H}$ *we have*

$$\psi = \sum_{j\in J} \langle\varphi_j, \psi\rangle\varphi_j. \tag{1.27}$$

(iii) For every vector $\psi \in \mathfrak{H}$ *we have*

$$\|\psi\|^2 = \sum_{j\in J} |\langle\varphi_j, \psi\rangle|^2. \tag{1.28}$$

(iv) $\langle\varphi_j, \psi\rangle = 0$ *for all* $j \in J$ *implies* $\psi = 0$*.*

Proof. We will use the notation from Theorem 1.2.
$(i) \Rightarrow (ii)$: If $\psi_\perp \neq 0$ than we can normalize $\psi_\perp$ to obtain a unit vector $\tilde\psi_\perp$ which is orthogonal to all vectors φ_j. But then $\{\varphi_j\}_{j\in J} \cup \{\tilde\psi_\perp\}$ would be a larger orthonormal set, contradicting maximality of $\{\varphi_j\}_{j\in J}$.
$(ii) \Rightarrow (iii)$: Follows since (ii) implies $\psi_\perp = 0$.
$(iii) \Rightarrow (iv)$: If $\langle\psi, \varphi_j\rangle = 0$ for all $j \in J$ we conclude $\|\psi\|^2 = 0$ and hence $\psi = 0$.
$(iv) \Rightarrow (i)$: If $\{\varphi_j\}_{j\in J}$ were not maximal, there would be a unit vector φ such that $\{\varphi_j\}_{j\in J} \cup \{\varphi\}$ is larger orthonormal set. But $\langle\varphi_j, \varphi\rangle = 0$ for all $j \in J$ implies $\varphi = 0$ by (iv), a contradiction. □

Since $\psi \to \psi_\|$ is continuous, it suffices to check conditions (ii) and (iii) on a dense set.

Example. The set of functions

$$\varphi_n(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{\mathrm{i}n\,x}, \qquad n \in \mathbb{Z}, \tag{1.29}$$

forms an orthonormal basis for $\mathfrak{H} = L^2(0, 2\pi)$. The corresponding orthogonal expansion is just the ordinary Fourier series. ⋄

A Hilbert space is **separable** if and only if there is a countable orthonormal basis. In fact, if $\mathfrak{H}$ is separable, then there exists a countable total set $\{\psi_j\}_{j=0}^N$. After throwing away some vectors we can assume that ψ_{n+1} cannot be expressed as a linear combinations of the vectors $\psi_0, \dots \psi_n$. Now we can construct an orthonormal set as follows: We begin by normalizing ψ_0

$$\varphi_0 = \frac{\psi_0}{\|\psi_0\|}. \tag{1.30}$$

Next we take ψ_1 and remove the component parallel to φ_0 and normalize again

$$\varphi_1 = \frac{\psi_1 - \langle \varphi_0, \psi_1 \rangle \varphi_0}{\|\psi_1 - \langle \varphi_0, \psi_1 \rangle \varphi_0\|}. \tag{1.31}$$

Proceeding like this we define recursively

$$\varphi_n = \frac{\psi_n - \sum_{j=0}^{n-1} \langle \varphi_j, \psi_n \rangle \varphi_j}{\|\psi_n - \sum_{j=0}^{n-1} \langle \varphi_j, \psi_n \rangle \varphi_j\|}. \tag{1.32}$$

This procedure is known as **Gram-Schmidt orthogonalization**. Hence we obtain an orthonormal set $\{\varphi_j\}_{j=0}^N$ such that $\mathrm{span}\{\varphi_j\}_{j=0}^n = \mathrm{span}\{\psi_j\}_{j=0}^n$ for any finite n and thus also for N. Since $\{\psi_j\}_{j=0}^N$ is total, we infer that $\{\varphi_j\}_{j=0}^N$ is an orthonormal basis.

If fact, if there is one countable basis, then it follows that every other basis is countable as well.

Theorem 1.4. *If $\mathfrak{H}$ is separable, then every orthonormal basis is countable.*

Proof. We know that there is at least one countable orthonormal basis $\{\varphi_j\}_{j\in J}$. Now let $\{\phi_k\}_{k\in K}$ be a second basis and consider the set $K_j = \{k \in K | \langle \phi_k, \varphi_j \rangle \neq 0\}$. Since these are the expansion coefficients of φ_j with respect to $\{\phi_k\}_{k\in K}$, this set is countable. Hence the set $\tilde{K} = \bigcup_{j\in J} K_j$ is countable as well. But $k \in K \backslash \tilde{K}$ implies $\phi_k = 0$ and hence $\tilde{K} = K$. □

We will assume all Hilbert spaces to be separable.

In particular, it can be shown that $L^2(M, d\mu)$ is separable. Moreover, it turns out that, up to unitary equivalence, there is only one (separable) infinite dimensional Hilbert space:

Let $\mathfrak{H}$ be an infinite dimensional Hilbert space and let $\{\varphi_j\}_{j\in\mathbb{N}}$ be any orthogonal basis. Then the map $U : \mathfrak{H} \to \ell^2(\mathbb{N})$, $\psi \mapsto (\langle\varphi_j, \psi\rangle)_{j\in\mathbb{N}}$ is unitary (by Theorem 1.3 (iii)). In particular, any infinite dimensional Hilbert space is unitarily equivalent to $\ell^2(\mathbb{N})$.

1.3. The projection theorem and the Riesz lemma

Let $M \subseteq \mathfrak{H}$ be a subset, then $M^\perp = \{\psi | \langle\varphi, \psi\rangle = 0, \forall\varphi \in M\}$ is called the **orthogonal complement** of M. By continuity of the scalar product it follows that $M^\perp$ is a closed linear subspace and by linearity that $(\overline{\text{span}(M)})^\perp = M^\perp$. For example we have $\mathfrak{H}^\perp = \{0\}$ since any vector in $\mathfrak{H}^\perp$ must be in particular orthogonal to all vectors in some orthonormal basis.

Theorem 1.5 (projection theorem)**.** *Let M be a closed linear subspace of a Hilbert space $\mathfrak{H}$, then every $\psi \in \mathfrak{H}$ can be uniquely written as $\psi = \psi_\parallel + \psi_\perp$ with $\psi_\parallel \in M$ and $\psi_\perp \in M^\perp$. One writes*

$$M \oplus M^\perp = \mathfrak{H} \tag{1.33}$$

in this situation.

Proof. Since M is closed, it is a Hilbert space and has an orthonormal basis $\{\varphi_j\}_{j\in J}$. Hence the result follows from Theorem 1.2. □

In other words, to every $\psi \in \mathfrak{H}$ we can assign a unique vector $\psi_\parallel$ which is the vector in M closest to ψ. The rest $\psi - \psi_\parallel$ lies in $M^\perp$. The operator $P_M\psi = \psi_\parallel$ is called the orthogonal projection corresponding to M. Clearly we have $P_{M^\perp}\psi = \psi - P_M\psi = \psi_\perp$.

Moreover, we see that the vectors in a closed subspace M are precisely those which are orthogonal to all vectors in $M^\perp$, that is, $M^{\perp\perp} = M$. If M is an arbitrary subset we have at least

$$M^{\perp\perp} = \overline{\text{span}(M)}. \tag{1.34}$$

Finally we turn to **linear functionals**, that is, to operators $\ell : \mathfrak{H} \to \mathbb{C}$. By the Cauchy-Schwarz inequality we know that $\ell_\varphi : \psi \mapsto \langle\varphi, \psi\rangle$ is a bounded linear functional (with norm $\|\varphi\|$). In turns out that in a Hilbert space every bounded linear functional can be written in this way.

Theorem 1.6 (Riesz lemma)**.** *Suppose ℓ is a bounded linear functional on a Hilbert space $\mathfrak{H}$. Then there is a vector $\varphi \in \mathfrak{H}$ such that $\ell(\psi) = \langle\varphi, \psi\rangle$ for all $\psi \in \mathfrak{H}$. In other words, a Hilbert space is equivalent to its own dual space $\mathfrak{H}^* = \mathfrak{H}$.*

Proof. If $\ell \equiv 0$ we can choose $\varphi = 0$. Otherwise $\mathrm{Ker}(\ell) = \{\psi | \ell(\psi) = 0\}$ is a proper subspace and we can find a unit vector $\tilde{\varphi} \in \mathrm{Ker}(\ell)^{\perp}$. For every $\psi \in \mathfrak{H}$ we have $\ell(\psi)\tilde{\varphi} - \ell(\tilde{\varphi})\psi \in \mathrm{Ker}(\ell)$ and hence

$$0 = \langle \tilde{\varphi}, \ell(\psi)\tilde{\varphi} - \ell(\tilde{\varphi})\psi \rangle = \ell(\psi) - \ell(\tilde{\varphi})\langle \tilde{\varphi}, \psi \rangle. \tag{1.35}$$

In other words, we can choose $\varphi = \ell(\tilde{\varphi})^* \tilde{\varphi}$. □

The following easy consequence is left as an exercise.

Corollary 1.7. *Suppose B is a bounded skew liner form, that is,*

$$|B(\psi, \varphi)| \le C \|\psi\| \, \|\varphi\|. \tag{1.36}$$

Then there is a unique bounded operator A such that

$$B(\psi, \varphi) = \langle A\psi, \varphi \rangle. \tag{1.37}$$

1.4. Orthogonal sums and tensor products

Given two Hilbert spaces $\mathfrak{H}_1$ and $\mathfrak{H}_2$ we define their **orthogonal sum** $\mathfrak{H}_1 \oplus \mathfrak{H}_2$ to be the set of all pairs $(\psi_1, \psi_2) \in \mathfrak{H}_1 \times \mathfrak{H}_2$ together with the scalar product

$$\langle (\varphi_1, \varphi_2), (\psi_1, \psi_2) \rangle = \langle \varphi_1, \psi_1 \rangle_1 + \langle \varphi_2, \psi_2 \rangle_2. \tag{1.38}$$

It is left as an exercise to verify that $\mathfrak{H}_1 \oplus \mathfrak{H}_2$ is again a Hilbert space. Moreover, $\mathfrak{H}_1$ can be identified with $\{(\psi_1, 0) | \psi_1 \in \mathfrak{H}_1\}$ and we can regard $\mathfrak{H}_1$ as a subspace of $\mathfrak{H}_1 \oplus \mathfrak{H}_2$. Similarly for $\mathfrak{H}_2$. It is also custom to write $\psi_1 + \psi_2$ instead of (ψ_1, ψ_2).

More generally, let $\mathfrak{H}_j$ $j \in \mathbb{N}$, be a countable collection of Hilbert spaces and define

$$\bigoplus_{j=1}^{\infty} \mathfrak{H}_j = \{ \sum_{j=1}^{\infty} \psi_j | \, \psi_j \in \mathfrak{H}_j, \, \sum_{j=1}^{\infty} \|\psi_j\|_j^2 < \infty \}, \tag{1.39}$$

which becomes a Hilbert space with the scalar product

$$\langle \sum_{j=1}^{\infty} \varphi_j, \sum_{j=1}^{\infty} \psi_j \rangle = \sum_{j=1}^{\infty} \langle \varphi_j, \psi_j \rangle_j. \tag{1.40}$$

Suppose $\mathfrak{H}$ and $\tilde{\mathfrak{H}}$ are two Hilbert spaces. Our goal is to construct their tensor product. The elements should be products $\psi \otimes \tilde{\psi}$ of elements $\psi \in \mathfrak{H}$ and $\tilde{\psi} \in \tilde{\mathfrak{H}}$. Hence we start with the set of all finite linear combinations of elements of $\mathfrak{H} \times \tilde{\mathfrak{H}}$

$$\mathcal{F}(\mathfrak{H}, \tilde{\mathfrak{H}}) = \{ \sum_{j=1}^{n} \alpha_j (\psi_j, \tilde{\psi}_j) | (\psi_j, \tilde{\psi}_j) \in \mathfrak{H} \times \tilde{\mathfrak{H}}, \, \alpha_j \in \mathbb{C} \}. \tag{1.41}$$

Since we want $(\psi_1+\psi_2)\otimes\tilde{\psi} = \psi_1\otimes\tilde{\psi}+\psi_2\otimes\tilde{\psi}$, $\psi\otimes(\tilde{\psi}_1+\tilde{\psi}_2) = \psi\otimes\tilde{\psi}_1+\psi\otimes\tilde{\psi}_2$, and $(\alpha\psi)\otimes\tilde{\psi} = \psi\otimes(\alpha\tilde{\psi})$ we consider $\mathcal{F}(\mathfrak{H},\tilde{\mathfrak{H}})/\mathcal{N}(\mathfrak{H},\tilde{\mathfrak{H}})$, where

$$\mathcal{N}(\mathfrak{H},\tilde{\mathfrak{H}}) = \operatorname{span}\{\sum_{j,k=1}^{n} \alpha_j\beta_k(\psi_j,\tilde{\psi}_k) - (\sum_{j=1}^{n}\alpha_j\psi_j, \sum_{k=1}^{n}\beta_k\tilde{\psi}_k)\} \tag{1.42}$$

and write $\psi\otimes\tilde{\psi}$ for the equivalence class of $(\psi,\tilde{\psi})$.

Next we define

$$\langle\psi\otimes\tilde{\psi}, \phi\otimes\tilde{\phi}\rangle = \langle\psi,\phi\rangle\langle\tilde{\psi},\tilde{\phi}\rangle \tag{1.43}$$

which extends to a skew linear form on $\mathcal{F}(\mathfrak{H},\tilde{\mathfrak{H}})/\mathcal{N}(\mathfrak{H},\tilde{\mathfrak{H}})$. To show that we obtain a scalar product, we need to ensure positivity. Let $\psi = \sum_i \alpha_i\psi_i\otimes\tilde{\psi}_i \neq 0$ and pick orthonormal bases φ_j, $\tilde{\varphi}_k$ for $\operatorname{span}\{\psi_i\}$, $\operatorname{span}\{\tilde{\psi}_i\}$, respectively. Then

$$\psi = \sum_{j,k}\alpha_{jk}\varphi_j\otimes\tilde{\varphi}_k, \quad \alpha_{jk} = \sum_i \alpha_i\langle\varphi_j,\psi_i\rangle\langle\tilde{\varphi}_k,\tilde{\psi}_i\rangle \tag{1.44}$$

and we compute

$$\langle\psi,\psi\rangle = \sum_{j,k}|\alpha_{jk}|^2 > 0. \tag{1.45}$$

The completion of $\mathcal{F}(\mathfrak{H},\tilde{\mathfrak{H}})/\mathcal{N}(\mathfrak{H},\tilde{\mathfrak{H}})$ with respect to the induced norm is called the **tensor product** $\mathfrak{H}\otimes\tilde{\mathfrak{H}}$ of $\mathfrak{H}$ and $\tilde{\mathfrak{H}}$.

Lemma 1.8. *If φ_j, $\tilde{\varphi}_k$ are orthonormal bases for $\mathfrak{H}$, $\tilde{\mathfrak{H}}$, respectively, then $\varphi_j\otimes\tilde{\varphi}_k$ is an orthonormal basis for $\mathfrak{H}\otimes\tilde{\mathfrak{H}}$.*

Proof. That $\varphi_j\otimes\tilde{\varphi}_k$ is an orthonormal set is immediate from (1.43). Moreover, since $\operatorname{span}\{\varphi_j\}$, $\operatorname{span}\{\tilde{\varphi}_k\}$ is dense in $\mathfrak{H}$, $\tilde{\mathfrak{H}}$, respectively, it is easy to see that $\varphi_j\otimes\tilde{\varphi}_k$ is dense in $\mathcal{F}(\mathfrak{H},\tilde{\mathfrak{H}})/\mathcal{N}(\mathfrak{H},\tilde{\mathfrak{H}})$. But the latter is dense in $\mathfrak{H}\otimes\tilde{\mathfrak{H}}$. □

Example. We have $\mathfrak{H}\otimes\mathbb{C}^n = \mathfrak{H}^n$. ◇

Example. Let $(M,d\mu)$ and $(\tilde{M},d\tilde{\mu})$ be two measure spaces. Then we have $L^2(M,d\mu)\otimes L^2(\tilde{M},d\tilde{\mu}) = L^2(M\times\tilde{M},d\mu\times d\tilde{\mu})$.

Clearly we have $L^2(M,d\mu)\otimes L^2(\tilde{M},d\tilde{\mu}) \subseteq L^2(M\times\tilde{M},d\mu\times d\tilde{\mu})$. Now take an orthonormal basis $\varphi_j\otimes\tilde{\varphi}_k$ for $L^2(M,d\mu)\otimes L^2(\tilde{M},d\tilde{\mu})$ as in our previous lemma. Then

$$\int_M\int_{\tilde{M}}(\varphi_j(x)\tilde{\varphi}_k(y))^* f(x,y)d\mu(x)d\tilde{\mu}(y) = 0 \tag{1.46}$$

implies

$$\int_M \varphi_j(x)^* f_k(x)d\mu(x) = 0, \qquad f_k(x) = \int_{\tilde{M}}\tilde{\varphi}_k(y)^* f(x,y)d\tilde{\mu}(y) \tag{1.47}$$

and hence $f_k(x) = 0$ μ-a.e. x. But this implies $f(x, y) = 0$ for μ-a.e. x and $\tilde{\mu}$-a.e. y and thus $f = 0$. Hence $\varphi_j \otimes \tilde{\varphi}_k$ is a basis for $L^2(M \times \tilde{M}, d\mu \times d\tilde{\mu})$ and equality follows. ⋄

It is straightforward to extend the tensor product to any finite number of Hilbert spaces. We even note

$$(\bigoplus_{j=1}^{\infty} \mathfrak{H}_j) \otimes \mathfrak{H} = \bigoplus_{j=1}^{\infty} (\mathfrak{H}_j \otimes \mathfrak{H}), \tag{1.48}$$

where equality has to be understood in the sense, that both spaces are unitarily equivalent by virtue of the identification

$$(\sum_{j=1}^{\infty} \psi_j) \otimes \psi = \sum_{j=1}^{\infty} \psi_j \otimes \psi. \tag{1.49}$$

Chapter 2

Self-adjointness and spectrum

2.1. Some quantum mechanics

In quantum mechanics, a single particle living in $\mathbb{R}^3$ is described by a complex-valued function (the **wave function**)

$$\psi(x,t), \qquad (x,t) \in \mathbb{R}^3 \times \mathbb{R}, \tag{2.1}$$

where x corresponds to a point in space and t corresponds to time. The quantity $\rho_t(x) = |\psi(x,t)|^2$ is interpreted as the **probability density** of the particle at the time t. In particular, ψ must be normalized according to

$$\int_{\mathbb{R}^3} |\psi(x,t)|^2 d^3x = 1, \qquad t \in \mathbb{R}. \tag{2.2}$$

The location x of the particle is a quantity which can be observed (i.e., measured) and is hence called **observable**. Due to our probabilistic interpretation it is also a random variable whose **expectation** is given by

$$\mathbb{E}_\psi(x) = \int_{\mathbb{R}^3} x|\psi(x,t)|^2 d^3x. \tag{2.3}$$

In a real life setting, it will not be possible to measure x directly and one will only be able to measure certain functions of x. For example, it is possible to check whether the particle is inside a certain area Ω of space (e.g., inside a detector). The corresponding observable is the characteristic function $\chi_\Omega(x)$ of this set. In particular, the number

$$\mathbb{E}_\psi(\chi_\Omega) = \int_{\mathbb{R}^3} \chi_\Omega(x)|\psi(x,t)|^2 d^3x = \int_\Omega |\psi(x,t)|^2 d^3x \tag{2.4}$$

corresponds to the probability of finding the particle inside $\Omega \subseteq \mathbb{R}^3$. An important point to observe is that, in contradistinction to classical mechanics, the particle is no longer localized at a certain point. In particular, the **mean-square deviation** $\Delta_\psi(x)^2 = \mathbb{E}_\psi(x^2) - \mathbb{E}_\psi(x)^2$ is always nonzero.

In general, the **configuration space** (or **phase space**) of a quantum system is a (complex) Hilbert space $\mathfrak{H}$ and the possible states of this system are represented by the elements ψ having norm one, $\|\psi\| = 1$.

An observable a corresponds to a linear operator A in this Hilbert space and its expectation, if the system is in the state ψ, is given by the real number

$$\mathbb{E}_\psi(A) = \langle \psi, A\psi \rangle, \tag{2.5}$$

where $\langle ., .. \rangle$ denotes the scalar product of $\mathfrak{H}$. From a physical point of view, (2.5) should make sense for any $\psi \in \mathfrak{H}$. However, this is not in the cards as our simple example of one particle already shows. In fact, the reader is invited to find a square integrable function $\psi(x)$ for which $x\psi(x)$ is no longer square integrable. The deeper reason behind this nuisance is that $\mathbb{E}_\psi(x)$ can attain arbitrary large values if the particle is not confined to a finite domain, which renders the corresponding operator unbounded. But unbounded operators cannot be defined on the entire Hilbert space in a natural way by the closed graph theorem (Theorem 2.7 below) .

Hence, A will only be defined on a subset $\mathfrak{D}(A) \subseteq \mathfrak{H}$ called the **domain** of A. Since we want A to be defined for *most* states, we require $\mathfrak{D}(A)$ to be dense.

However, it should be noted that there is no general prescription how to find the operator corresponding to a given observable.

Now let us turn to the time evolution of such a quantum mechanical system. Given an initial state $\psi(0)$ of the system, there should be a unique $\psi(t)$ representing the state of the system at time $t \in \mathbb{R}$. We will write

$$\psi(t) = U(t)\psi(0). \tag{2.6}$$

Moreover, it follows from physical experiments, that **superposition of states** holds, that is, $U(t)(\alpha_1\psi_1(0) + \alpha_2\psi_2(0)) = \alpha_1\psi_1(t) + \alpha_2\psi_2(t)$ ($|\alpha_1|^2 + |\alpha_2|^2 = 1$). In other words, $U(t)$ should be a linear operator. Moreover, since $\psi(t)$ is a state (i.e., $\|\psi(t)\| = 1$), we have

$$\|U(t)\psi\| = \|\psi\|. \tag{2.7}$$

Such operators are called **unitary**. Next, since we have assumed uniqueness of solutions to the initial value problem, we must have

$$U(0) = \mathbb{I}, \qquad U(t+s) = U(t)U(s). \tag{2.8}$$

A family of unitary operators $U(t)$ having this property is called a **one-parameter unitary group**. In addition, it is natural to assume that this group is **strongly continuous**

$$\lim_{t\to t_0} U(t)\psi = U(t_0)\psi, \qquad \psi \in \mathfrak{H}. \tag{2.9}$$

Each such group has an **infinitesimal generator** defined by

$$H\psi = \lim_{t\to 0} \frac{\mathrm{i}}{t}(U(t)\psi - \psi), \quad \mathfrak{D}(H) = \{\psi \in \mathfrak{H} | \lim_{t\to 0} \frac{1}{t}(U(t)\psi - \psi) \text{ exists}\}. \tag{2.10}$$

This operator is called the **Hamiltonian** and corresponds to the energy of the system. If $\psi(0) \in \mathfrak{D}(H)$, then $\psi(t)$ is a solution of the **Schrödinger equation** (in suitable units)

$$\mathrm{i}\frac{d}{dt}\psi(t) = H\psi(t). \tag{2.11}$$

This equation will be the main subject of our course.

In summary, we have the following **axioms of quantum mechanics**.

Axiom 1. The configuration space of a quantum system is a complex separable Hilbert space $\mathfrak{H}$ and the possible states of this system are represented by the elements of $\mathfrak{H}$ which have norm one.

Axiom 2. Each observable a corresponds to a linear operator A defined maximally on a dense subset $\mathfrak{D}(A)$. Moreover, the operator corresponding to a polynomial $P_n(a) = \sum_{j=0}^n \alpha_j a^j$, $\alpha_j \in \mathbb{R}$, is $P_n(A) = \sum_{j=0}^n \alpha_j A^j$, $\mathfrak{D}(P_n(A)) = \mathfrak{D}(A^n) = \{\psi \in \mathfrak{D}(A) | A\psi \in \mathfrak{D}(A^{n-1})\}$ ($A^0 = \mathbb{I}$).

Axiom 3. The expectation value for a measurement of a, when the system is in the state $\psi \in \mathfrak{D}(A)$, is given by (2.5), which must be real for all $\psi \in \mathfrak{D}(A)$.

Axiom 4. The time evolution is given by a strongly continuous one-parameter unitary group $U(t)$. The generator of this group corresponds to the energy of the system.

In the following sections we will try to draw some mathematical consequences from these assumptions:

First we will see that Axiom 2 and 3 imply that observables correspond to self-adjoint operators. Hence these operators play a central role in quantum mechanics and we will derive some of their basic properties. Another crucial role is played by the (closure of the) set of all possible expectation values for the measurement of a, which will be identified as the spectrum $\sigma(A)$ of the corresponding operator A.

The problem of defining functions of an observable will lead us to the spectral theorem (in the next chapter), which generalizes the diagonalization of symmetric matrices.

Axiom 4 will be the topic of Chapter 5.

2.2. Self-adjoint operators

Let $\mathfrak{H}$ be a (complex separable) Hilbert space. A **linear operator** is a linear mapping

$$A : \mathfrak{D}(A) \to \mathfrak{H}, \tag{2.12}$$

where $\mathfrak{D}(A)$ is a linear subspace of $\mathfrak{H}$, called the **domain** of A. It is called **bounded** if the operator norm

$$\|A\| = \sup_{\|\psi\|=1} \|A\psi\| = \sup_{\|\varphi\|=\|\psi\|=1} |\langle\psi, A\varphi\rangle| \tag{2.13}$$

is finite. The second equality follows since equality in $|\langle\psi, A\varphi\rangle| \le \|\psi\|\,\|A\varphi\|$ is attained when $A\varphi = z\psi$ for some $z \in \mathbb{C}$. If A is bounded it is no restriction to assume $\mathfrak{D}(A) = \mathfrak{H}$ and we will always do so. The Banach space of all bounded linear operators is denoted by $\mathfrak{L}(\mathfrak{H})$.

The expression $\langle\psi, A\psi\rangle$ encountered in the previous section is called the **quadratic form**

$$q_A(\psi) = \langle\psi, A\psi\rangle, \qquad \psi \in \mathfrak{D}(A), \tag{2.14}$$

associated to A. An operator can be reconstructed from its quadratic form via the **polarization identity**

$$\langle\varphi, A\psi\rangle = \frac{1}{4}\left(q_A(\varphi+\psi) - q_A(\varphi-\psi) + \mathrm{i}q_A(\varphi-\mathrm{i}\psi) - \mathrm{i}q_A(\varphi+\mathrm{i}\psi)\right). \tag{2.15}$$

A densely defined linear operator A is called **symmetric** (or **Hermitian**) if

$$\langle\varphi, A\psi\rangle = \langle A\varphi, \psi\rangle, \qquad \psi, \varphi \in \mathfrak{D}(A). \tag{2.16}$$

The justification for this definition is provided by the following

Lemma 2.1. *A densely defined operator A is symmetric if and only if the corresponding quadratic form is real valued.*

Proof. Clearly (2.16) implies that $\mathrm{Im}(q_A(\psi)) = 0$. Conversely, taking the imaginary part of the identity

$$q_A(\psi + \mathrm{i}\varphi) = q_A(\psi) + q_A(\varphi) + \mathrm{i}(\langle\psi, A\varphi\rangle - \langle\varphi, A\psi\rangle) \tag{2.17}$$

shows $\mathrm{Re}\langle A\varphi, \psi\rangle = \mathrm{Re}\langle\varphi, A\psi\rangle$. Replacing φ by $\mathrm{i}\varphi$ in this last equation shows $\mathrm{Im}\langle A\varphi, \psi\rangle = \mathrm{Im}\langle\varphi, A\psi\rangle$ and finishes the proof. □

In other words, a densely defined operator A is symmetric if and only if

$$\langle\psi, A\psi\rangle = \langle A\psi, \psi\rangle, \qquad \psi \in \mathfrak{D}(A). \tag{2.18}$$

This already narrows the class of admissible operators to the class of symmetric operators by Axiom 3. Next, let us tackle the issue of the correct domain.

By Axiom 2, A should be defined maximally, that is, if $\tilde{A}$ is another symmetric operator such that $A \subseteq \tilde{A}$, then $A = \tilde{A}$. Here we write $A \subseteq \tilde{A}$ if $\mathfrak{D}(A) \subseteq \mathfrak{D}(\tilde{A})$ and $A\psi = \tilde{A}\psi$ for all $\psi \in \mathfrak{D}(A)$. In addition, we write $A = \tilde{A}$ if both $\tilde{A} \subseteq A$ and $A \subseteq \tilde{A}$ hold.

The **adjoint operator** A^* of a densely defined linear operator A is defined by

$$\begin{array}{rcl} \mathfrak{D}(A^*) & = & \{\psi \in \mathfrak{H} | \exists \tilde{\psi} \in \mathfrak{H} : \langle \psi, A\varphi \rangle = \langle \tilde{\psi}, \varphi \rangle, \forall \varphi \in \mathfrak{D}(A)\} \\ A^*\psi & = & \tilde{\psi} \end{array} . \qquad (2.19)$$

The requirement that $\mathfrak{D}(A)$ is dense implies that A^* is well-defined. However, note that $\mathfrak{D}(A^*)$ might not be dense in general. In fact, it might contain no vectors other than 0.

For later use, note that

$$\mathrm{Ker}(A^*) = \mathrm{Ran}(A)^{\perp}. \qquad (2.20)$$

For symmetric operators we clearly have $A \subseteq A^*$. If in addition, $A = A^*$ holds, then A is called **self-adjoint**. Our goal is to show that observables correspond to self-adjoint operators. This is for example true in the case of the position operator x, which is a special case of a multiplication operator.

Example. (Multiplication operator) Consider the multiplication operator

$$(Af)(x) = A(x)f(x), \quad \mathfrak{D}(A) = \{f \in L^2(\mathbb{R}^n, d\mu) \,|\, Af \in L^2(\mathbb{R}^n, d\mu)\}, \qquad (2.21)$$

given by multiplication with the measurable function $A : \mathbb{R}^n \to \mathbb{C}$. First of all note that $\mathfrak{D}(A)$ is dense. In fact, consider $\Omega_n = \{x \in \mathbb{R}^n \,|\, |A(x)| \le n\}$. Then, for every $f \in L^2(\mathbb{R}^n, d\mu)$ the function $f_n = \chi_{\Omega_n} f \in \mathfrak{D}(A)$ converges to f as $n \to \infty$.

Next, let us compute the adjoint of A. Performing a formal computation we have for $h, f \in \mathfrak{D}(A)$ that

$$\langle h, Af \rangle = \int h(x)^* A(x) f(x) d\mu(x) = \int (A(x)^* h(x))^* f(x) d\mu(x) = \langle \tilde{A}h, f \rangle, \qquad (2.22)$$

where $\tilde{A}$ is multiplication by $A(x)^*$,

$$(\tilde{A}f)(x) = A(x)^* f(x), \quad \mathfrak{D}(\tilde{A}) = \{f \in L^2(\mathbb{R}^n, d\mu) \,|\, \tilde{A}f \in L^2(\mathbb{R}^n, d\mu)\}. \qquad (2.23)$$

Note $\mathfrak{D}(\tilde{A}) = \mathfrak{D}(A)$. At first sight this seems to show that the adjoint of A is $\tilde{A}$. But for our calculation we had to assume $h \in \mathfrak{D}(A)$ and there might be some functions in $\mathfrak{D}(A^*)$ which do not satisfy this requirement! In particular, our calculation only shows $\tilde{A} \subseteq A^*$. To show that equality holds, we need to work a little harder:

If $h \in \mathfrak{D}(A^*)$ we must have

$$\int h(x)^* A(x) f(x) d\mu(x) = \int g(x)^* f(x) d\mu(x), \qquad f \in \mathfrak{D}(A), \tag{2.24}$$

and thus

$$\int (h(x)A(x)^* - g(x))^* f(x) d\mu(x) = 0, \qquad f \in \mathfrak{D}(A). \tag{2.25}$$

In particular,

$$\int \chi_{\Omega_n}(x)(h(x)A(x)^* - g(x))^* f(x) d\mu(x) = 0, \qquad f \in L^2(\mathbb{R}^n, d\mu), \tag{2.26}$$

which shows that $\chi_{\Omega_n}(h(x)A(x)^* - g(x))^* f(x) \in L^2(\mathbb{R}^n, d\mu)$ vanishes. Since n is arbitrary, we even have $h(x)A(x)^* = g(x) \in L^2(\mathbb{R}^n, d\mu)$ and thus A^* is multiplication by $A(x)^*$ and $\mathfrak{D}(A^*) = \mathfrak{D}(A)$.

In particular, A is self-adjoint if A is real-valued. In the general case we have at least $\|Af\| = \|A^* f\|$ for all $f \in \mathfrak{D}(A) = \mathfrak{D}(A^*)$. Such operators are called normal. ⋄

Now note that

$$A \subseteq B \quad \Rightarrow \quad B^* \subseteq A^*, \tag{2.27}$$

that is, increasing the domain of A implies decreasing the domain of A^*. Thus there is no point in trying to extend the domain of a self-adjoint operator further. In fact, if A is self-adjoint and B is a symmetric extension, we infer $A \subseteq B \subseteq B^* \subseteq A^* = A$ implying $A = B$.

Corollary 2.2. *Self-adjoint operators are maximal, that is, they do not have any symmetric extensions.*

Furthermore, if A^* is densely defined (which is the case if A is symmetric) we can consider A^{**}. From the definition (2.19) it is clear that $A \subseteq A^{**}$ and thus A^{**} is an extension of A. This extension is closely related to extending a linear subspace M via $M^{\perp\perp} = \overline{M}$ (as we will see a bit later) and thus is called the **closure** $\overline{A} = A^{**}$ of A.

If A is symmetric we have $A \subseteq A^*$ and hence $\overline{A} = A^{**} \subseteq A^*$, that is, $\overline{A}$ lies between A and A^*. Moreover, $\langle \psi, A^* \varphi \rangle = \langle \overline{A} \psi, \varphi \rangle$ for all $\psi \in \mathfrak{D}(\overline{A})$, $\varphi \in \mathfrak{D}(A^*)$ implies that $\overline{A}$ is symmetric since $A^* \varphi = \overline{A} \varphi$ for $\varphi \in \mathfrak{D}(\overline{A})$.

Example. (Differential operator) Take $\mathfrak{H} = L^2(0, 2\pi)$.

(i). Consider the operator

$$A_0 f = -\mathrm{i} \frac{d}{dx} f, \quad \mathfrak{D}(A_0) = \{ f \in C^1[0, 2\pi] \,|\, f(0) = f(2\pi) = 0 \}. \tag{2.28}$$

That A_0 is symmetric can be shown by a simple integration by parts. However, this will also follow once we have computed A_0^*. If $g \in \mathfrak{D}(A_0^*)$ we must have

$$\int_0^{2\pi} \mathrm{i} f'(x)^* g(x) dx = \int_0^{2\pi} \mathrm{i} f(x)^* \tilde{g}(x) dx \tag{2.29}$$

for some $\tilde{g} \in L^2(0, 2\pi)$. Integration by parts shows

$$\int_0^{2\pi} \mathrm{i} f'(x)^* \left(g(x) + \mathrm{i} \int_0^x \tilde{g}(t) dt \right) dx = 0 \tag{2.30}$$

and hence $g(x) + \mathrm{i} \int_0^x \tilde{g}(t) dt \in \{f' | f \in \mathfrak{D}(A_0)\}^\perp$. But $\{f' | f \in \mathfrak{D}(A_0)\} = \{h \in \mathfrak{H} | \int_0^{2\pi} h(t) dt = 0\} = \{1\}^\perp$ implying $g(x) = g(0) - \mathrm{i} \int_0^x \tilde{g}(t) dt$ since $\{1\}^{\perp\perp} = \mathrm{span}\{1\}$. Thus $g \in AC[0, 2\pi]$, where

$$AC[a, b] = \{f \in C[a, b] | f(x) = f(a) + \int_a^x g(t) dt,\ g \in L^1(a, b)\} \tag{2.31}$$

denotes the set of all **absolutely continuous functions**. In summary, $g \in AC[0, 2\pi] \subseteq \mathfrak{D}(A_0^*)$ and $A_0^* g = \tilde{g} = -\mathrm{i} g'$. Conversely, for every $g \in AC[0, 2\pi]$ (2.29) holds with $\tilde{g} = -\mathrm{i} g'$ and we conclude

$$A_0^* f = -\mathrm{i} \frac{d}{dx} f, \quad \mathfrak{D}(A_0^*) = AC[0, 2\pi]. \tag{2.32}$$

In particular, A is symmetric but not self-adjoint. Since $A^{**} \subseteq A^*$ we compute

$$0 = \langle g, \overline{A_0} f \rangle - \langle A_0^* g, f \rangle = \mathrm{i}(f(0) g(0)^* - f(2\pi) g(2\pi)^*) \tag{2.33}$$

and since the boundary values of $g \in \mathfrak{D}(A_0^*)$ can be prescribed arbitrary, we must have $f(0) = f(2\pi) = 0$. Thus

$$\overline{A_0} f = -\mathrm{i} \frac{d}{dx} f, \quad \mathfrak{D}(\overline{A_0}) = \{f \in AC[0, 2\pi] \,|\, f(0) = f(2\pi) = 0\}. \tag{2.34}$$

(ii). Now let us take

$$Af = -\mathrm{i} \frac{d}{dx} f, \quad \mathfrak{D}(A) = \{f \in C^1[0, 2\pi] \,|\, f(0) = f(2\pi)\}. \tag{2.35}$$

which is clearly an extension of A_0. Thus $A^* \subseteq A_0^*$ and we compute

$$0 = \langle g, Af \rangle - \langle A^* g, f \rangle = \mathrm{i} f(0)(g(0)^* - g(2\pi)^*). \tag{2.36}$$

Since this must hold for all $f \in \mathfrak{D}(A)$ we conclude $g(0) = g(2\pi)$ and

$$A^* f = -\mathrm{i} \frac{d}{dx} f, \quad \mathfrak{D}(A^*) = \{f \in AC[0, 2\pi] \,|\, f(0) = f(2\pi)\}. \tag{2.37}$$

Similarly, as before, $\overline{A} = A^*$ and thus $\overline{A}$ is self-adjoint. ⋄

One might suspect that there is no big difference between the two symmetric operators A_0 and A from the previous example, since they coincide on a dense set of vectors. However, the converse is true: For example, the first operator A_0 has no eigenvectors at all (i.e., solutions of the equation

$A_0\psi = z\psi$, $z \in \mathbb{C}$) whereas the second one has an orthonormal basis of eigenvectors!

Example. Compute the eigenvectors of A_0 and A from the previous example.

(i). By definition an eigenvector is a (nonzero) solution of $A_0\psi = z\psi$, $z \in \mathbb{C}$, that is, a solution of the ordinary differential equation

$$\psi'(x) = z\psi(x) \tag{2.38}$$

satisfying the boundary conditions $\psi(0) = \psi(2\pi) = 0$ (since we must have $\psi \in \mathfrak{D}(A_0)$. The general solution of the differential equation is $\psi(x) = \psi(0)\mathrm{e}^{zx}$ and the boundary conditions imply $\psi(x) = 0$. Hence there are no eigenvectors.

(ii). Now we look for solutions of $A\psi = z\psi$, that is the same differential equation as before, but now subject to the boundary condition $\psi(0) = \psi(2\pi)$. Again the general solution is $\psi(x) = \psi(0)\mathrm{e}^{zx}$ and the boundary condition requires $\psi(0) = \psi(0)\mathrm{e}^{2\pi z}$. Thus there are two possibilities. Either $\psi(0) = 0$ (which is of no use for us) or $z \in \mathbb{Z}$. In particular, we see that all eigenfunctions are given by

$$\psi_n(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{nx}, \qquad n \in \mathbb{Z}, \tag{2.39}$$

which are well-known to form an orthonormal basis. ⋄

We will see a bit later that this is a consequence of self-adjointness of $\overline{A}$. Hence it will be important to know whether a given operator is self-adjoint or not. Our example shows that symmetry is easy to check (in case of differential operators it usually boils down to integration by parts), but computing the adjoint of an operator is a nontrivial job even in simple situations. However, we will learn soon that self-adjointness is a much stronger property than symmetry justifying the additional effort needed to prove it.

On the other hand, if a given symmetric operator A turns out not to be self-adjoint, this raises the question of self-adjoint extensions. Two cases need to be distinguished. If $\overline{A}$ is self-adjoint, then there is only one self-adjoint extension (if B is another one, we have $\overline{A} \subseteq B$ and hence $\overline{A} = B$ by Corollary 2.2). In this case A is called **essentially self-adjoint** and $\mathfrak{D}(A)$ is called a **core** for $\overline{A}$. Otherwise there might be more than one self-adjoint extension or none at all. This situation is more delicate and will be investigated in Section 2.5.

Since we have seen that computing A^* is not always easy, a criterion for self-adjointness not involving A^* will be useful.

Lemma 2.3. *Let A be symmetric such that* $\mathrm{Ran}(A+z) = \mathrm{Ran}(A+z^*) = \mathfrak{H}$ *for one $z \in \mathbb{C}$. Then A is self-adjoint.*

Proof. Let $\psi \in \mathfrak{D}(A^*)$ and $A^*\psi = \tilde{\psi}$. Since $\mathrm{Ran}(A + z^*) = \mathfrak{H}$, there is a $\vartheta \in \mathfrak{D}(A)$ such that $(A + z^*)\vartheta = \tilde{\psi} + z^*\psi$. Now we compute

$$\langle \psi, (A+z)\varphi \rangle = \langle \tilde{\psi} + z^*\psi, \varphi \rangle = \langle (A+z^*)\vartheta, \varphi \rangle = \langle \vartheta, (A+z)\varphi \rangle, \quad \varphi \in \mathfrak{D}(A), \tag{2.40}$$

and hence $\psi = \vartheta \in \mathfrak{D}(A)$ since $\mathrm{Ran}(A + z) = \mathfrak{H}$. □

To proceed further, we will need more information on the closure of an operator. We will use a different approach which avoids the use of the adjoint operator. We will establish equivalence with our original definition in Lemma 2.4.

The simplest way of extending an operator A is to take the closure of its **graph** $\Gamma(A) = \{(\psi, A\psi) | \psi \in \mathfrak{D}(A)\} \subset \mathfrak{H}^2$. That is, if $(\psi_n, A\psi_n) \to (\psi, \tilde{\psi})$ we might try to define $A\psi = \tilde{\psi}$. For $A\psi$ to be well-defined, we need that $(\psi_n, A\psi_n) \to (0, \tilde{\psi})$ implies $\tilde{\psi} = 0$. In this case A is called **closable** and the unique operator $\overline{A}$ which satisfies $\Gamma(\overline{A}) = \overline{\Gamma(A)}$ is called the **closure** of A. Clearly, A is called **closed** if $\overline{A} = A$, which is the case if and only if the graph of A is closed. A bounded operator is closed if and only if its domain is closed.

Example. Let us compute the closure of the operator A_0 from the previous example without the use of the adjoint operator. Let $f \in \mathfrak{D}(\overline{A_0})$ and let $f_n \in \mathfrak{D}(A_0)$ be a sequence such that $f_n \to f$, $A_0 f_n \to -\mathrm{i}g$. Then $f_n' \to g$ and hence $f(x) = \int_0^x g(t)dt$. Thus $f \in AC[0, 2\pi]$ and $f(0) = 0$. Moreover $f(2\pi) = \lim_{n\to 0} \int_0^{2\pi} f_n'(t)dt = 0$. Conversely, any such f can be approximated by functions in $\mathfrak{D}(A_0)$ (show this). ⋄

Next, let us collect a few important results.

Lemma 2.4. *Suppose A is a densely defined operator.*

(1) *A^* is closed.*

(2) *A is closable if and only if $\mathfrak{D}(A^*)$ is dense and $\overline{A} = A^{**}$ respectively $(\overline{A})^* = A^*$ in this case.*

(3) *If A is injective and the $\mathrm{Ran}(A)$ is dense, then $(A^*)^{-1} = (A^{-1})^*$. If A is closable and $\overline{A}$ is injective, then $\overline{A}^{-1} = \overline{A^{-1}}$.*

Proof. Let us consider the following two unitary operators from $\mathfrak{H}^2$ to itself

$$U(\varphi, \psi) = (\psi, -\varphi), \qquad V(\varphi, \psi) = (\psi, \varphi). \tag{2.41}$$

(1). From

$$\begin{aligned}\Gamma(A^*) &= \{(\varphi,\tilde\varphi)|\langle\varphi, A\psi\rangle = \langle\tilde\varphi,\psi\rangle\ \forall\varphi\in\mathfrak{D}(A^*)\}\\ &= \{(\varphi,\tilde\varphi)|\langle(-\tilde\varphi,\varphi),(\psi,\tilde\psi)\rangle = 0\ \forall(\psi,\tilde\psi)\in\Gamma(A)\}\\ &= U(\Gamma(A)^\perp) = (U\Gamma(A))^\perp \end{aligned}\tag{2.42}$$

we conclude that A^* is closed.

(2). From

$$\begin{aligned}\overline{\Gamma(A)} &= \Gamma(A)^{\perp\perp} = (U\Gamma(A^*))^\perp\\ &= \{(\psi,\tilde\psi)|\,\langle\psi, A^*\varphi\rangle - \langle\tilde\psi,\varphi\rangle = 0, \forall\varphi\in\mathfrak{D}(A^*)\}\end{aligned}\tag{2.43}$$

we see that $(0,\tilde\psi)\in\overline{\Gamma(A)}$ if and only if $\tilde\psi\in\mathfrak{D}(A^*)^\perp$. Hence A is closable if and only if $\mathfrak{D}(A^*)$ is dense. In this case, equation (2.42) also shows $\overline{A}^* = A^*$. Moreover, replacing A by A^* in (2.42) and comparing with the last formula shows $A^{**} = \overline{A}$.

(3). Next note that (provided A is injective)

$$\Gamma(A^{-1}) = V\Gamma(A).\tag{2.44}$$

Hence if $\mathrm{Ran}(A)$ is dense, then $\mathrm{Ker}(A^*) = \mathrm{Ran}(A)^\perp = \{0\}$ and

$$\Gamma((A^*)^{-1}) = V\Gamma(A^*) = VU\Gamma(A)^\perp = UV\Gamma(A)^\perp = U(V\Gamma(A))^\perp\tag{2.45}$$

shows that $(A^*)^{-1} = (A^{-1})^*$. Similarly, if A is closable and $\overline{A}$ is injective, then $\overline{A}^{-1} = \overline{A^{-1}}$ by

$$\Gamma(\overline{A}^{-1}) = V\Gamma(\overline{A}) = \overline{V\Gamma(A)} = \Gamma(\overline{A^{-1}}).\tag{2.46}$$

□

Furthermore, if $A\in\mathfrak{L}(\mathfrak{H})$ we clearly have $\mathfrak{D}(A^*) = \mathfrak{H}$ and hence

$$\|A^*\| = \sup_{\|\varphi\|=\|\psi\|=1} |\langle\psi, A^*\varphi\rangle| = \sup_{\|\varphi\|=\|\psi\|=1} |\langle A\psi,\varphi\rangle| = \|A\|.\tag{2.47}$$

In particular, since $\overline{A} = A^{**}$ we obtain

Theorem 2.5. *We have $\overline{A}\in\mathfrak{L}(\mathfrak{H})$ if and only if $A^*\in\mathfrak{L}(\mathfrak{H})$ Moreover, $\|A^*\| = \|A\|$ in this case.*

Now we can also generalize Lemma 2.3 to the case of essential self-adjoint operators.

Lemma 2.6. *A symmetric operator A is essentially self-adjoint if and only if one of the following conditions holds for one $z\in\mathbb{C}\backslash\mathbb{R}$.*

- $\overline{\mathrm{Ran}(A+z)} = \overline{\mathrm{Ran}(A+z^*)} = \mathfrak{H}$.
- $\mathrm{Ker}(A^*+z) = \mathrm{Ker}(A^*+z^*) = \{0\}$.

If A is non-negative, that is $\langle\psi, A\psi\rangle \ge 0$ for all $\psi \in \mathfrak{D}(A)$, we can also admit $z \in (-\infty, 0)$.

Proof. As noted earlier $\mathrm{Ker}(A^*) = \mathrm{Ran}(A)^\perp$, and hence the two conditions are equivalent. By taking the closure of A it is no restriction to assume that A is closed. Let $z = x + \mathrm{i}y$. From

$$\|(A-z)\psi\|^2 = \|(A-x)\psi - \mathrm{i}y\psi\|^2 = \|(A-x)\psi\|^2 + y^2\|\psi\|^2 \ge y^2\|\psi\|^2, \quad (2.48)$$

we infer that $\mathrm{Ker}(A-z) = \{0\}$ and hence $(A-z)^{-1}$ exists. Moreover, setting $\psi = (A-z)^{-1}\varphi$ $(y \ne 0)$ shows $\|(A-z)^{-1}\| \le |y|^{-1}$. Hence $(A-z)^{-1}$ is bounded and closed. Since it is densely defined by assumption, its domain $\mathrm{Ran}(A+z)$ must be equal to $\mathfrak{H}$. Replacing z by z^* and applying Lemma 2.3 finishes the general case. The argument for the non-negative case with $z < 0$ is similar using $\varepsilon\|\psi\|^2 \le |\langle\psi, (A+\varepsilon)\psi\rangle|^2 \le \|\psi\|\|(A+\varepsilon)\psi\|$ which shows $(A+\varepsilon)^{-1} \le \varepsilon^{-1}$, $\varepsilon > 0$. □

In addition, we can also prove the closed graph theorem which shows that an unbounded operator cannot be defined on the entire Hilbert space.

Theorem 2.7 (closed graph)**.** *Let $\mathfrak{H}_1$ and $\mathfrak{H}_2$ be two Hilbert spaces and A an operator defined on all of $\mathfrak{H}_1$. Then A is bounded if and only if $\Gamma(A)$ is closed.*

Proof. If A is bounded than it is easy to see that $\Gamma(A)$ is closed. So let us assume that $\Gamma(A)$ is closed. Then A^* is well defined and for all unit vectors $\varphi \in \mathfrak{D}(A^*)$ we have that the linear functional $\ell_\varphi(\psi) = \langle A^*\varphi, \psi\rangle$ is pointwise bounded

$$\|\ell_\varphi(\psi)\| = |\langle\varphi, A\psi\rangle| \le \|A\psi\|. \quad (2.49)$$

Hence by the uniform boundedness principle there is a constant C such that $\|\ell_\varphi\| = \|A^*\varphi\| \le C$. That is, A^* is bounded and so is $A = A^{**}$. □

Finally we want to draw some some further consequences of Axiom 2 and show that observables correspond to self-adjoint operators. Since self-adjoint operators are already maximal, the difficult part remaining is to show that an observable has at least one self-adjoint extension. There is a good way of doing this for non-negative operators and hence we will consider this case first.

An operator is called **non-negative** (resp. **positive**) if $\langle\psi, A\psi\rangle \ge 0$ (resp. > 0 for $\psi \ne 0$) for all $\psi \in \mathfrak{D}(A)$. If A is positive, the map $(\varphi, \psi) \mapsto \langle\varphi, A\psi\rangle$ is a scalar product. However, there might be sequences which are Cauchy with respect to this scalar product but not with respect to our original one. To avoid this, we introduce the scalar product

$$\langle\varphi, \psi\rangle_A = \langle\varphi, (A+1)\psi\rangle, \qquad A \ge 0, \quad (2.50)$$

defined on $\mathfrak{D}(A)$, which satisfies $\|\psi\| \le \|\psi\|_A$. Let $\mathfrak{H}_A$ be the completion of $\mathfrak{D}(A)$ with respect to the above scalar product. We claim that $\mathfrak{H}_A$ can be regarded as a subspace of $\mathfrak{H}$, that is, $\mathfrak{D}(A) \subseteq \mathfrak{H}_A \subseteq \mathfrak{H}$.

If (ψ_n) is a Cauchy sequence in $\mathfrak{D}(A)$, then it is also Cauchy in $\mathfrak{H}$ (since $\|\psi\| \le \|\psi\|_A$ by assumption) and hence we can identify it with the limit of (ψ_n) regarded as a sequence in $\mathfrak{H}$. For this identification to be unique, we need to show that if $(\psi_n) \subset \mathfrak{D}(A)$ is a Cauchy sequence in $\mathfrak{H}_A$ such that $\|\psi_n\| \to 0$, then $\|\psi_n\|_A \to 0$. This follows from

$$\begin{aligned} \|\psi_n\|_A^2 &= \langle \psi_n, \psi_n - \psi_m \rangle_A + \langle \psi_n, \psi_m \rangle_A \\ &\le \|\psi_n\|_A \|\psi_n - \psi_m\|_A + \|\psi_n\| \|(A+1)\psi_m\| \end{aligned} \tag{2.51}$$

since the right hand side can be made arbitrarily small choosing m, n large.

Clearly the quadratic form q_A can be extended to every $\psi \in \mathfrak{H}_A$ by setting

$$q_A(\psi) = \langle \psi, \psi \rangle_A - \|\psi\|^2, \qquad \psi \in \mathfrak{Q}(A) = \mathfrak{H}_A. \tag{2.52}$$

The set $\mathfrak{Q}(A)$ is also called the **form domain** of A.

Now we come to our extension result. Note that $A + 1$ is injective and the best we can hope for is that for a non-negative extension $\tilde{A}$, $\tilde{A} + 1$ is a bijection from $\mathfrak{D}(\tilde{A})$ onto $\mathfrak{H}$.

Lemma 2.8. *Suppose A is a non-negative operator, then there is a non-negative extension $\tilde{A}$ such that* $\mathrm{Ran}(\tilde{A} + 1) = \mathfrak{H}$.

Proof. Let us define an operator $\tilde{A}$ by

$$\begin{aligned} \mathfrak{D}(\tilde{A}) &= \{\psi \in \mathfrak{H}_A | \exists \tilde{\psi} \in \mathfrak{H} : \langle \varphi, \psi \rangle_A = \langle \varphi, \tilde{\psi} \rangle, \forall \varphi \in \mathfrak{H}_A\} \\ \tilde{A}\psi &= \tilde{\psi} - \psi \end{aligned} . \tag{2.53}$$

Since $\mathfrak{H}_A$ is dense, $\tilde{\psi}$ is well-defined. Moreover, it is straightforward to see that $\tilde{A}$ is a non-negative extension of A.

It is also not hard to see that $\mathrm{Ran}(\tilde{A} + 1) = \mathfrak{H}$. Indeed, for any $\tilde{\psi} \in \mathfrak{H}$, $\varphi \mapsto \langle \tilde{\psi}, \varphi \rangle$ is bounded linear functional on $\mathfrak{H}_A$. Hence there is an element $\psi \in \mathfrak{H}_A$ such that $\langle \tilde{\psi}, \varphi \rangle = \langle \psi, \varphi \rangle_A$ for all $\varphi \in \mathfrak{H}_A$. By the definition of $\tilde{A}$, $(\tilde{A} + 1)\psi = \tilde{\psi}$ and hence $\tilde{A} + 1$ is onto. □

Now it is time for an

Example. Let us take $\mathfrak{H} = L^2(0, \pi)$ and consider the operator

$$Af = -\frac{d^2}{dx^2} f, \quad \mathfrak{D}(A) = \{f \in C^2[0, \pi] \,|\, f(0) = f(\pi) = 0\}, \tag{2.54}$$

which corresponds to the one-dimensional model of a particle confined to a box.

(i). First of all, using integration by parts twice, it is straightforward to check that A is symmetric

$$\int_0^\pi g(x)^*(-f'')(x)dx = \int_0^\pi g'(x)^* f'(x)dx = \int_0^\pi (-g'')(x)^* f(x)dx. \quad (2.55)$$

Note that the *boundary conditions* $f(0) = f(\pi) = 0$ are chosen such that the boundary terms occurring from integration by parts vanish. Moreover, the same calculation also shows that A is positive

$$\int_0^\pi f(x)^*(-f'')(x)dx = \int_0^\pi |f'(x)|^2 dx > 0, \quad f \neq 0. \quad (2.56)$$

(ii). Next let us show $\mathfrak{H}_A = \{f \in AC[0,\pi] \,|\, f(0) = f(\pi) = 0\}$. In fact, since

$$\langle g, f\rangle_A = \int_0^\pi \big(g'(x)^* f'(x) + g(x)^* f(x)\big)\, dx, \quad (2.57)$$

we see that f_n is Cauchy in $\mathfrak{H}_A$ if and only if both f_n and f'_n are Cauchy in $L^2(0,\pi)$. Thus $f_n \to f$ and $f'_n \to g$ in $L^2(0,\pi)$ and $f_n(x) = \int_0^x f'_n(t)dt$ implies $f(x) = \int_0^x g(t)dt$. Thus $f \in AC[0,\pi]$. Moreover, $f(0) = 0$ is obvious and from $0 = f_n(\pi) = \int_0^\pi f'_n(t)dt$ we have $f(\pi) = \lim_{n\to\infty} \int_0^\pi f'_n(t)dt = 0$. So we have $\mathfrak{H}_A \subseteq \{f \in AC[0,\pi] \,|\, f(0) = f(\pi) = 0\}$. To see the converse approximate f' by smooth functions g_n. Using $g_n - \int_0^\pi g_n(t)dt$ instead of g_n it is no restriction to assume $\int_0^\pi g_n(t)dt = 0$. Now define $f_n(x) = \int_0^x g_n(t)dt$ and note $f_n \in \mathfrak{D}(A) \to f$.

(iii). Finally, let us compute the extension $\tilde{A}$. We have $f \in \mathfrak{D}(\tilde{A})$ if for all $g \in \mathfrak{H}_A$ there is an $\tilde{f}$ such that $\langle g, f\rangle_A = \langle g, \tilde{f}\rangle$. That is,

$$\int_0^\pi g'(x)^* f'(x)dx = \int_0^\pi g(x)^*(\tilde{f}(x) - f(x))dx. \quad (2.58)$$

Integration by parts on the right hand side shows

$$\int_0^\pi g'(x)^* f'(x)dx = -\int_0^\pi g'(x)^* \int_0^x (\tilde{f}(t) - f(t))dt\, dx \quad (2.59)$$

or equivalently

$$\int_0^\pi g'(x)^* \left(f'(x) + \int_0^x (\tilde{f}(t) - f(t))dt \right) dx = 0. \quad (2.60)$$

Now observe $\{g' \in \mathfrak{H} | g \in \mathfrak{H}_A\} = \{g \in \mathfrak{H} | \int_0^\pi g(t)dt = 0\} = \{1\}^\perp$ and thus $f'(x) + \int_0^x (\tilde{f}(t) - f(t))dt \in \{1\}^{\perp\perp} = \mathrm{span}\{1\}$. So we see $f \in AC^1[0,\pi] = \{f \in AC[0,\pi] | f' \in AC[0,\pi]\}$ and $\tilde{A}f = -f''$. The converse is easy and hence

$$\tilde{A}f = -\frac{d^2}{dx^2} f, \quad \mathfrak{D}(\tilde{A}) = \{f \in AC^1[0,\pi] \,|\, f(0) = f(\pi) = 0\}. \quad (2.61)$$

◇

Now let us apply this result to operators A corresponding to observables. Since A will, in general, not satisfy the assumptions of our lemma, we will consider $1+A^2$, $\mathfrak{D}(1+A^2)=\mathfrak{D}(A^2)$, instead, which has a symmetric extension whose range is $\mathfrak{H}$. By our requirement for observables, $1+A^2$ is maximally defined and hence is equal to this extension. In other words, $\mathrm{Ran}(1+A^2)=\mathfrak{H}$. Moreover, for any $\varphi\in\mathfrak{H}$ there is a $\psi\in\mathfrak{D}(A^2)$ such that

$$(A-\mathrm{i})(A+\mathrm{i})\psi=(A+\mathrm{i})(A-\mathrm{i})\psi=\varphi \tag{2.62}$$

and since $(A\pm\mathrm{i})\psi\in\mathfrak{D}(A)$, we infer $\mathrm{Ran}(A\pm\mathrm{i})=\mathfrak{H}$. As an immediate consequence we obtain

Corollary 2.9. *Observables correspond to self-adjoint operators.*

But there is another important consequence of the results which is worth while mentioning.

Theorem 2.10 (Friedrichs extension)**.** *Let A be a* **semi-bounded** *symmetric operator, that is,*

$$q_A(\psi)=\langle\psi,A\psi\rangle\geq\gamma\|\psi\|^2,\qquad \gamma\in\mathbb{R}. \tag{2.63}$$

Then there is a self-adjoint extension $\tilde{A}$ which is also bounded from below by γ and which satisfies $\mathfrak{D}(\tilde{A})\subseteq\mathfrak{H}_{A-\gamma}$.

Proof. Replacing A by $A-\gamma$ we can reduce it to the case considered in Lemma 2.8. The rest is straightforward. □

2.3. Resolvents and spectra

Let A be a (densely defined) closed operator. The **resolvent set** of A is defined by

$$\rho(A)=\{z\in\mathbb{C}|(A-z)^{-1}\in\mathfrak{L}(\mathfrak{H})\}. \tag{2.64}$$

More precisely, $z\in\rho(A)$ if and only if $(A-z):\mathfrak{D}(A)\to\mathfrak{H}$ is bijective and its inverse is bounded. By the closed graph theorem (Theorem 2.7), it suffices to check that $A-z$ is bijective. The complement of the resolvent set is called the **spectrum**

$$\sigma(A)=\mathbb{C}\backslash\rho(A) \tag{2.65}$$

of A. In particular, $z\in\sigma(A)$ if $A-z$ has a nontrivial kernel. A vector $\psi\in\mathrm{Ker}(A-z)$ is called an **eigenvector** and z is called **eigenvalue** in this case.

The function

$$\begin{array}{rccc} R_A: & \rho(A) & \to & \mathfrak{L}(\mathfrak{H}) \\ & z & \mapsto & (A-z)^{-1} \end{array} \tag{2.66}$$

is called **resolvent** of A. Note the convenient formula

$$R_A(z)^* = ((A-z)^{-1})^* = ((A-z)^*)^{-1} = (A^*-z^*)^{-1} = R_{A^*}(z^*). \quad (2.67)$$

In particular,

$$\rho(A^*) = \rho(A)^*. \quad (2.68)$$

If $z, z' \in \rho(A)$, we have the **first resolvent formula**

$$R_A(z) - R_A(z') = (z-z')R_A(z)R_A(z') = (z-z')R_A(z')R_A(z). \quad (2.69)$$

In fact,

$$\begin{aligned} &(A-z)^{-1} - (z-z')(A-z)^{-1}(A-z')^{-1} = \\ &\qquad (A-z)^{-1}(1-(z-A+A-z')(A-z')^{-1}) = (A-z')^{-1}, \end{aligned} \quad (2.70)$$

which proves the first equality. The second follows after interchanging z and z'. Now fix $z' = z_0$ and use (2.69) recursively to obtain

$$R_A(z) = \sum_{j=0}^{n} (z-z_0)^j R_A(z_0)^{j+1} + (z-z_0)^{n+1} R_A(z_0)^{n+1} R_A(z). \quad (2.71)$$

The sequence of bounded operators

$$R_n = \sum_{j=0}^{n} (z-z_0)^j R_A(z_0)^{j+1} \quad (2.72)$$

converges to a bounded operator if $|z-z_0| < \|R_A(z_0)\|^{-1}$ and clearly we expect $z \in \rho(A)$ and $R_n \to R_A(z)$ in this case. Let $R_\infty = \lim_{n\to\infty} R_n$ and set $\varphi_n = R_n\psi$, $\varphi = R_\infty\psi$ for some $\psi \in \mathfrak{H}$. Then a quick calculation shows

$$AR_n\psi = \psi + (z-z_0)\varphi_{n-1} + z\varphi_n. \quad (2.73)$$

Hence $(\varphi_n, A\varphi_n) \to (\varphi, \psi + z\varphi)$ shows $\varphi \in \mathfrak{D}(A)$ (since A is closed) and $(A-z)R_\infty\psi = \psi$. Similarly, for $\psi \in \mathfrak{D}(A)$,

$$R_n A\psi = \psi + (z-z_0)\varphi_{n-1} + z\varphi_n \quad (2.74)$$

and hence $R_\infty(A-z)\psi = \psi$ after taking the limit. Thus $R_\infty = R_A(z)$ as anticipated.

If A is bounded, a similar argument verifies the **Neumann series** for the resolvent

$$\begin{aligned} R_A(z) &= -\sum_{j=0}^{n-1} \frac{A^j}{z^{j+1}} + \frac{1}{z^n} A^n R_A(z) \\ &= -\sum_{j=0}^{\infty} \frac{A^j}{z^{j+1}}, \quad |z| > \|A\|. \end{aligned} \quad (2.75)$$

In summary we have proved the following

Theorem 2.11. *The resolvent set $\rho(A)$ is open and $R_A : \rho(A) \to \mathfrak{L}(\mathfrak{H})$ is holomorphic, that is, it has an absolutely convergent power series expansion around every point $z_0 \in \rho(A)$. In addition,*

$$\|R_A(z)\| \geq \operatorname{dist}(z, \sigma(A))^{-1} \tag{2.76}$$

and if A is bounded we have $\{z \in \mathbb{C} | \, |z| > \|A\|\} \subseteq \rho(A)$.

As a consequence we obtain the useful

Lemma 2.12. *We have $z \in \sigma(A)$ if there is a sequence $\psi_n \in \mathfrak{D}(A)$ such that $\|\psi_n\| = 1$ and $\|(A - z)\psi_n\| \to 0$. If z is a boundary point of $\rho(A)$, then the converse is also true. Such a sequence is called* **Weyl sequence***.*

Proof. Let ψ_n be a Weyl sequence. Then $z \in \rho(A)$ is impossible by $1 = \|\psi_n\| = \|R_A(z)(A - z)\psi_n\| \leq \|R_A(z)\|\|(A - z)\psi_n\| \to 0$. Conversely, by (2.76) there is a sequence $z_n \to z$ and corresponding vectors $\varphi_n \in \mathfrak{H}$ such that $\|R_A(z)\varphi_n\|\|\varphi_n\|^{-1} \to \infty$. Let $\psi_n = R_A(z_n)\varphi_n$ and rescale φ_n such that $\|\psi_n\| = 1$. Then $\|\varphi_n\| \to 0$ and hence

$$\|(A - z)\psi_n\| = \|\varphi_n + (z_n - z)\psi_n\| \leq \|\varphi_n\| + |z - z_n| \to 0 \tag{2.77}$$

shows that ψ_n is a Weyl sequence. □

Let us also note the following spectral mapping result.

Lemma 2.13. *Suppose A is injective, then*

$$\sigma(A^{-1})\backslash\{0\} = (\sigma(A)\backslash\{0\})^{-1}. \tag{2.78}$$

In addition, we have $A\psi = z\psi$ if and only if $A^{-1}\psi = z^{-1}\psi$.

Proof. Suppose $z \in \rho(A)\backslash\{0\}$. Then we claim

$$R_{A^{-1}}(z^{-1}) = -zAR_A(z) = -z - R_A(z). \tag{2.79}$$

In fact, the right hand side is a bounded operator from $\mathfrak{H} \to \operatorname{Ran}(A) = \mathfrak{D}(A^{-1})$ and

$$(A^{-1} - z^{-1})(-zAR_A(z))\varphi = (-z + A)R_A(z)\varphi = \varphi, \quad \varphi \in \mathfrak{H}. \tag{2.80}$$

Conversely, if $\psi \in \mathfrak{D}(A^{-1}) = \operatorname{Ran}(A)$ we have $\psi = A\varphi$ and hence

$$(-zAR_A(z))(A^{-1} - z^{-1})\psi = AR_A(z)((A - z)\varphi) = A\varphi = \psi. \tag{2.81}$$

Thus $z^{-1} \in \rho(A^{-1})$. The rest follows after interchanging the roles of A and A^{-1}. □

Next, let us characterize the spectra of self-adjoint operators.

Theorem 2.14. *Let A be symmetric. Then A is self-adjoint if and only if $\sigma(A) \subseteq \mathbb{R}$ and $A \geq 0$ if and only if $\sigma(A) \subseteq [0, \infty)$. Moreover, $\|R_A(z)\| \leq |\operatorname{Im}(z)|^{-1}$ and, if $A \geq 0$, $\|R_A(\lambda)\| \leq |\lambda|^{-1}$, $\lambda < 0$.*

Proof. If $\sigma(A) \subseteq \mathbb{R}$, then $\mathrm{Ran}(A+z) = \mathfrak{H}$, $z \in \mathbb{C}\backslash\mathbb{R}$, and hence A is self-adjoint by Lemma 2.6. Conversely, if A is self-adjoint (resp. $A \geq 0$), then $R_A(z)$ exists for $z \in \mathbb{C}\backslash\mathbb{R}$ (resp. $z \in \mathbb{C}\backslash(-\infty, 0]$) and satisfies the given estimates as has been shown in the proof of Lemma 2.6. □

In particular, the converse part in Lemma 2.12 holds for self-adjoint operators and hence the closure of the set of all possible expectation values equals the spectrum of the corresponding operator:

Theorem 2.15. *Let A be self-adjoint, then*

$$\sigma(A) = \overline{\{\langle\psi, A\psi\rangle \,|\, \psi \in \mathfrak{D}(A),\ \|\psi\| = 1\}} \tag{2.82}$$

and in particular,

$$\inf \sigma(A) = \inf_{\psi \in \mathfrak{D}(A),\ \|\psi\|=1} \langle\psi, A\psi\rangle. \tag{2.83}$$

For the eigenvalues and corresponding eigenfunctions we have:

Lemma 2.16. *Let A be symmetric. Then all eigenvalues are real and eigenvectors corresponding to different eigenvalues are orthogonal.*

Proof. If $A\psi_j = \lambda_j\psi_j$, $j = 1, 2$, we have

$$\lambda_1\|\psi_1\|^2 = \langle\psi_1, \lambda_1\psi_1\rangle = \langle\psi_1, A\psi_1\rangle = \langle\psi_1, A\psi_1\rangle = \langle\lambda_1\psi_1, \psi_1\rangle = \lambda_1^*\|\psi_1\|^2 \tag{2.84}$$

and

$$(\lambda_1 - \lambda_2)\langle\psi_1, \psi_2\rangle = \langle A\psi_1, \psi_2\rangle - \langle A\psi_1, \psi_2\rangle = 0, \tag{2.85}$$

finishing the proof. □

The result does not imply that two linearly independent eigenfunctions to the same eigenvalue are orthogonal. However, it is no restriction to assume that they are since we can use Gram-Schmidt to find an orthonormal basis for $\mathrm{Ker}(A - \lambda)$. If $\mathfrak{H}$ is finite dimensional, we can always find an orthonormal basis of eigenvectors. In the infinite dimensional case this is no longer true in general. However, if there is an orthonormal basis of eigenvectors, then A is essentially self-adjoint.

Theorem 2.17. *Suppose A is a symmetric operator which has an orthonormal basis of eigenfunctions $\{\varphi_j\}$, then A is essentially self-adjoint. In particular, it is essentially self-adjoint on* $\mathrm{span}\{\varphi_j\}$.

Proof. Consider the set of all finite linear combinations $\psi = \sum_{j=0}^n c_j\varphi_j$ which is dense in $\mathfrak{H}$. Then $\phi = \sum_{j=0}^n \frac{c_j}{\lambda_j \pm \mathrm{i}}\varphi_j \in \mathfrak{D}(A)$ and $(A \pm \mathrm{i})\phi = \psi$ shows that $\mathrm{Ran}(A \pm \mathrm{i})$ is dense. □

In addition, we note the following asymptotic expansion for the resolvent.

Lemma 2.18. *Suppose A is self-adjoint. For every $\psi \in \mathfrak{H}$ we have*

$$\lim_{\mathrm{Im}(z)\to\infty} \|AR_A(z)\psi\| = 0. \tag{2.86}$$

In particular, if $\psi \in \mathfrak{D}(A^n)$, then

$$R_A(z)\psi = -\sum_{j=0}^{n} \frac{A^j\psi}{z^{j+1}} + o(\frac{1}{z^{n+1}}), \quad \textit{as} \quad \mathrm{Im}(z) \to \infty. \tag{2.87}$$

Proof. It suffices to prove the first claim since the second then follows as in (2.75).

Write $\psi = \tilde{\psi} + \varphi$, where $\tilde{\psi} \in \mathfrak{D}(A)$ and $\|\varphi\| \le \varepsilon$. Then

$$\begin{aligned} \|AR_A(z)\psi\| &\le \|R_A(z)A\tilde{\psi}\| + \|AR_A(z)\varphi\| \\ &\le \frac{\|A\tilde{\psi}\|}{\mathrm{Im}(z)} + \|\varphi\|, \end{aligned} \tag{2.88}$$

by (2.48), finishing the proof. □

Similarly, we can characterize the spectra of unitary operators. Recall that a bijection U is called unitary if $\langle U\psi, U\psi\rangle = \langle \psi, U^*U\psi\rangle = \langle \psi, \psi\rangle$. Thus U is unitary if and only if

$$U^* = U^{-1}. \tag{2.89}$$

Theorem 2.19. *Let U be unitary, then $\sigma(U) \subseteq \{z \in \mathbb{C} | |z| = 1\}$. All eigenvalues have modulus one and eigenvectors corresponding to different eigenvalues are orthogonal.*

Proof. Since $\|U\| \le 1$ we have $\sigma(U) \subseteq \{z \in \mathbb{C} | |z| \le 1\}$. Moreover, U^{-1} is also unitary and hence $\sigma(U) \subseteq \{z \in \mathbb{C} | |z| \ge 1\}$ by Lemma 2.13. If $U\psi_j = z_j\psi_j$, $j = 1, 2$ we have

$$(z_1 - z_2)\langle \psi_1, \psi_2\rangle = \langle U^*\psi_1, \psi_2\rangle - \langle \psi_1, U\psi_2\rangle = 0 \tag{2.90}$$

since $U\psi = z\psi$ implies $U^*\psi = U^{-1}\psi = z^{-1}\psi = z^*\psi$. □

2.4. Orthogonal sums of operators

Let $\mathfrak{H}_j$, $j = 1, 2$, be two given Hilbert spaces and let $A_j : \mathfrak{D}(A_j) \to \mathfrak{H}_j$ be two given operators. Setting $\mathfrak{H} = \mathfrak{H}_1 \oplus \mathfrak{H}_2$ we can define an operator

$$A = A_1 \oplus A_2, \qquad \mathfrak{D}(A) = \mathfrak{D}(A_1) \oplus \mathfrak{D}(A_2) \tag{2.91}$$

by setting $A(\psi_1 + \psi_2) = A_1\psi_1 + A_2\psi_2$ for $\psi_j \in \mathfrak{D}(A_j)$. Clearly A is closed, (essentially) self-adjoint, etc., if and only if both A_1 and A_2 are. Moreover, it is straightforward to verify

Theorem 2.20. *Suppose A_j are self-adjoint operators on $\mathfrak{H}_j$, then $A = A_1 \oplus A_2$ is self-adjoint and*

$$R_A(z) = R_{A_1}(z) \oplus R_{A_2}(z), \qquad z \in \rho(A) = \rho(A_1) \cap \rho(A_2). \tag{2.92}$$

In particular,

$$\sigma(A) = \sigma(A_1) \cup \sigma(A_2). \tag{2.93}$$

Conversely, given an operator A it might be useful to write A as orthogonal sum and investigate each part separately.

Let $\mathfrak{H}_1 \subseteq \mathfrak{H}$ be a closed subspace and let P_1 be the corresponding projector. We say that $\mathfrak{H}_1$ **reduces** the operator A if $P_1 A \subseteq A P_1$. Note that this implies $P_1\mathfrak{D}(A) \subseteq \mathfrak{D}(A)$. Moreover, if we set $\mathfrak{H}_2 = \mathfrak{H}_1^\perp$, we have $\mathfrak{H} = \mathfrak{H}_1 \oplus \mathfrak{H}_2$ and $P_2 = \mathbb{I} - P_1$ reduces A as well.

Lemma 2.21. *Suppose $\mathfrak{H}_1 \subseteq \mathfrak{H}$ reduces A, then $A = A_1 \oplus A_2$, where*

$$A_j \psi = A\psi, \qquad \mathfrak{D}(A_j) = P_j \mathfrak{D}(A) \subseteq \mathfrak{D}(A). \tag{2.94}$$

If A is closable, then $\mathfrak{H}_1$ also reduces $\overline{A}$ and

$$\overline{A} = \overline{A_1} \oplus \overline{A_2}. \tag{2.95}$$

Proof. As already noted, $P_1\mathfrak{D}(A) \subseteq \mathfrak{D}(A)$ and hence $P_2\mathfrak{D}(A) = (\mathbb{I} - P_1)\mathfrak{D}(A) \subseteq \mathfrak{D}(A)$. Thus we see $\mathfrak{D}(A) = \mathfrak{D}(A_1) \oplus \mathfrak{D}(A_2)$. Moreover, if $\psi \in \mathfrak{D}(A_j)$ we have $A\psi = AP_j\psi = P_jA\psi \in \mathfrak{H}_j$ and thus $A_j : \mathfrak{D}(A_j) \to \mathfrak{H}_j$ which proves the first claim.

Now let us turn to the second claim. Clearly $\overline{A} \subseteq \overline{A_1} \oplus \overline{A_2}$. Conversely, suppose $\psi \in \mathfrak{D}(\overline{A})$, then there is a sequence $\psi_n \in \mathfrak{D}(A)$ such that $\psi_n \to \psi$ and $A\psi_n \to A\psi$. Then $P_j\psi_n \to P_j\psi$ and $AP_j\psi_n = P_jA\psi_n \to PA\psi$. In particular, $P_j\psi \in \mathfrak{D}(\overline{A})$ and $AP_j\psi = PA\psi$. □

The same considerations apply to countable orthogonal sums

$$A = \bigoplus_j A_j, \qquad \mathfrak{D}(A) = \bigoplus_j \mathfrak{D}(A_j). \tag{2.96}$$

If A is self-adjoint, then $\mathfrak{H}_1$ reduces A if $P_1\mathfrak{D}(A) \subseteq \mathfrak{D}(A)$ and $AP_1\psi \in \mathfrak{H}_1$ for every $\psi \in \mathfrak{D}(A)$. In fact, if $\psi \in \mathfrak{D}(A)$ we can write $\psi = \psi_1 \oplus \psi_2$, $\psi_j = P_j\psi \in \mathfrak{D}(A)$. Since $AP_1\psi = A\psi_1$ and $P_1A\psi = P_1A\psi_1 + P_1A\psi_2 = A\psi_1 + P_1A\psi_2$ we need to show $P_1A\psi_2 = 0$. But this follows since

$$\langle \varphi, P_1A\psi_2 \rangle = \langle AP_1\varphi, \psi_2 \rangle = 0 \tag{2.97}$$

for every $\varphi \in \mathfrak{D}(A)$.

2.5. Self-adjoint extensions

In many physical applications a symmetric operator is given. If this operator turns out to be essentially self-adjoint, there is a unique self-adjoint extension and everything is fine. However, if it is not, it is important to find out if there are self-adjoint extensions at all (for physical problems there better are) and to classify them. It is safe to skip this entire section on first reading.

In Section 2.2 we have seen that A is essentially self-adjoint if $\mathrm{Ker}(A^* - z) = \mathrm{Ker}(A^* - z^*) = \{0\}$ for one $z \in \mathbb{C}\backslash\mathbb{R}$. Hence self-adjointness is related to the dimension of these spaces and one calls the numbers

$$d_\pm(A) = \dim K_\pm, \quad K_\pm = \mathrm{Ran}(A \pm \mathrm{i})^\perp = \mathrm{Ker}(A^* \mp \mathrm{i}), \tag{2.98}$$

defect indices of A (we have chosen $z = \mathrm{i}$ for simplicity, every other $z \in \mathbb{C}\backslash\mathbb{R}$ would be as good). If $d_-(A) = d_+(A) = 0$ there is one self-adjoint extension of A, namely $\overline{A}$. But what happens in the general case? Is there more than one extension, or maybe none at all? These questions can be answered by virtue of the **Cayley transform**

$$V = (A - \mathrm{i})(A + \mathrm{i})^{-1} : \mathrm{Ran}(A + \mathrm{i}) \to \mathrm{Ran}(A - \mathrm{i}). \tag{2.99}$$

Theorem 2.22. *The Cayley transform is a bijection from the set of all symmetric operators A to the set of all isometric operators V (i.e., $\|V\varphi\| = \|\varphi\|$ for all $\varphi \in \mathfrak{D}(V)$) for which $\mathrm{Ran}(1 + V)$ is dense.*

Proof. Since A is symmetric we have $\|(A \pm \mathrm{i})\psi\|^2 = \|A\psi\|^2 + \|\psi\|^2$ for all $\psi \in \mathfrak{D}(A)$ by a straightforward computation. And thus for every $\varphi = (A + \mathrm{i})\psi \in \mathfrak{D}(V) = \mathrm{Ran}(A + \mathrm{i})$ we have

$$\|V\varphi\| = \|(A - \mathrm{i})\psi\| = \|(A + \mathrm{i})\psi\| = \|\varphi\|. \tag{2.100}$$

Next observe

$$1 \pm V = ((A - \mathrm{i}) \pm (A + \mathrm{i}))(A + \mathrm{i})^{-1} = \begin{cases} 2A(A + \mathrm{i})^{-1} \\ 2\mathrm{i}(A + \mathrm{i})^{-1} \end{cases}, \tag{2.101}$$

which shows $\mathfrak{D}(A) = \mathrm{Ran}(1 - V)$ and

$$A = \mathrm{i}(1 + V)(1 - V)^{-1}. \tag{2.102}$$

Conversely, let V be given and use the last equation to define A.

Since A is symmetric we have $\langle (1 \pm V)\varphi, (1 \mp V)\varphi\rangle = \pm 2\mathrm{i}\langle V\varphi, \varphi\rangle$ for all $\varphi \in \mathfrak{D}(V)$ by a straightforward computation. And thus for every $\psi = (1 - V)\varphi \in \mathfrak{D}(A) = \mathrm{Ran}(1 - V)$ we have

$$\langle A\psi, \psi\rangle = -\mathrm{i}\langle (1{+}V)\varphi, (1{+}V)\varphi\rangle = \mathrm{i}\langle (1{+}V)\varphi, (1{+}V)\varphi\rangle = \langle \psi, A\psi\rangle, \tag{2.103}$$

that is, A is symmetric. Finally observe

$$A \pm \mathrm{i} = ((1 + V) \pm (1 - V))(1 - V)^{-1} = \begin{cases} 2\mathrm{i}(1 - V)^{-1} \\ 2\mathrm{i}V(1 - V)^{-1} \end{cases}, \tag{2.104}$$

which shows that A is the Cayley transform of V and finishes the proof. □

Thus A is self-adjoint if and only if its Cayley transform V is unitary. Moreover, finding a self-adjoint extension of A is equivalent to finding a unitary extensions of V. Finding a unitary extension V is equivalent to (taking the closure and) finding a unitary operator from $\mathfrak{D}(V)^\perp$ to $\mathrm{Ran}(V)^\perp$. This is possible if and only if both spaces have the same dimension, that is, if and only if $d_+(A) = d_-(A)$.

Theorem 2.23. *A symmetric operator has self-adjoint extensions if and only if its defect indices are equal.*

In this case let A_1 be a self-adjoint extension, V_1 its Cayley transform. Then

$$\mathfrak{D}(A_1) = \mathfrak{D}(A) + (1 - V_1)K_+ = \{\psi + \varphi_+ - V_1\varphi_+ | \psi \in \mathfrak{D}(A),\ \varphi_+ \in K_+\} \tag{2.105}$$

and

$$A_1(\psi + \varphi_+ - V_1\varphi_+) = A\psi + \mathrm{i}\varphi_+ + \mathrm{i}V_1\varphi_+. \tag{2.106}$$

Moreover,

$$(A_1 \pm \mathrm{i})^{-1} = (A \pm \mathrm{i})^{-1} \oplus \frac{\mp\mathrm{i}}{2}\sum_j \langle\varphi_j^\pm, .\rangle(\varphi_j^\pm - \varphi_j^\mp), \tag{2.107}$$

where $\{\varphi_j^+\}$ is an orthonormal basis for K_+ and $\varphi_j^- = V_1\varphi_j^+$.

Concerning closures we note that a bounded operator is closed if and only if its domain is closed and any operator is closed if and only if its inverse is closed. Hence we have

Lemma 2.24. *The following items are equivalent.*

- *A is closed.*
- $\mathfrak{D}(V) = \mathrm{Ran}(A + \mathrm{i})$ *is closed.*
- $\mathrm{Ran}(V) = \mathrm{Ran}(A - \mathrm{i})$ *is closed.*
- *V is closed.*

Next, we give a useful criterion for the existence of self-adjoint extensions. A skew linear map $C : \mathfrak{H} \to \mathfrak{H}$ is called a **conjugation** if it satisfies $C^2 = \mathbb{I}$ and $\langle C\psi, C\varphi\rangle = \langle\psi, \varphi\rangle$. The prototypical example is of course complex conjugation $C\psi = \psi^*$. An operator A is called C-**real** if

$$C\mathfrak{D}(A) \subseteq \mathfrak{D}(A), \quad \text{and} \quad AC\psi = CA\psi, \quad \psi \in \mathfrak{D}(A). \tag{2.108}$$

Note that in this case $C\mathfrak{D}(A) = \mathfrak{D}(A)$, since $\mathfrak{D}(A) = C^2\mathfrak{D}(A) \subseteq C\mathfrak{D}(A)$.

Theorem 2.25. *Suppose the symmetric operator A is C-real, then its defect indices are equal.*

Proof. Let $\{\varphi_j\}$ be an orthonormal set in $\mathrm{Ran}(A+\mathrm{i})^\perp$. Then $\{K\varphi_j\}$ is an orthonormal set in $\mathrm{Ran}(A-\mathrm{i})^\perp$. Hence $\{\varphi_j\}$ is an orthonormal basis for $\mathrm{Ran}(A+\mathrm{i})^\perp$ if and only if $\{K\varphi_j\}$ is an orthonormal basis for $\mathrm{Ran}(A-\mathrm{i})^\perp$. Hence the two spaces have the same dimension. □

Finally, we note the following useful formula for the difference of resolvents of self-adjoint extensions.

Lemma 2.26. *Suppose A is a closed symmetric operator with equal defect indices $d = d_+(A) = d_-(A)$. Then $\dim \mathrm{Ker}(A^* - z^*) = d$ for all z. Moreover, if A_j, $j = 1, 2$ are self-adjoint extensions and if $\{\varphi_j(z)\}$ is an orthonormal basis for $\mathrm{Ker}(A^* - z^*)$, then*

$$(A_1 - z)^{-1} - (A_2 - z)^{-1} = \sum_{j,k} (\alpha^1_{jk}(z) - \alpha^2_{jk}(z)) \langle \varphi_k(z), .\rangle \varphi_k(z^*), \quad (2.109)$$

where

$$\alpha^l_{jk}(z) = \langle \varphi_j(z^*), (A_l - z)^{-1} \varphi_k(z)\rangle. \quad (2.110)$$

Proof. First of all we note that instead of $z = \mathrm{i}$ we could use $V(z) = (A + z^*)(A + z)^{-1}$ for any $z \in \mathbb{C}\backslash\mathbb{R}$. Let $d_\pm(z) = \dim K_\pm(z)$, $K_+(z) = \mathrm{Ran}(A + z)^\perp$ respectively $K_-(z) = \mathrm{Ran}(A + z^*)^\perp$. The same arguments as before show that there is a one to one correspondence between the self-adjoint extensions of A and the unitary operators on $\mathbb{C}^{d(z)}$. Hence $d(z_1) = d(z_2) = d_\pm(A)$.

Now note that $((A_1-z)^{-1}-(A_2-z)^{-1})\varphi$ is zero for every $\varphi \in \mathrm{Ran}(A-z)$. Hence it suffices to consider it for vectors $\varphi = \sum_j \langle \varphi_j(z), \varphi\rangle \varphi_j(z) \in \mathrm{Ran}(A-z)^\perp$. Hence we have

$$(A_1 - z)^{-1} - (A_2 - z)^{-1} = \sum_j \langle \varphi_j(z), .\rangle \psi_j(z), \quad (2.111)$$

where

$$\psi_j(z) = ((A_1 - z)^{-1} - (A_2 - z)^{-1})\varphi_j(z). \quad (2.112)$$

Now computation the adjoint once using $((A_j - z)^{-1})^* = (A_j - z^*)^{-1}$ and once using $(\sum_j \langle \psi_j, .\rangle \varphi_j)^* = \sum_j \langle \varphi_j, .\rangle \psi_j$ we obtain

$$\sum_j \langle \varphi_j(z^*), .\rangle \psi_j(z^*) = \sum_j \langle \psi_j(z), .\rangle \varphi_j(z). \quad (2.113)$$

Evaluating at $\varphi_k(z)$ implies

$$\psi_k(z) = \sum_j \langle \psi_j(z^*), \varphi_k(z)\rangle \varphi_j(z^*) \quad (2.114)$$

and finishes the proof. □

Chapter 3

The spectral theorem

The time evolution of a quantum mechanical system is governed by the Schrödinger equation

$$\mathrm{i}\frac{d}{dt}\psi(t) = H\psi(t). \tag{3.1}$$

If $\mathfrak{H} = \mathbb{C}^n$, and H is hence a matrix, this system of ordinary differential equations is solved by the matrix exponential

$$\psi(t) = \exp(-\mathrm{i}tH)\psi(0). \tag{3.2}$$

This matrix exponential can be defined by a convergent power series. For this approach the boundedness of H is crucial, which might not be the case for a a quantum system. However, the best way to compute the matrix exponential, and to understand the underlying dynamics, is to diagonalize H. But how do we diagonalize a self-adjoint operator? The answer is know as the spectral theorem.

3.1. The spectral theorem

In this section we want to address the problem of defining functions of a self-adjoint operator A in a natural way, that is, such that

$$(f+g)(A) = f(A)+g(A), \quad (fg)(A) = f(A)g(A), \quad (f^*)(A) = f(A)^*. \tag{3.3}$$

As long as f and g are polynomials, no problems arise. If we want to extend this definition to a larger class of functions, we will need to perform some limiting procedure. Hence we could consider convergent power series or equip the space of polynomials with the sup norm. In both cases this only works if the operator A is bounded. To overcome this limitation, we will use characteristic functions $\chi_\Omega(A)$ instead of powers A^j. Since $\chi_\Omega(\lambda)^2 = \chi_\Omega(\lambda)$, the corresponding operators should be orthogonal projections. Moreover, we should also have $\chi_{\mathbb{R}}(A) = \mathbb{I}$ and $\chi_\Omega(A) = \sum_{j=1}^n \chi_{\Omega_j}(A)$ for any finite union $\Omega = \bigcup_{j=1}^n \Omega_j$ of disjoint sets. The only remaining problem is of course the definition of $\chi_\Omega(A)$. However, we will defer this problem and begin by developing a functional calculus for a family of characteristic functions $\chi_\Omega(A)$.

Denote the Borel sigma algebra of $\mathbb{R}$ by $\mathfrak{B}$. A **projection-valued measure** is a map

$$P : \mathfrak{B} \to \mathfrak{L}(\mathfrak{H}), \qquad \Omega \mapsto P(\Omega), \tag{3.4}$$

from the Borel sets to the set of orthogonal projections, that is, $P(\Omega)^* = P(\Omega)$ and $P(\Omega)^2 = P(\Omega)$, such that the following two conditions hold:

(1) $P(\mathbb{R}) = \mathbb{I}$.

(2) If $\Omega = \bigcup_n \Omega_n$ with $\Omega_n \cap \Omega_m = \emptyset$ for $n \neq m$, then $\sum_n P(\Omega_n)\psi = P(\Omega)\psi$ for every $\psi \in \mathfrak{H}$ (strong convergence).

Note that we require $\sum_n P(\Omega_n)\psi = P(\Omega)\psi$ rather than $\sum_n P(\Omega_n) = P(\Omega)$. That is, we do not require the sum to converge in the operator norm. In general, a sequence of bounded operators A_n is said to **converge strongly** to a bounded operator A if $A_n\psi \to A\psi$ for every $\psi \in \mathfrak{H}$. Clearly norm convergence implies strong convergence. For the sake of completeness, a sequence of bounded operators A_n is said to **converge weakly** to a bounded operator A if $\langle \varphi, A_n\psi\rangle \to \langle \varphi, A\psi\rangle$ for every $\varphi, \psi \in \mathfrak{H}$. Clearly strong convergence implies weak convergence.

It is straightforward to verify that any projection-valued measure satisfies

$$P(\emptyset) = 0, \qquad P(\mathbb{R}\backslash\Omega) = \mathbb{I} - P(\Omega), \tag{3.5}$$

and

$$P(\Omega_1 \cup \Omega_2) + P(\Omega_1 \cap \Omega_2) = P(\Omega_1) + P(\Omega_2). \tag{3.6}$$

Moreover, we also have

$$P(\Omega_1)P(\Omega_2) = P(\Omega_1 \cap \Omega_2). \tag{3.7}$$

Indeed, suppose $\Omega_1 \cap \Omega_2 = \emptyset$ first. Then, taking the square of (3.6) we infer

$$P(\Omega_1)P(\Omega_2) + P(\Omega_2)P(\Omega_1) = 0. \tag{3.8}$$

Multiplying this equation from the left by $P(\Omega_1)P(\Omega_2)$ and from the right by $P(\Omega_2)$ we have $(P(\Omega_1)P(\Omega_2))^2 = 0$ and hence $P(\Omega_1)P(\Omega_2) = 0$. For the general case $\Omega_1 \cap \Omega_2 \neq \emptyset$ we now have

$$\begin{aligned} P(\Omega_1)P(\Omega_2) &= (P(\Omega_1 - \Omega_2) + P(\Omega_1 \cap \Omega_2))(P(\Omega_2 - \Omega_1) + P(\Omega_1 \cap \Omega_2)) \\ &= P(\Omega_1 \cap \Omega_2) \end{aligned} \tag{3.9}$$

as stated.

We will abbreviate $P(\lambda) = P((-\infty, \lambda])$. Picking $\psi \in \mathfrak{H}$, we obtain a finite Borel measure $\mu_\psi(\Omega) = \langle \psi, P(\Omega)\psi\rangle$, $\mu_\psi(\mathbb{R}) = \|\psi\|^2 < \infty$.

Using the polarization identity (2.15) we also have the following complex Borel measures

$$\mu_{\varphi,\psi}(\Omega) = \langle \varphi, P(\Omega)\psi\rangle = \frac{1}{4}(\mu_{\varphi+\psi}(\Omega) - \mu_{\varphi-\psi}(\Omega) + \mathrm{i}\mu_{\varphi-\mathrm{i}\psi}(\Omega) - \mathrm{i}\mu_{\varphi+\mathrm{i}\psi}(\Omega)). \tag{3.10}$$

Note also that, by Cauchy-Schwarz, $|\mu_{\varphi,\psi}(\Omega)| \le \|\varphi\| \, \|\psi\|$. Now let us turn to integration with respect to our projection-valued measure. For any simple function $f = \sum_{j=1}^n c_j \chi_{\Omega_j}$ we set

$$P(f) \equiv \int_{\mathbb{R}} f(\lambda) dP(\lambda) = \sum_{j=1}^n c_j P(\Omega_j). \tag{3.11}$$

In particular, $P(\chi_\Omega) = P(\Omega)$. The operator P is a linear map from the set of simple functions into the set of bounded linear operators on $\mathfrak{H}$ satisfying

$$\langle \varphi, P(f)\psi\rangle = \int_{\mathbb{R}} f(\lambda) d\mu_{\varphi,\psi}(\lambda) \tag{3.12}$$

and

$$\|P(f)\psi\|^2 = \int_{\mathbb{R}} |f(\lambda)|^2 d\mu_\psi(\lambda) \le \sup_{\lambda\in\mathbb{R}} |f(\lambda)|^2 \|\psi\|^2. \tag{3.13}$$

Equipping the set of simple functions with the sup norm, there is a unique extension of P to a bounded linear operator $P : B(\mathbb{R}) \to \mathfrak{L}(\mathfrak{H})$ (whose norm is one) from the bounded Borel functions on $\mathbb{R}$ (with sup norm) to the set of bounded linear operators on $\mathfrak{H}$. In particular, (3.12) and (3.13) remain true. In addition, observe

$$d\mu_{P(g)\varphi,P(f)\psi} = g^* f d\mu_{\varphi,\psi} \tag{3.14}$$

and

$$\langle P(g)\varphi, P(f)\psi\rangle = \int_{\mathbb{R}} g^*(\lambda) f(\lambda) d\mu_{\varphi,\psi}(\lambda). \tag{3.15}$$

There is some additional structure behind this extension. Let us recall some definitions first. A **Banach algebra** $\mathcal{A}$ (with identity) is a Banach space which is at the same time an algebra such that

(1) $\|ab\| \le \|a\| \, \|b\|$, $a, b \in \mathcal{A}$,

(2) $\|\mathbb{I}\| = 1$.

If, in addition, there is a conjugate linear mapping $a \mapsto a^*$ such that

(1) $\|a^*\| = \|a\|$,

(2) $\|aa^*\| = \|a\| \, \|a^*\|$, $a \in \mathcal{A}$,

then $\mathcal{A}$ is called a C^* **algebra**. The element a^* is called the **adjoint** of a.

An element $a \in \mathcal{A}$ is called **normal** if $aa^* = a^*a$, **self-adjoint** if $a = a^*$, **unitary** if $aa^* = a^*a = \mathbb{I}$, (orthogonal) **projection** if $a = a^* = a^2$, and **positive** if $a = bb^*$ for some $b \in \mathcal{A}$.

A $*$-subalgebra is a subalgebra which is closed under the adjoint map. An **ideal** is a subspace $\mathcal{I} \subseteq \mathcal{A}$ such that $a \in \mathcal{I}$, $b \in \mathcal{A}$ implies $ab \in \mathcal{I}$ and $ba \in \mathcal{I}$. If it is closed under the adjoint map it is called a $*$-ideal.

For example, the set of all bounded Borel measurable functions $B(\mathbb{R})$ (with the sup norm) and the set $\mathfrak{L}(\mathfrak{H})$ of all bounded linear mappings on $\mathfrak{H}$ are C^* algebras.

A C^* algebra homomorphism ϕ is a linear map between two C^* algebras which respects both the multiplication and the adjoint, that is, $\phi(ab) = \phi(a)\phi(b)$, $\phi(a^{-1}) = \phi(a)^{-1}$ and $\phi(a^*) = \phi(a)^*$. Any C^* algebra homomorphism has norm one

$$\|\phi(a)\| \leq \|a\| \tag{3.16}$$

and is positivity preserving.

Theorem 3.1. *Let $P(\Omega)$ be a projection-valued measure on $\mathfrak{H}$. Then the operator*

$$\begin{array}{rccl} P: & B(\mathbb{R}) & \to & \mathfrak{L}(\mathfrak{H}) \\ & f & \mapsto & \int_{\mathbb{R}} f(\lambda) dP(\lambda) \end{array} \tag{3.17}$$

is a C^* algebra homomorphism.

In addition, if $f_n(x) \to f(x)$ pointwise and if the sequence $\sup_{\lambda\in\mathbb{R}} |f_n(\lambda)|$ is bounded, then $P(f_n) \to P(f)$ strongly.

Proof. The properties $P(1) = \mathbb{I}$, $P(f^*) = P(f)^*$, and $P(fg) = P(f)P(g)$ are straightforward for simple functions f. For general f they follow from continuity. Hence P is a C^* algebra homomorphism.

The last claim follows from the dominated convergence theorem and (3.13). □

Next we want to define this operator for unbounded Borel functions. Since we expect the resulting operator to be unbounded, we need a suitable domain first

$$\mathfrak{D}_f = \{\psi \in \mathfrak{H} | \int_{\mathbb{R}} |f(\lambda)|^2 d\mu_\psi(\lambda) < \infty\}. \tag{3.18}$$

This is clearly a linear subspace of $\mathfrak{H}$ since $\mu_{\alpha\psi}(\Omega) = |\alpha|^2 \mu_\psi(\Omega)$ and since $\mu_{\varphi+\psi}(\Omega) \leq 2(\mu_\varphi(\Omega) + \mu_\psi(\Omega))$ (by the triangle inequality).

For every $\psi \in \mathfrak{D}_f$, the bounded Borel function

$$f_n = \chi_{\Omega_n} f, \qquad \Omega_n = \{\lambda | \, |f(\lambda)| \leq n\}, \tag{3.19}$$

converges to f in the sense of $L^2(\mathbb{R}, d\mu_\psi)$. Moreover, because of (3.13), $P(f_n)\psi$ converges to some vector $\tilde{\psi}$. We define $P(f)\psi = \tilde{\psi}$. By construction, $P(f)$ is a linear operator such that (3.12) and (3.13) hold.

In addition, $\mathfrak{D}_f$ is dense. Indeed, let Ω_n be defined as in (3.19) and abbreviate $\psi_n = P(\Omega_n)\psi$. Now observe that $d\mu_{\psi_n} = \chi_{\Omega_n} d\mu_\psi$ and hence $\psi_n \in \mathfrak{D}_f$. Moreover, $\psi_n \to \psi$ by (3.13) since $\chi_{\Omega_n} \to 1$ in $L^2(\mathbb{R}, d\mu_\psi)$.

The operator $P(f)$ has some additional properties. One calls an unbounded operator A **normal** if $\mathfrak{D}(A) = \mathfrak{D}(A^*)$ and $\|A\psi\| = \|A^*\psi\|$ for all $\psi \in \mathfrak{D}(A)$.

Theorem 3.2. *For any Borel function f, the operator*

$$P(f) = \int_{\mathbb{R}} f(\lambda) dP(\lambda), \qquad \mathfrak{D}(P(f)) = \mathfrak{D}_f, \tag{3.20}$$

is normal and satisfies

$$P(f)^* = P(f^*). \tag{3.21}$$

Proof. Let f be given and define f_n, Ω_n as above. Since (3.21) holds for f_n by our previous theorem, we get

$$\langle \varphi, P(f)\psi \rangle = \langle P(f^*)\varphi, \psi \rangle \tag{3.22}$$

for any $\varphi, \psi \in \mathfrak{D}_f = \mathfrak{D}(f^*)$ by continuity. Thus it remains to show that $\mathfrak{D}(P(f)^*) \subseteq \mathfrak{D}_f$. If $\psi \in \mathfrak{D}(P(f)^*)$ we have $\langle \psi, P(f)\varphi \rangle = \langle \tilde{\psi}, \varphi \rangle$ for all $\varphi \in \mathfrak{D}_f$ by definition. Now observe that $P(f_n^*)\psi = P(\Omega_n)\tilde{\psi}$ since we have

$$\langle P(f_n^*)\psi, \varphi \rangle = \langle \psi, P(f_n)\varphi \rangle = \langle \psi, P(f)P(\Omega_n)\varphi \rangle = \langle P(\Omega_n)\tilde{\psi}, \varphi \rangle \tag{3.23}$$

for any $\varphi \in \mathfrak{H}$. To see the second equality use $P(f_n)\varphi = P(f_m\chi_n)\varphi = P(f_m)P(\Omega_n)\varphi$ for $m \geq n$ and let $m \to \infty$. This proves existence of the limit

$$\lim_{n\to\infty} \int_{\mathbb{R}} |f_n|^2 d\mu_\psi(\lambda) = \lim_{n\to\infty} \|P(f_n^*)\psi\|^2 = \lim_{n\to\infty} \|P(\Omega_n)\tilde{\psi}\|^2 = \|\tilde{\psi}\|^2, \tag{3.24}$$

which implies $f \in L^2(\mathbb{R}, d\mu_\psi)$, that is, $\psi \in \mathfrak{D}_f$. That $P(f)$ is normal follows from $\|P(f)\psi\| = \|P(f^*)\psi\| = \int_R |f(\lambda)|^2 d\mu_\psi$. □

These considerations seem to indicate some kind of correspondence between the operators $P(f)$ in $\mathfrak{H}$ and f in $L^2(\mathbb{R}, d\mu_\psi)$. Recall that $U : \mathfrak{H} \to \tilde{\mathfrak{H}}$ is called unitary if it is a bijection which preserves scalar products $\langle U\varphi, U\psi \rangle = \langle \varphi, \psi \rangle$. The operators A in $\mathfrak{H}$ and $\tilde{A}$ in $\tilde{\mathfrak{H}}$ are said to be **unitarily equivalent** if

$$UA = \tilde{A}U, \qquad U\mathfrak{D}(A) = \mathfrak{D}(\tilde{A}). \tag{3.25}$$

Clearly, A is self-adjoint if and only if $\tilde{A}$ is and $\sigma(A) = \sigma(\tilde{A})$.

Now let us return to our original problem and consider the subspace

$$\mathfrak{H}_\psi = \{P(f)\psi | f \in L^2(\mathbb{R}, d\mu_\psi)\} \subseteq \mathfrak{H}. \tag{3.26}$$

The vector ψ is called **cyclic** if $\mathfrak{H}_\psi = \mathfrak{H}$. By (3.13), the relation

$$U_\psi(P(f)\psi) = f \tag{3.27}$$

defines a unique unitary operator $U_\psi : \mathfrak{H}_\psi \to L^2(\mathbb{R}, d\mu_\psi)$ such that

$$U_\psi P(f) = f U_\psi, \tag{3.28}$$

where f is identified with its corresponding multiplication operator. Moreover, if f is unbounded we have $U_\psi(\mathfrak{D}_f \cap \mathfrak{H}_\psi) = \mathfrak{D}(f) = \{g \in L^2(\mathbb{R}, d\mu_\psi) | fg \in L^2(\mathbb{R}, d\mu_\psi)\}$ (since $\varphi = P(f)\psi$ implies $d\mu_\varphi = f d\mu_\psi$) and the above equation still holds.

If ψ is cyclic, our picture is complete. Otherwise we need to extend this approach. A set $\{\psi_j\}_{j\in J}$ (J some index set) is called a set of spectral vectors if $\|\psi_j\| = 1$ and $\mathfrak{H}_{\psi_i} \perp \mathfrak{H}_{\psi_j}$ for all $i \neq j$. A maximal set of spectral vectors is called a **spectral basis**. By Zorn's lemma there exists a spectral basis.

It is important to observe that the cardinality of a spectral basis is *not* well-defined (in contradistinction to the cardinality of an ordinary basis of the Hilbert space). However, it can be at most equal to the cardinality of an ordinary basis. In particular, since $\mathfrak{H}$ is separable, it is at most countable. The minimal cardinality of a spectral basis is called **spectral multiplicity** of P. If the spectral multiplicity is one, the spectrum is called **simple**.

For a spectral basis $\{\psi_j\}_{j\in J}$ we claim $\bigoplus_j \mathfrak{H}_{\psi_j} = \mathfrak{H}$. Indeed, if equality would not hold, we could find a $g \perp \bigoplus_j \mathfrak{H}_{\psi_j}$ and $\{\psi_j\}_{j\in J} \cup \{g\}$ would be a larger spectral set contradicting maximality.

In summary we have,

Lemma 3.3. *For every projection valued measure P, there is an (at most countable) spectral basis $\{\psi_n\}$ such that*

$$\mathfrak{H} = \bigoplus_n \mathfrak{H}_{\psi_n} \tag{3.29}$$

and a corresponding unitary operator

$$U = \bigoplus_n U_{\psi_n} : \mathfrak{H} \to \bigoplus_n L^2(\mathbb{R}, d\mu_{\psi_n}) \tag{3.30}$$

such that for any Borel function f,

$$UP(f) = fU, \qquad U\mathfrak{D}_f = \mathfrak{D}(f). \tag{3.31}$$

Using this canonical form of projection valued measures it is straightforward to prove

Lemma 3.4. *Let f, g be Borel functions and $\alpha, \beta \in \mathbb{C}$. Then we have*

$$\alpha P(f) + \beta P(g) \subseteq P(\alpha f + \beta g), \quad \mathfrak{D}(\alpha P(f) + \beta P(g)) = \mathfrak{D}_{|f|+|g|} \tag{3.32}$$

and

$$P(f)P(g) \subseteq P(f\,g), \quad \mathfrak{D}(P(f)P(g)) = \mathfrak{D}_g \cap \mathfrak{D}_{f\,g}. \tag{3.33}$$

Now observe, that to every projection valued measure P we can assign a self-adjoint operator $A = \int_{\mathbb{R}} \lambda dP(\lambda)$. The question is whether we can invert this map. To do this, we consider the resolvent $R_A(z) = \int_{\mathbb{R}} (\lambda - z)^{-1} dP(\lambda)$. By (3.12) the corresponding quadratic form is given by

$$F_\psi(z) = \langle \psi, R_A(z)\psi \rangle = \int_{\mathbb{R}} \frac{1}{\lambda - z} d\mu_\psi(\lambda), \tag{3.34}$$

which is know as the **Borel transform** of the measure μ_ψ. It can be shown (see Section 3.4) that $F_\psi(z)$ is a holomorphic map from the upper half plane to itself. Such functions are called **Herglotz functions**. Moreover, the measure μ_ψ can be reconstructed from $F_\psi(z)$ by **Stieltjes inversion formula**

$$\mu_\psi(\lambda) = \lim_{\delta \downarrow 0} \lim_{\varepsilon \downarrow 0} \frac{1}{\pi} \int_{-\infty}^{\lambda + \delta} \mathrm{Im}(F_\psi(\lambda + \mathrm{i}\varepsilon)) d\lambda. \tag{3.35}$$

Conversely, if $F_\psi(z)$ is a Herglotz function satisfying $|F(z)| \le \frac{M}{\mathrm{Im}(z)}$, then it is the Borel transform of a unique measure μ_ψ (given by Stieltjes inversion formula).

So let A be a given self-adjoint operator and consider the expectation of the resolvent of A,

$$F_\psi(z) = \langle \psi, R_A(z)\psi \rangle. \tag{3.36}$$

This function is holomorphic for $z \in \rho(A)$ and satisfies

$$F_\psi(z^*) = F_\psi(z)^* \quad \text{and} \quad |F_\psi(z)| \le \frac{\|\psi\|^2}{\mathrm{Im}(z)} \tag{3.37}$$

(see Theorem 2.14). Moreover, the first resolvent formula (2.69) shows

$$\mathrm{Im}(F_\psi(z)) = \mathrm{Im}(z) \|R_A(z)\psi\|^2 \tag{3.38}$$

that it maps the upper half plane to itself, that is, it is a Herglotz function. So by our above remarks, there is a corresponding measure $\mu_\psi(\lambda)$ given by Stieltjes inversion formula. It is called **spectral measure** corresponding to ψ.

More generally, by polarization, for each $\varphi, \psi \in \mathfrak{H}$ we can find a corresponding complex measure $\mu_{\varphi,\psi}$ such that

$$\langle \varphi, R_A(z)\psi \rangle = \int_{\mathbb{R}} \frac{1}{\lambda - z} d\mu_{\varphi,\psi}(\lambda). \tag{3.39}$$

The measure $\mu_{\varphi,\psi}$ is conjugate linear in φ and linear in ψ. Moreover, a comparison with our previous considerations begs us to define a family of operators $P_A(\Omega)$ via

$$\langle\varphi, P_A(\Omega)\psi\rangle = \int_{\mathbb{R}} \chi_\Omega(\lambda) d\mu_{\varphi,\psi}(\lambda). \tag{3.40}$$

This is indeed possible by the Riesz lemma since $|\langle\varphi, P_A(\Omega)\psi\rangle| = |\mu_{\varphi,\psi}(\Omega)| \leq \|\varphi\|\,\|\psi\|$. The operators $P_A(\Omega)$ are non negative ($0 \leq \langle\psi, P_A(\Omega)\psi\rangle \leq 1$) and hence self-adjoint.

Lemma 3.5. *The family of operators $P_A(\Omega)$ forms a projection valued measure.*

Proof. We first show $P_A(\Omega_1)P_A(\Omega_2) = P_A(\Omega_1 \cap \Omega_2)$ in two steps. First observe (using the first resolvent formula (2.69))

$$\begin{aligned}\int_{\mathbb{R}} \frac{1}{\lambda - \tilde{z}} d\mu_{R_A(z^*)\varphi,\psi}(\lambda) &= \langle R_A(z^*)\varphi, R_A(\tilde{z})\psi\rangle = \langle\varphi, R_A(z)R_A(\tilde{z})\psi\rangle \\ &= \frac{1}{z - \tilde{z}} \left(\langle\varphi, R_A(z)\psi\rangle - \langle\varphi, R_A(\tilde{z})\psi\rangle\right) \\ &= \frac{1}{z - \tilde{z}} \int_{\mathbb{R}} \left(\frac{1}{\lambda - z} - \frac{1}{\lambda - \tilde{z}}\right) d\mu_{\varphi,\psi}(\lambda) = \int_{\mathbb{R}} \frac{1}{\lambda - \tilde{z}} \frac{d\mu_{\varphi,\psi}(\lambda)}{\lambda - z}\end{aligned} \tag{3.41}$$

implying $d\mu_{R_A(z^*)\varphi,\psi}(\lambda) = (\lambda - z)^{-1} d\mu_{\varphi,\psi}(\lambda)$ since a Herglotz function is uniquely determined by its measure. Secondly we compute

$$\begin{aligned}\int_{\mathbb{R}} \frac{1}{\lambda - z} d\mu_{\varphi,P_A(\Omega)\psi}(\lambda) &= \langle\varphi, R_A(z)P_A(\Omega)\psi\rangle = \langle R_A(z^*)\varphi, P_A(\Omega)\psi\rangle \\ &= \int_{\mathbb{R}} \chi_\Omega(\lambda) d\mu_{R_A(z^*)\varphi,\psi}(\lambda) = \int_{\mathbb{R}} \frac{1}{\lambda - z} \chi_\Omega(\lambda) d\mu_{\varphi,\psi}(\lambda)\end{aligned}$$

implying $d\mu_{\varphi,P_A(\Omega)\psi}(\lambda) = \chi_\Omega(\lambda) d\mu_{\varphi,\psi}(\lambda)$. Equivalently we have

$$\langle\varphi, P_A(\Omega_1)P_A(\Omega_2)\psi\rangle = \langle\varphi, P_A(\Omega_1 \cap \Omega_2)\psi\rangle \tag{3.42}$$

since $\chi_{\Omega_1}\chi_{\Omega_2} = \chi_{\Omega_1 \cap \Omega_2}$. In particular, choosing $\Omega_1 = \Omega_2$, we see that $P_A(\Omega_1)$ is a projector.

The relation $P_A(\mathbb{R}) = \mathbb{I}$ follows from (3.86) below and Lemma 2.18 which imply $\mu_\psi(\mathbb{R}) = \|\psi\|^2$.

Now let $\Omega = \bigcup_{n=1}^\infty \Omega_n$ with $\Omega_n \cap \Omega_m = \emptyset$ for $n \neq m$. Then

$$\sum_{j=1}^n \langle\psi, P_A(\Omega_j)\psi\rangle \to \langle\psi, P_A(\Omega)\psi\rangle \tag{3.43}$$

since $\mu_\psi(\Omega) = \sum_{j=1}^\infty \mu_\psi(\Omega_j)$. Furthermore, using

$$P_A(\Omega) - \sum_{j=1}^n P_A(\Omega_j) = P_A(\Omega \backslash \bigcup_{j=1}^n \Omega_j) \tag{3.44}$$

and

$$\langle \psi, P_A(\Omega \backslash \bigcup_{j=1}^{n} \Omega_j)\psi\rangle = \|P_A(\Omega \backslash \bigcup_{j=1}^{n} \Omega_j)\psi\|^2 \qquad (3.45)$$

we see that $\sum_{j=1}^{n} P_A(\Omega_j)\psi \to P_A(\Omega)\psi$ for any $\psi \in \mathfrak{H}$. □

Now we can prove the **spectral theorem** for self-adjoint operators.

Theorem 3.6 (Spectral theorem)**.** *To every self-adjoint operator A there corresponds a unique projection valued measure P_A such that*

$$A = \int_{\mathbb{R}} \lambda dP_A(\lambda). \qquad (3.46)$$

Proof. Existence has already been established. Moreover, Lemma 3.4 shows that $P_A((\lambda - z)^{-1}) = R_A(z)$, $z \in \mathbb{C}\backslash\mathbb{R}$. Since the measures $\mu_{\varphi,\psi}$ are uniquely determined by the resolvent and the projection valued measure is uniquely determined by the measures $\mu_{\varphi,\psi}$ we are done. □

The quadratic form of A is given by

$$q_A(\psi) = \int_{\mathbb{R}} \lambda d\mu_\psi(\lambda) \qquad (3.47)$$

and can be defined for every ψ in the **form domain** self-adjoint operator

$$\mathfrak{Q}(A) = \{\psi \in \mathfrak{H} | \int_{\mathbb{R}} |\lambda| d\mu_\psi(\lambda) < \infty\}. \qquad (3.48)$$

This extends our previous definition for non-negative operators.

Note, that if A and $\tilde{A}$ are unitarily equivalent as in (3.25), then $UR_A(z) = R_{\tilde{A}}(z)U$ and hence

$$d\mu_\psi = d\tilde{\mu}_{U\psi}. \qquad (3.49)$$

In particular, we have $UP_A(f) = P_{\tilde{A}}(f)U$, $U\mathfrak{D}(P_A(f)) = \mathfrak{D}(P_{\tilde{A}}(f))$.

Finally, let us give a characterization of the spectrum of A in terms of the associated projectors.

Theorem 3.7. *The spectrum of A is given by*

$$\sigma(A) = \{\lambda \in \mathbb{R} | P_A((\lambda - \varepsilon, \lambda + \varepsilon)) \neq 0 \text{ for all } \varepsilon > 0\}. \qquad (3.50)$$

Proof. Let $\Omega_n = (\lambda_0 - \frac{1}{n}, \lambda_0 + \frac{1}{n})$. Suppose $P_A(\Omega_n) \neq 0$. Then we can find a $\psi_n \in P_A(\Omega_n)\mathfrak{H}$ with $\|\psi_n\| = 1$. Since

$$\begin{aligned} \|(A - \lambda_0)\psi_n\|^2 &= \|(A - \lambda_0)P_A(\Omega_n)\psi_n\|^2 \\ &= \int_{\mathbb{R}} (\lambda - \lambda_0)^2 \chi_{\Omega_n}(\lambda) d\mu_{\psi_n}(\lambda) \leq \frac{1}{n^2} \end{aligned} \qquad (3.51)$$

we conclude $\lambda_0 \in \sigma(A)$ by Lemma 2.12.

Conversely, if $P_A((\lambda_0-\varepsilon,\lambda_0+\varepsilon))=0$, set $f_\varepsilon(\lambda)=\chi_{\mathbb{R}\backslash(\lambda_0-\varepsilon,\lambda_0+\varepsilon)}(\lambda)(\lambda-\lambda_0)^{-1}$. Then

$$\begin{aligned}(A-\lambda_0)P_A(f_\varepsilon) = P_A(f_\varepsilon)(A-\lambda_0) &= \\ P_A(f_\varepsilon(\lambda)(\lambda-\lambda_0)) = P_A(\mathbb{R}\backslash(\lambda_0-\varepsilon,\lambda_0+\varepsilon)) &= \mathbb{I} \end{aligned} \tag{3.52}$$

and hence $\lambda_0\in\rho(A)$. □

Thus $P_A((\lambda_1,\lambda_2))=0$ if and only if $(\lambda_1,\lambda_2)\subseteq\rho(A)$ and we have

$$P_A(\sigma(A))=\mathbb{I} \qquad \text{and} \qquad P_A(\mathbb{R}\cap\rho(A))=0 \tag{3.53}$$

and consequently

$$P_A(f)=P_A(\sigma(A))P_A(f)=P_A(\chi_{\sigma(A)}f). \tag{3.54}$$

In other words, $P_A(f)$ is not affected by the values of f on $\mathbb{R}\backslash\sigma(A)$!

It is clearly more intuitive to write $P_A(f)=f(A)$ and we will do so from now on. This notation is justified by the elementary observation

$$P_A(\sum_{j=0}^n \alpha_j\lambda^j)=\sum_{j=0}^n \alpha_j A^j. \tag{3.55}$$

Moreover, this also shows that if A is bounded and $f(A)$ can be defined via a convergent power series, then this agrees with our present definition by Theorem 3.1.

3.2. More on Borel measures

Section 3.1 showed that in order to understand self-adjoint operators, one needs to understand multiplication operators on $L^2(\mathbb{R},d\mu)$, where $d\mu$ is a finite Borel measure. This is the purpose of the present section.

The set of all **growth points**, that is,

$$\sigma(\mu)=\{\lambda\in\mathbb{R}|\mu((\lambda-\varepsilon,\lambda+\varepsilon))>0 \text{ for all } \varepsilon>0\}, \tag{3.56}$$

is called the spectrum of μ. Invoking Morea's together with Fubini's theorem shows that the Borel transform

$$F(z)=\int_{\mathbb{R}}\frac{1}{\lambda-z}d\mu(\lambda) \tag{3.57}$$

is holomorphic for $z\in\mathbb{C}\backslash\sigma(\mu)$. The converse following from Stieltjes inversion formula. Associated with this measure is the operator

$$Af(\lambda)=\lambda f(\lambda), \qquad \mathfrak{D}(A)=\{f\in L^2(\mathbb{R},d\mu)|\lambda f(\lambda)\in L^2(\mathbb{R},d\mu)\}. \tag{3.58}$$

By Theorem 3.7 the spectrum of A is precisely the spectrum of μ, that is,

$$\sigma(A)=\sigma(\mu). \tag{3.59}$$

Note that $1 \in L^2(\mathbb{R}, d\mu)$ is a cyclic vector for A and that

$$d\mu_{g,f}(\lambda) = g(\lambda)^* f(\lambda) d\mu(\lambda). \tag{3.60}$$

Now what can we say about the function $f(A)$ (which is precisely the multiplication operator by f) of A? We are only interested in the case where f is real valued. Introduce the measure

$$(f_\star\mu)(\Omega) = \mu(f^{-1}(\Omega)), \tag{3.61}$$

then

$$\int_{\mathbb{R}} g(\lambda) d(f_\star\mu)(\lambda) = \int_{\mathbb{R}} g(f(\lambda)) d\mu(\lambda). \tag{3.62}$$

In fact, it suffices to check this formula for simple functions g which follows since $\chi_\Omega \circ f = \chi_{f^{-1}(\Omega)}$. In particular, we have

$$P_{f(A)}(\Omega) = \chi_{f^{-1}(\Omega)}. \tag{3.63}$$

It is tempting to conjecture that $f(A)$ is unitarily equivalent to multiplication by λ in $L^2(\mathbb{R}, d(f_\star\mu))$ via the map

$$L^2(\mathbb{R}, d(f_\star\mu)) \to L^2(\mathbb{R}, d\mu), \quad g \mapsto g \circ f. \tag{3.64}$$

However, this map is only unitary if its range is $L^2(\mathbb{R}, d\mu)$, which is equivalent to 1 being also cyclic for $f(A)$.

For example, let $f(\lambda) = \lambda^2$, then $(g \circ f)(\lambda) = g(\lambda^2)$ and the range of the above map is given by the symmetric functions. Note that we can still get a unitary map $L^2(\mathbb{R}, d(f_\star\mu)) \oplus L^2(\mathbb{R}, \chi d(f_\star\mu)) \to L^2(\mathbb{R}, d\mu)$, $(g_1, g_2) \mapsto g_1(\lambda^2) + g_2(\lambda^2)(\chi(\lambda) - \chi(-\lambda))$, where $\chi = \chi_{(0,\infty)}$.

Lemma 3.8. *Let f be real valued. The spectrum of $f(A)$ is given by*

$$\sigma(f(A)) = \sigma(f_\star\mu). \tag{3.65}$$

In particular,

$$\sigma(f(A)) \subseteq \overline{f(\sigma(A))}, \tag{3.66}$$

where equality holds if f is continuous and the closure can be dropped if, in addition, $\sigma(A)$ is bounded (i.e., compact).

Proof. If $\lambda_0 \in \sigma(f_\star\mu)$, then the sequence $g_n = \mu(\Omega_n)^{-1/2}\chi_{\Omega_n}$, $\Omega_n = f^{-1}((\lambda_0 - \frac{1}{n}, \lambda_0 + \frac{1}{n}))$, satisfies $\|g_n\| = 1$, $\|(f(A) - \lambda_0)g_n\| < n^{-1}$ and hence $\lambda_0 \in \sigma(f(A))$. Conversely, if $\lambda_0 \notin \sigma(f_\star\mu)$, then $\mu(\Omega_n) = 0$ for some n and hence we can change f on Ω_n such that $f(\mathbb{R}) \cap (\lambda_0 - \frac{1}{n}, \lambda_0 + \frac{1}{n}) = \emptyset$ without changing the corresponding operator. Thus $(f(A) - \lambda_0)^{-1} = (f(\lambda) - \lambda_0)^{-1}$ exists and is bounded, implying $\lambda_0 \notin \sigma(f(A))$.

If f is continuous, $f^{-1}(f(\lambda) - \varepsilon, f(\lambda) + \varepsilon)$ contains an open interval around λ and hence $f(\lambda) \in \sigma(f(A))$ if $\lambda \in \sigma(A)$. If, in addition, $\sigma(A)$ is compact, then $f(\sigma(A))$ is compact and hence closed. □

Next we recall the unique decomposition of μ with respect to Lebesgue measure,

$$d\mu = d\mu_{ac} + d\mu_s, \tag{3.67}$$

where μ_{ac} is **absolutely continuous** with respect to Lebesgue measure (i.e., we have $\mu_{ac}(B) = 0$ for all B with Lebesgue measure zero) and μ_s is **singular** with respect to Lebesgue measure (i.e., μ_s is supported, $\mu_s(\mathbb{R}\backslash B) = 0$, on a set B with Lebesgue measure zero). The singular part μ_s can be further decomposed into a **(singularly) continuous** and a **pure point** part,

$$d\mu_s = d\mu_{sc} + d\mu_{pp}, \tag{3.68}$$

where μ_{sc} is continuous on $\mathbb{R}$ and μ_{pp} is a step function. Since the measures $d\mu_{ac}$, $d\mu_{sc}$, and $d\mu_{pp}$ are mutually singular, they have mutually disjoint supports M_{ac}, M_{sc}, and M_{pp}. Note that these sets are *not* unique. We will choose them such that M_{pp} is the set of all jumps of $\mu(\lambda)$ and such that M_{sc} has Lebesgue measure zero.

To the sets M_{ac}, M_{sc}, and M_{pp} correspond projectors $P^{ac} = \chi_{M_{ac}}(A)$, $P^{sc} = \chi_{M_{sc}}(A)$, and $P^{pp} = \chi_{M_{pp}}(A)$ satisfying $P^{ac} + P^{sc} + P^{pp} = \mathbb{I}$. In other words, we have a corresponding direct sum decomposition of both our Hilbert space

$$L^2(\mathbb{R}, d\mu) = L^2(\mathbb{R}, d\mu_{ac}) \oplus L^2(\mathbb{R}, d\mu_{sc}) \oplus L^2(\mathbb{R}, d\mu_{pp}) \tag{3.69}$$

and our operator A

$$A = (AP^{ac}) \oplus (AP^{sc}) \oplus (AP^{pp}). \tag{3.70}$$

3.3. Spectral types

Our next aim is to transfer the results of the previous section to arbitrary self-adjoint operators using Lemma 3.3. First note that to each spectral basis $\{\psi_n\}$ we can assign a trace measure

$$d\mu = \sum_n \varepsilon_n d\mu_{\psi_n}, \qquad 0 < \varepsilon_n \le 1, \sum_n \varepsilon_n = 1. \tag{3.71}$$

Then, we have $\sigma(A) = \sigma(\mu)$ and the following generalization of Lemma 3.8 holds.

Theorem 3.9 (Spectral mapping)**.** *Let μ be the trace measure of some spectral basis and let f be real-valued. Then the spectrum of $f(A)$ is given by*

$$\sigma(f(A)) = \{\lambda \in \mathbb{R} | \mu(f^{-1}(\lambda - \varepsilon, \lambda + \varepsilon)) > 0 \text{ for all } \varepsilon > 0\}. \tag{3.72}$$

In particular,

$$\sigma(f(A)) \subseteq \overline{f(\sigma(A))}, \tag{3.73}$$

where equality holds if f is continuous and the closure can be dropped if, in addition, $\sigma(A)$ is bounded.

Next, we want to introduce the splitting (3.69) for arbitrary self-adjoint operators A. It is tempting to pick a spectral basis and treat each summand in the direct sum separately. However, since it is not clear that this approach is independent of the spectral basis chosen, we use the more sophisticated definition

$$\begin{aligned}\mathfrak{H}_{ac} &= \{\psi \in \mathfrak{H} | \mu_\psi \text{ is absolutely continuous}\},\\ \mathfrak{H}_{sc} &= \{\psi \in \mathfrak{H} | \mu_\psi \text{ is singularly continuous}\},\\ \mathfrak{H}_{pp} &= \{\psi \in \mathfrak{H} | \mu_\psi \text{ is pure point}\}.\end{aligned} \tag{3.74}$$

Lemma 3.10. *We have*

$$\mathfrak{H} = \mathfrak{H}_{ac} \oplus \mathfrak{H}_{sc} \oplus \mathfrak{H}_{pp}. \tag{3.75}$$

There are Borel sets M_{xx} such that the projector onto $\mathfrak{H}_{xx}$ is given by $P^{xx} = \chi_{M_{xx}}(A)$, $xx \in \{ac, sc, pp\}$. In particular, the subspaces $\mathfrak{H}_{xx}$ reduce A. For the sets M_{xx} one can choose the corresponding supports of some trace measure μ.

Proof. We will use the unitary operator U of Lemma 3.3. Pick $\varphi \in \mathfrak{H}$ and write $\varphi = \sum_n \varphi_n$ with $\varphi_n \in \mathfrak{H}_{\psi_n}$. Let $f_n = U\varphi_n$, then, by construction of the unitary operator U, $\varphi_n = f_n(A)\psi_n$ and hence $d\mu_{\varphi_n} = |f_n|^2 d\mu_{\psi_n}$. Moreover, since the subspaces $\mathfrak{H}_{\psi_n}$ are orthogonal, we have

$$d\mu_\varphi = \sum_n |f_n|^2 d\mu_{\psi_n} \tag{3.76}$$

and hence

$$d\mu_{\varphi,xx} = \sum_n |f_n|^2 d\mu_{\psi_n,xx}, \qquad xx \in \{ac, sc, pp\}. \tag{3.77}$$

This shows

$$U\mathfrak{H}_{xx} = \bigoplus_n L^2(\mathbb{R}, d\mu_{\psi_n,xx}), \qquad xx \in \{ac, sc, pp\} \tag{3.78}$$

and reduces our problem to the considerations of the previous section. □

The **absolutely continuous**, **singularly continuous**, and **pure point spectrum** of A are defined as

$$\sigma_{ac}(A) = \sigma(A|_{\mathfrak{H}_{ac}}), \quad \sigma_{sc}(A) = \sigma(A|_{\mathfrak{H}_{sc}}), \quad \text{and} \quad \sigma_{pp}(A) = \sigma(A|_{\mathfrak{H}_{pp}}), \tag{3.79}$$

respectively. If A and $\tilde{A}$ are unitarily equivalent via U, then so are $A|_{\mathfrak{H}_{xx}}$ and $\tilde{A}|_{\tilde{\mathfrak{H}}_{xx}}$ by (3.49). In particular, $\sigma_{xx}(A) = \sigma_{xx}(\tilde{A})$.

It is important to observe that $\sigma_{pp}(A)$ is in general not equal to the set of eigenvalues

$$\sigma_p(A) = \{\lambda \in \mathbb{R} | \lambda \text{ is an eigenvalue of } A\} \tag{3.80}$$

since we only have $\sigma_{pp}(A) = \overline{\sigma_p(A)}$.

3.4. Appendix: The Herglotz theorem

A holomorphic function $F : \mathbb{C}_+ \to \mathbb{C}_+$, $\mathbb{C}_\pm = \{z \in \mathbb{C} | \pm\mathrm{Im}(z) > 0\}$, is called a Herglotz function. We can define F on $\mathbb{C}_-$ using $F(z^*) = F(z)^*$.

Suppose μ is a finite Borel measure. Then its Borel transform is defined via

$$F(z) = \int_{\mathbb{R}} \frac{d\mu(\lambda)}{\lambda - z}. \tag{3.81}$$

Theorem 3.11. *The Borel transform of a finite Borel measure is a Herglotz function satisfying*

$$|F(z)| \le \frac{\mu(\mathbb{R})}{\mathrm{Im}(z)}, \qquad z \in \mathbb{C}_+. \tag{3.82}$$

Moreover, the measure μ can be reconstructed via Stieltjes inversion formula

$$\frac{1}{2}\left(\mu((\lambda_1,\lambda_2)) + \mu([\lambda_1,\lambda_2])\right) = \lim_{\varepsilon\downarrow 0} \frac{1}{\pi} \int_{\lambda_1}^{\lambda_2} \mathrm{Im}(F(\lambda + \mathrm{i}\varepsilon)) d\lambda. \tag{3.83}$$

Proof. By Morea's and Fubini's theorem, F is holomorphic on $\mathbb{C}_+$ and the remaining properties follow from $0 < \mathrm{Im}((\lambda - z)^{-1})$ and $|\lambda - z|^{-1} \le \mathrm{Im}(z)^{-1}$. Stieltjes inversion formula follows from Fubini's theorem and the dominated convergence theorem since

$$\frac{1}{2\pi \mathrm{i}} \int_{\lambda_1}^{\lambda_2} (\frac{1}{x - \lambda - \mathrm{i}\varepsilon} - \frac{1}{x - \lambda + \mathrm{i}\varepsilon}) d\lambda \to \frac{1}{2}\left(\chi_{[\lambda_1,\lambda_2]}(x) + \chi_{(\lambda_1,\lambda_2)}(x)\right) \tag{3.84}$$

pointwise. □

Observe

$$\mathrm{Im}(F(z)) = \mathrm{Im}(z) \int_{\mathbb{R}} \frac{d\mu(\lambda)}{|\lambda - z|^2} \tag{3.85}$$

and

$$\lim_{\lambda\to\infty} \lambda\, \mathrm{Im}(F(\mathrm{i}\lambda)) = \mu(\mathbb{R}). \tag{3.86}$$

The converse of the previous theorem is also true

Theorem 3.12. *Suppose F is a Herglotz function satisfying*

$$|F(z)| \le \frac{M}{\mathrm{Im}(z)}, \qquad z \in \mathbb{C}_+. \tag{3.87}$$

Then there is a unique Borel measure μ, satisfying $\mu(\mathbb{R}) \le M$, such that F is the Borel transform of μ.

Proof. We abbreviate $F(z) = v(z) + \mathrm{i}\, w(z)$ and $z = x + \mathrm{i}\, y$. Next we choose a contour

$$\Gamma = \{x + \mathrm{i}\varepsilon + \lambda | \lambda \in [-R, R]\} \cup \{x + \mathrm{i}\varepsilon + R\mathrm{e}^{\mathrm{i}\varphi} | \varphi \in [0, \pi]\}. \tag{3.88}$$

and note that z lies inside Γ and $z^* + 2\mathrm{i}\varepsilon$ lies outside Γ if $0 < \varepsilon < y < R$. Hence we have by Cauchy's formula

$$F(z) = \frac{1}{2\pi\mathrm{i}} \int_\Gamma \left(\frac{1}{\zeta - z} - \frac{1}{\zeta - z^* - 2\mathrm{i}\varepsilon} \right) F(\zeta) d\zeta. \tag{3.89}$$

Inserting the explicit form of Γ we see

$$\begin{aligned} F(z) &= \frac{1}{\pi} \int_{-R}^{R} \frac{y - \varepsilon}{\lambda^2 + (y - \varepsilon)^2} F(x + \mathrm{i}\varepsilon + \lambda) d\lambda \\ &\quad + \frac{\mathrm{i}}{\pi} \int_0^\pi \frac{y - \varepsilon}{R^2 \mathrm{e}^{2\mathrm{i}\varphi} + (y - \varepsilon)^2} F(x + \mathrm{i}\varepsilon + R\mathrm{e}^{\mathrm{i}\varphi}) R\mathrm{e}^{\mathrm{i}\varphi} d\varphi. \end{aligned} \tag{3.90}$$

The integral over the semi circle vanishes as $R \to \infty$ and hence we obtain

$$F(z) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{y - \varepsilon}{(\lambda - x)^2 + (y - \varepsilon)^2} F(\lambda + \mathrm{i}\varepsilon) d\lambda \tag{3.91}$$

and taking imaginary parts

$$w(z) = \int_{\mathbb{R}} \phi_\varepsilon(\lambda) w_\varepsilon(\lambda) d\lambda, \tag{3.92}$$

where $\phi_\varepsilon(\lambda) = (y - \varepsilon)/((\lambda - x)^2 + (y - \varepsilon)^2)$ and $w_\varepsilon(\lambda) = w(\lambda + \mathrm{i}\varepsilon)/\pi$. Letting $y \to \infty$ we infer from our bound

$$\int_{\mathbb{R}} w_\varepsilon(\lambda) d\lambda \le M. \tag{3.93}$$

In particular, since $|\phi_\varepsilon(\lambda) - \phi_0(\lambda)| \le const\, \varepsilon$ we have

$$w(z) = \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}} \phi_0(\lambda) d\mu_\varepsilon(\lambda), \tag{3.94}$$

where $\mu_\varepsilon(\lambda) = \int_{-\infty}^{\lambda} w_\varepsilon(x) dx$. It remains to establish that the monotone functions μ_ε converge properly. Since $0 \le \mu_\varepsilon(\lambda) \le M$, there is a convergent subsequence for fixed λ. Moreover, by the standard diagonal trick, there is even a subsequence ε_n such that $\mu_{\varepsilon_n}(\lambda)$ converges for each rational λ. For irrational λ we set $\mu(\lambda_0) = \inf_{\lambda \ge \lambda_0} \{\mu(\lambda) | \lambda \text{ rational}\}$. Then $\mu(\lambda)$ is monotone, $0 \le \mu(\lambda_1) \le \mu(\lambda_2) \le M$, $\lambda_1 \le \lambda_2$, and we claim

$$w(z) = \int_{\mathbb{R}} \phi_0(\lambda) d\mu(\lambda). \tag{3.95}$$

Fix $\delta > 0$ and let $\lambda_1 < \lambda_2 < \cdots < \lambda_{m+1}$ be rational numbers such that

$$|\lambda_{j+1} - \lambda_j| \le \delta \quad \text{and} \quad \lambda_1 \le x - \frac{\delta}{y^3},\ \lambda_{m+1} \ge x + \frac{\delta}{y^3}. \tag{3.96}$$

Then

$$|\phi_0(\lambda) - \phi_0(\lambda_j)| \le \frac{\delta}{y^2}, \qquad \lambda_j \le \lambda \le \lambda_{j+1}, \tag{3.97}$$

and

$$|\phi_0(\lambda)| \le \frac{\delta}{y^2}, \qquad \lambda \le \lambda_1 \text{ or } \lambda_{m+1} \le \lambda. \tag{3.98}$$

Now observe

$$\begin{aligned} | \int_{\mathbb{R}} \phi_0(\lambda) d\mu(\lambda) - \int_{\mathbb{R}} \phi_0(\lambda) d\mu_{\varepsilon_n}(\lambda)| \le \\ | \int_{\mathbb{R}} \phi_0(\lambda) d\mu(\lambda) - \sum_{j=1}^m \phi_0(\lambda_j)(\mu(\lambda_{j+1}) - \mu(\lambda_j))| \\ + | \sum_{j=1}^m \phi_0(\lambda_j)(\mu(\lambda_{j+1}) - \mu(\lambda_j) - \mu_{\varepsilon_n}(\lambda_{j+1}) + \mu_{\varepsilon_n}(\lambda_j)) \\ + | \int_{\mathbb{R}} \phi_0(\lambda) d\mu_{\varepsilon_n}(\lambda) - \sum_{j=1}^m \phi_0(\lambda_j)(\mu_{\varepsilon_n}(\lambda_{j+1}) - \mu_{\varepsilon_n}(\lambda_j))| \end{aligned} \tag{3.99}$$

The first and third term can be bounded by $2M\delta/y^2$. Moreover, since $\phi_0(y) \le 1/y$ we can find an $N \in \mathbb{N}$ such that

$$|\mu(\lambda_j) - \mu_{\varepsilon_n}(\lambda_j)| \le \frac{y}{2m}\delta, \qquad n \ge N, \tag{3.100}$$

and hence the second term is bounded by δ. In summary, the difference in (3.99) can be made arbitrarily small.

Now $F(z)$ and $\int_{\mathbb{R}}(\lambda - z)^{-1} d\mu(\lambda)$ have the same imaginary part and thus the only differ by a real constant. By our bound this constant must be zero. □

The Radon-Nikodym derivative of μ can be obtained from the boundary values of F.

Theorem 3.13. *Let μ be a finite Borel measure and F its Borel transform, then*

$$(\underline{D}\mu)(\lambda) \le \liminf_{\varepsilon\downarrow 0} \frac{1}{\pi} F(\lambda + \mathrm{i}\varepsilon) \le \limsup_{\varepsilon\downarrow 0} \frac{1}{\pi} F(\lambda + \mathrm{i}\varepsilon) \le (\overline{D}\mu)(\lambda). \tag{3.101}$$

Proof. We need to estimate

$$\mathrm{Im}(F(\lambda + \mathrm{i}\varepsilon)) = \int_{\mathbb{R}} K_\varepsilon(t) d\mu(t), \qquad K_\varepsilon(t) = \frac{\varepsilon}{t^2 + \varepsilon^2}. \tag{3.102}$$

We first split the integral into two parts

$$\mathrm{Im}(F(\lambda+\mathrm{i}\varepsilon)) = \int_{I_\delta} K_\varepsilon(t-\lambda) d\mu(t) + \int_{\mathbb{R}\setminus I_\delta} K_\varepsilon(t-\lambda)\mu(t), \quad I_\delta = (\lambda-\delta, \lambda+\delta). \tag{3.103}$$

Clearly the second part can be estimated by

$$\int_{\mathbb{R}\backslash I_\delta} K_\varepsilon(t-\lambda)\mu(t) \le K_\varepsilon(\delta)\mu(\mathbb{R}). \tag{3.104}$$

To estimate the first part we integrate

$$K'_\varepsilon(s)\, ds\, d\mu(t) \tag{3.105}$$

over the triangle $\{(s,t)|\lambda - s < t < \lambda + s,\, 0 < s < \delta\} = \{(s,t)|\lambda - \delta < t < \lambda + \delta,\, t - \lambda < s < \delta\}$ and obtain

$$\int_0^\delta \mu(I_s) K'_\varepsilon(s) ds = \int_{I_\delta} (K(\delta) - K_\varepsilon(t-\lambda)) d\mu(t). \tag{3.106}$$

Now suppose there is are constants c and C such that $c \le \frac{\mu(I_s)}{2s} \le C$, $0 \le s \le \delta$, then

$$2c \arctan(\frac{\delta}{\varepsilon}) \le \int_{I_\delta} K_\varepsilon(t-\lambda) d\mu(t) \le 2C \arctan(\frac{\delta}{\varepsilon}) \tag{3.107}$$

since

$$\delta K_\varepsilon(\delta) + \int_0^\delta -s K'_\varepsilon(s) ds = \arctan(\frac{\delta}{\varepsilon}). \tag{3.108}$$

Thus the claim follows combining both estimates. □

As a consequence of Theorem 0.10 and Theorem 0.11 we obtain

Theorem 3.14. *Let μ be a finite Borel measure and F its Borel transform, then the limit*

$$\mathrm{Im}(F(\lambda)) = \lim_{\varepsilon \downarrow 0} \frac{1}{\pi} \mathrm{Im}(F(\lambda + \mathrm{i}\varepsilon)) \tag{3.109}$$

exists a.e. with respect to μ (finite or infinite) and

$$(D\mu)(\lambda) = \frac{1}{\pi} \mathrm{Im}(F(\lambda)) \tag{3.110}$$

whenever the latter is finite.

Moreover, the set $\{\lambda | F(\lambda) = \infty\}$ is a support for the singularly and $\{\lambda | F(\lambda) < \infty\}$ is a support for the absolutely continuous part.

In particular,

Corollary 3.15. *The measure μ is purely absolutely continuous on I if $\limsup_{\varepsilon\downarrow 0} \mathrm{Im}(F(\lambda + \mathrm{i}\varepsilon)) < \infty$ for all $\lambda \in I$.*

Chapter 4

Applications of the spectral theorem

Now let us show how the spectral theorem can be used. This chapter can be skipped on first reading. We will give a few typical applications:

Firstly we will derive an operator valued version of of Stieltjes' inversion formula. To do this, we need to show how to integrate a family of functions of A with respect to a parameter. Moreover, we will show that these integrals can be evaluated by computing the corresponding integrals of the complex valued functions.

Secondly we will consider commuting operators and show how certain facts, which are known to hold for the resolvent of an operator A, can be established for a larger class of functions.

Finally, we will show how the dimension of $\mathrm{Ran} P_A(\Omega)$ can be estimated.

4.1. Integral formulas

We begin with the first task by having a closer look at the projector $P_A(\Omega)$. They project onto subspaces corresponding to expectation values in the set Ω. In particular, the number

$$\langle\psi, \chi_\Omega(A)\psi\rangle \tag{4.1}$$

is the probability for a measurement of a to lie in Ω. In addition, we have

$$\langle\psi, A\psi\rangle = \int_\Omega \lambda\, d\mu_\psi(\lambda) \in \mathrm{hull}(\Omega), \quad \psi \in P_A(\Omega)\mathfrak{H},\ \|\psi\| = 1, \tag{4.2}$$

where $\mathrm{hull}(\Omega)$ is the convex hull of Ω.

The space $\mathrm{Ran}\chi_{\{\lambda_0\}}(A)$ is called the **eigenspace** corresponding to λ_0 since we have

$$\langle\varphi, A\psi\rangle = \int_{\mathbb{R}} \lambda\, \chi_{\{\lambda_0\}}(\lambda) d\mu_{\varphi,\psi}(\lambda) = \lambda_0 \int_{\mathbb{R}} d\mu_{\varphi,\psi}(\lambda) = \lambda_0\langle\varphi, \psi\rangle \tag{4.3}$$

and hence $A\psi = \lambda_0\psi$ for all $\psi \in \mathrm{Ran}\chi_{\{\lambda_0\}}(A)$. The dimension of the eigenspace is called the **multiplicity** of the eigenvalue.

Moreover, since

$$\lim_{\varepsilon\downarrow 0} \frac{-\mathrm{i}\varepsilon}{\lambda - \lambda_0 - \mathrm{i}\varepsilon} = \chi_{\{\lambda_0\}}(\lambda) \tag{4.4}$$

we infer from Theorem 3.1 that

$$\lim_{\varepsilon\downarrow 0} -\mathrm{i}\varepsilon R_A(\lambda_0 + \mathrm{i}\varepsilon)\psi = \chi_{\{\lambda_0\}}(A)\psi. \tag{4.5}$$

Similarly, we can obtain an operator valued version of Stieltjes' inversion formula. But first we need to recall a few facts from integration in Banach spaces.

We will consider the case of mappings $f : I \to X$ where $I = [t_0, t_1] \subset \mathbb{R}$ is a compact interval and X is a Banach space. As before, a function $f : I \to X$ is called simple if the image of f is finite, $f(I) = \{x_i\}_{i=1}^n$, and if each inverse image $f^{-1}(x_i)$, $1 \le i \le n$, is a Borel set. The set of simple functions $S(I, X)$ forms a linear space and can be equipped with the sup norm. The corresponding Banach space obtained after completion is called the set of regulated functions $R(I, X)$.

Observe that $C(I, X) \subset R(I, X)$. In fact, consider the simple function $f_n = \sum_{i=0}^{n-1} f(s_i)\chi_{[s_i,s_{i+1})}$, where $s_i = t_0 + i\frac{t_1-t_0}{n}$. Since $f \in C(I, X)$ is uniformly continuous, we infer that f_n converges uniformly to f.

For $f \in S(I, X)$ we can define a linear map $\int : S(I, X) \to X$ by

$$\int_I f(t)dt = \sum_{i=1}^n x_i |f^{-1}(x_i)|, \tag{4.6}$$

where $|\Omega|$ denotes the Lebesgue measure of Ω. This map satisfies

$$\|\int_I f(t)dt\| \le \|f\|(t_1 - t_0) \tag{4.7}$$

and hence it can be extended uniquely to a linear map $\int : R(I, X) \to X$ with the same norm $(t_1 - t_0)$. We even have

$$\|\int_I f(t)dt\| \le \int_I \|f(t)\|dt, \tag{4.8}$$

which clearly holds for $f \in S(I,X)$ und thus for all $f \in R(I,X)$ by continuity. In addition, if $\ell \in X^*$ is a continuous linear functional, then

$$\ell(\int_I f(t)dt) = \int_I \ell(f(t))dt, \qquad f \in R(I,X). \tag{4.9}$$

If $I = \mathbb{R}$, we say that $f : I \to X$ is integrable if $f \in R([-r,r],X)$ for all $r > 0$ and if $\|f(t)\|$ is integrable. In this case we can set

$$\int_{\mathbb{R}} f(t)dt = \lim_{r\to\infty} \int_{[-r,r]} f(t)dt \tag{4.10}$$

and (4.8) and (4.9) still hold.

We will use the standard notation $\int_{t_2}^{t_3} f(s)ds = \int_I \chi_{(t_2,t_3)}(s)f(s)ds$ and $\int_{t_3}^{t_2} f(s)ds = -\int_{t_2}^{t_3} f(s)ds$.

We write $f \in C^1(I,X)$ if

$$\frac{d}{dt}f(t) = \lim_{\varepsilon\to 0} \frac{f(t+\varepsilon) - f(t)}{\varepsilon} \tag{4.11}$$

exists for all $t \in I$. In particular, if $f \in C(I,X)$, then $F(t) = \int_{t_0}^t f(s)ds \in C^1(I,X)$ and $dF/dt = f$ as can be seen from

$$|F(t+\varepsilon) - F(t) - f(t)\varepsilon| = |\int_t^{t+\varepsilon} (f(s) - f(t))ds| \le |\varepsilon| \sup_{s\in[t,t+\varepsilon]} |f(s) - f(t)|. \tag{4.12}$$

The important facts for us are the following two results.

Lemma 4.1. *Suppose $f : I \times \mathbb{R} \to \mathbb{C}$ is a bounded Borel function such that $f(.,\lambda)$ is integrable for every λ and set $F(\lambda) = \int_I f(t,\lambda)dt$. For any self-adjoint operator A we have*

$$F(A) = \int_I f(t,A)dt \quad \textit{respectively} \quad F(A)\psi = \int_I f(t,A)\psi\, dt. \tag{4.13}$$

Moreover, suppose $A : I \to \mathfrak{L}(\mathfrak{H})$ is integrable, then

$$\left(\int_I A(t)dt\right)\psi = \int_I (A(t)\psi)dt. \tag{4.14}$$

Proof. We compute

$$\begin{aligned}\langle\varphi,(\int_I f(t,A)dt)\psi\rangle &= \int_I \langle\varphi, f(t,A)\psi\rangle dt\\ &= \int_I\int_{\mathbb{R}} f(t,\lambda)d\mu_{\varphi,\psi}(\lambda)dt\\ &= \int_{\mathbb{R}}\int_I f(t,\lambda)dt\, d\mu_{\varphi,\psi}(\lambda)\\ &= \int_{\mathbb{R}} F(\lambda)d\mu_{\varphi,\psi}(\lambda) = \langle\varphi, F(A)\psi\rangle \end{aligned} \qquad (4.15)$$

by Fubini's theorem and hence the first claim follows. The remaining claims are similar. □

Lemma 4.2. *Suppose $F:\mathbb{R}\to\mathfrak{L}(\mathfrak{H})$ is integrable and $A\in\mathfrak{L}(\mathfrak{H})$. Then*

$$A\int_{\mathbb{R}} f(t)dt = \int_{\mathbb{R}} Af(t)dt \quad \textit{respectively} \quad \int_{\mathbb{R}} f(t)dtA = \int_{\mathbb{R}} f(t)Adt. \qquad (4.16)$$

Proof. It suffices to prove the case where f is simple and of compact support. But for such functions the claim is straightforward. □

Now we can prove Stone's formula.

Theorem 4.3 (Stone's formula)**.** *Let A be self-adjoint, then*

$$\frac{1}{2\pi\mathrm{i}}\int_{\lambda_1}^{\lambda_2}(R_A(\lambda+\mathrm{i}\varepsilon)-R_A(\lambda-\mathrm{i}\varepsilon))d\lambda \to \frac{1}{2}\left(P_A([\lambda_1,\lambda_2])+P_A((\lambda_1,\lambda_2))\right) \qquad (4.17)$$

strongly.

Proof. The result follows easily combining the previous lemma with Theorem 3.1 and (3.84). □

Let Γ be a differentiable Jordan curve in $\rho(A)$. Then the following integral

$$\int_\Gamma R_A(z)dz \qquad (4.18)$$

is well-defined by our above analysis. Furthermore, it only depends on the homotopy class of Γ since

$$\langle\varphi,(\int_\Gamma R_A(z)dz)\psi\rangle = \int_\Gamma\langle\varphi, R_A(z)\psi\rangle dz \qquad (4.19)$$

and since $\langle \varphi, R_A(z)\psi\rangle$ is holomorphic in $\rho(A)$. Moreover, the result can be easily computed

$$\begin{aligned}\int_\Gamma \langle \varphi, R_A(z)\psi\rangle dz &= \int_\Gamma \int_{\mathbb{R}} \frac{1}{\lambda - z} d\mu_{\varphi,\psi}(\lambda) dz \\ &= \int_{\mathbb{R}} \int_\Gamma \frac{1}{\lambda - z} dz\, d\mu_{\varphi,\psi}(\lambda) \\ &= \int_{\mathbb{R}} \chi_\Omega(z) d\mu_{\varphi,\psi}(\lambda), \end{aligned} \tag{4.20}$$

where Ω is the intersection of the interior of Γ with $\mathbb{R}$. Hence

$$\chi_\Omega(A) = \int_\Gamma R_A(z) dz. \tag{4.21}$$

4.2. Commuting operators

Now we come to commuting operators. We first recall the Stone-Weierstrass theorem.

Theorem 4.4 (Stone-Weierstrass)**.** *Let X be a compact Hausdorff space and let B be a subalgabra of $C(X,\mathbb{R})$ which separates points, that is for every $x, y \in X$ there is an $f \in B$ with $f(x) \neq f(y)$. Then the closure of B is either $C(X,\mathbb{R})$ of $\{f \in C(X,\mathbb{R}) | f(x_0) = 0\}$ for some $x_0 \in X$.*

We are interested in the complex case $C(X,\mathbb{C})$, where the same conclusion holds if B is a $*$-subalgebra (i.e., closed under conjugation), which follows easily by applying the Stone-Weierstrass to the real and imaginary part.

As a preparation we can now prove

Lemma 4.5. *Let $K \subseteq \mathbb{R}$ be closed. And let $C_\infty(K)$ be the set of all continuous functions on K which vanish at ∞ (if K is unbounded) with the sup norm. The $*$-algebra generated by the function*

$$\lambda \mapsto \frac{1}{\lambda - z} \tag{4.22}$$

for one $z \in \mathbb{C}\backslash K$ is dense in $C_\infty(K)$.

Proof. If K is compact, the claim follows directly from the complex Stone-Weierstrass theorem since $(\lambda_1 - z)^{-1} = (\lambda_2 - z)^{-1}$ implies $\lambda_1 = \lambda_2$. Otherwise, replace K by $\tilde{K} = K \cup \{\infty\}$, which is compact, and set $(\infty - z)^{-1} = 0$. Then we can again apply the complex Stone-Weierstrass theorem to conclude that our $*$-subalgebra is equal to $\{f \in C(\tilde{K}) | f(\infty) = 0\}$ which is equivalent to $C_\infty(K)$. □

We say that two bounded operators A, B **commute** if

$$[A, B] = AB - BA = 0. \tag{4.23}$$

If A or B is unbounded, we soon run into trouble with this definition since the above expression might not even make sense for any nonzero vector (e.g., take $B = \langle\varphi, .\rangle\psi$ with $\psi \notin \mathfrak{D}(A)$). To avoid this nuisance we will replace A by a bounded function of A. A good candidate is the resolvent. Hence if A is self-adjoint and B is bounded we will say that A and B commute if

$$[R_A(z), B] = [R_A(z^*), B] = 0 \tag{4.24}$$

for one $z \in \rho(A)$.

Lemma 4.6. *Suppose A is self-adjoint and commutes with the bounded operator B. Then*

$$[f(A), B] = 0 \tag{4.25}$$

for any bounded Borel function f. If f is unbounded, the claim holds for any $\psi \in \mathfrak{D}(f(A))$.

Proof. Equation (4.24) tell us that (4.25) holds for any f in the $*$-subalgebra generated by $R_A(z)$. Since this subalgebra is dense in $C_\infty(\sigma(A))$, the claim follows for all such $f \in C_\infty(\sigma(A))$. Next fix $\psi \in \mathfrak{H}$ and let f be bounded. Choose a sequence $f_n \in C_\infty(\sigma(A))$ converging to f in $L^2(\mathbb{R}, d\mu_\psi)$. Then

$$Bf(A)\psi = \lim_{n\to\infty} Bf_n(A)\psi = \lim_{n\to\infty} f_n(A)B\psi = f(A)B\psi. \tag{4.26}$$

If f is unbounded, let $\psi \in \mathfrak{D}(f(A))$ and choose f_n as in (3.19). Then

$$f(A)B\psi = \lim_{n\to\infty} f_n(A)B\psi = \lim_{n\to\infty} Bf_n(A)\psi \tag{4.27}$$

shows $f \in L^2(\mathbb{R}, d\mu_{B\psi})$ (i.e., $B\psi \in \mathfrak{D}(f(A))$) and $f(A)B\psi = BF(A)\psi$. □

Corollary 4.7. *If A is self-adjoint and bounded, then (4.24) holds if and only if (4.23) holds.*

Proof. Since $\sigma(A)$ is compact, we have $\lambda \in C_\infty(\sigma(A))$ and hence (4.23) follows from (4.25) by our lemma. Conversely, since B commutes with any polynomial of A, the claim follows from the Neumann series. □

As another consequence we obtain

Theorem 4.8. *Suppose A is self-adjoint and has simple spectrum. A bounded operator B commutes with A if and only if $B = f(A)$ for some bounded Borel function.*

Proof. Let ψ be a cyclic vector for A. By our unitary equivalence it is no restriction to assume $\mathfrak{H} = L^2(\mathbb{R}, d\mu_\psi)$. Then

$$Bg(\lambda) = Bg(\lambda) \cdot 1 = g(\lambda)(B1)(\lambda) \tag{4.28}$$

since B commutes with the multiplication operator $g(\lambda)$. Hence B is multiplication by $f(\lambda) = (B1)(\lambda)$. □

The assumption that the spectrum of A is simple is crucial as the example $A = \mathbb{I}$ shows. Note also that the functions $\exp(-\mathrm{i}tA)$ can also be used instead of resolvents.

Lemma 4.9. *Suppose A is self-adjoint and B is bounded. Then B commutes with A if and only if*

$$[\mathrm{e}^{-\mathrm{i}At}, B] = 0 \tag{4.29}$$

for all $t \in \mathbb{R}$.

Proof. It suffices to show $[\hat{f}(A), B] = 0$ for $f \in \mathcal{S}(\mathbb{R})$, since these functions are dense in $C_\infty(\mathbb{R})$ by the complex Stone-Weierstrass theorem. Here $\hat{f}$ denotes the Fourier transform of f, see Section 6.1. But for such f we have

$$[\hat{f}(A), B] = \frac{1}{\sqrt{2\pi}}[\int_{\mathbb{R}} f(t)\mathrm{e}^{-\mathrm{i}At}dt, B] = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(t)[\mathrm{e}^{-\mathrm{i}At}, B]dt = 0 \tag{4.30}$$

by Lemma 4.2. □

The extension to the case where B is self-adjoint and unbounded is straightforward. We say that A and B commute in this case if

$$[R_A(z_1), R_B(z_2)] = [R_A(z_1^*), R_B(z_2)] = 0 \tag{4.31}$$

for one $z_1 \in \rho(A)$ and one $z_2 \in \rho(B)$ (the claim for z_2^* follows by taking adjoints). From our above analysis it follows that this is equivalent to

$$[\mathrm{e}^{-\mathrm{i}At}, \mathrm{e}^{-\mathrm{i}Bs}] = 0, \quad t, s \in \mathbb{R}, \tag{4.32}$$

respectively

$$[f(A), g(B)] = 0 \tag{4.33}$$

for arbitrary bounded Borel functions f and g.

4.3. The min-max theorem

In many applications a self-adjoint operator has a number of eigenvalues below the bottom of the essential spectrum. The essential spectrum is obtained from the spectrum by removing all discrete eigenvalues with finite multiplicity (we will have a closer look at it in Section 8.2). In general there is no way of computing the lowest eigenvalues and their corresponding eigenfunctions explicitly. However, one often has some idea how the eigenfunctions might approximately look like.

So suppose we have a normalized function ψ_1 which is an approximation for the eigenfunction φ_1 of the lowest eigenvalue E_1. Then by Theorem 2.15 we know that

$$\langle\psi_1, A\psi_1\rangle \geq \langle\varphi_1, A\varphi_1\rangle = E_1. \tag{4.34}$$

If we add some free parameters to ψ_1, one can optimize them and obtain quite good upper bounds for the first eigenvalue.

But is there also something one can say about the next eigenvalues? Suppose we know the first eigenfunction φ_1, then we can restrict A to the orthogonal complement of φ_1 and proceed as before: E_2 will be the infimum over all expectations restricted to this subspace. If we restrict to the orthogonal complement of an approximating eigenfunction ψ_1, there will still be a component in the direction of φ_1 left and hence the infimum of the expectations will be lower than E_2. Thus the optimal choice $\psi_1 = \varphi_1$ will give the maximal value E_2.

More precisely, let $\{\varphi_j\}_{j=1}^N$ be an orthonormal basis for the space spanned by the eigenfunctions corresponding to eigenvalues below the essential spectrum. Assume they satisfy $(A - E_j)\varphi_j = 0$, where $E_j \leq E_{j+1}$ are the eigenvalues (counted according to their multiplicity). If the number of eigenvalues N is finite we set $E_j = \inf\sigma_{ess}(A)$ for $j > N$ and choose φ_j such that $\|(A - E_j)\varphi_j = \| \leq \varepsilon$.

Define

$$U(\psi_1, \dots, \psi_n) = \{\psi \in \mathfrak{D}(A) |\, \|\psi\| = 1,\ \psi \in \operatorname{span}\{\psi_1, \dots, \psi_n\}^\perp\}. \tag{4.35}$$

(i) We have

$$\inf_{\psi\in U(\psi_1,\dots,\psi_n)} \langle\psi, A\psi\rangle \leq E_n + O(\varepsilon). \tag{4.36}$$

In fact, set $\psi = \sum_{j=1}^n \alpha_j\varphi_j$ and choose α_j such that $\psi \in U(\psi_1, \dots, \psi_{n-1})$, then

$$\langle\psi, A\psi\rangle = \sum_{j=1}^n |\alpha_j|^2 E_j + O(\varepsilon) \leq E_n + O(\varepsilon) \tag{4.37}$$

and the claim follows.

(ii) We have

$$\inf_{\psi\in U(\psi_1,\dots,\psi_n)} \langle\psi, A\psi\rangle \geq E_n + O(\varepsilon). \tag{4.38}$$

In fact, set $\psi = \varphi_n$ and proceed as before.

Since ε can be chosen arbitrarily small we have proven

Theorem 4.10 (Min-Max). *Let A be self-adjoint and let $E_1 \leq E_2 \leq E_3 \cdots$ be the eigenvalues of A below the essential spectrum respectively the infimum* *of the essential spectrum once there are no more eigenvalues left. Then*

$$E_n = \sup_{\psi_1,\dots,\psi_{n-1}} \inf_{\psi\in U(\psi_1,\dots,\psi_{n-1})} \langle\psi, A\psi\rangle. \tag{4.39}$$

Clearly the same result holds if $\mathfrak{D}(A)$ is replaced by the quadratic form domain $\mathfrak{Q}(A)$ in the definition of U. In addition, as long as E_n is an eigenvalue, the sup and inf are in fact max and min, explaining the name.

Corollary 4.11. *Suppose A and B are self-adjoint operators with $A \ge B$ (i.e. $A - B \ge 0$), then $E_n(A) \ge E_n(B)$.*

4.4. Estimating eigenspaces

Next, we show that the dimension of the range of $P_A(\Omega)$ can be estimated if we have some functions which lie approximately in this space.

Theorem 4.12. *Suppose A is a bounded self-adjoint operator and ψ_j, $1 \le j \le k$, are linearly independent elements of a $\mathfrak{H}$.*
(i). Let $\lambda \in \mathbb{R}$, $\psi_j \in \mathfrak{Q}(A)$. If

$$\langle\psi, A\psi\rangle < \lambda\|\psi\|^2 \tag{4.40}$$

for any nonzero linear combination $\psi = \sum_{j=1}^k c_j\psi_j$, then

$$\dim \operatorname{Ran} P_A((-\infty,\lambda)) \ge k. \tag{4.41}$$

Similarly, $\langle\psi, A\psi\rangle > \lambda\|\psi\|^2$ implies $\dim \operatorname{Ran} P_A((\lambda,\infty)) \ge k$.
(ii). Let $\lambda_1 < \lambda_2$, $\psi_j \in \mathfrak{D}(A)$. If

$$\|(A - \frac{\lambda_2+\lambda_1}{2})\psi\| < \frac{\lambda_2-\lambda_1}{2}\|\psi\| \tag{4.42}$$

for any nonzero linear combination $\psi = \sum_{j=1}^k c_j\psi_j$, then

$$\dim \operatorname{Ran} P_A((\lambda_1,\lambda_2)) \ge k. \tag{4.43}$$

Proof. (i). Let $M = \operatorname{span}\{\psi_j\} \subseteq \mathfrak{H}$. We claim $\dim P_A((-\infty,\lambda))M = \dim M = k$. For this it suffices to show $\operatorname{Ker} P_A((-\infty,\lambda))|_M = \{0\}$. Suppose $P_A((-\infty,\lambda))\psi = 0$, $\psi \ne 0$. Then we see that for any nonzero linear combination ψ

$$\begin{aligned}\langle\psi, A\psi\rangle &= \int_{\mathbb{R}} \eta\, d\mu_\psi(\eta) = \int_{[\lambda,\infty)} \eta\, d\mu_\psi(\eta)\\ &\ge \lambda\int_{[\lambda,\infty)} d\mu_\psi(\eta) = \lambda\|\psi\|^2.\end{aligned} \tag{4.44}$$

This contradicts our assumption (4.40).
(ii). Using the same notation as before we need to show $\operatorname{Ker} P_A((\lambda_1,\lambda_2))|_M =$

$\{0\}$. If $P_A((\lambda_1,\lambda_2))\psi = 0$, $\psi \neq 0$, then,

$$\|(A - \frac{\lambda_2+\lambda_1}{2})\psi\|^2 = \int_{\mathbb{R}} (x - \frac{\lambda_2+\lambda_1}{2})^2 d\mu_\psi(x) = \int_\Omega x^2 d\mu_\psi(x + \frac{\lambda_2+\lambda_1}{2})$$
$$\geq \frac{(\lambda_2-\lambda_1)^2}{4} \int_\Omega d\mu_\psi(x + \frac{\lambda_2+\lambda_1}{2}) = \frac{(\lambda_2-\lambda_1)^2}{4}\|\psi\|^2, \tag{4.45}$$

where $\Omega = (-\infty, -(\lambda_2-\lambda_1)/2] \cup [(\lambda_2-\lambda_1)/2, \infty)$. But this is a contradiction as before. □

4.5. Tensor products of operators

Suppose A_j, $1 \leq j \leq n$, are self-adjoint operators on $\mathfrak{H}_j$. For every monomial $\lambda_1^{n_1} \cdots \lambda_n^{n_n}$ we can define

$$(A_1^{n_1} \otimes \cdots \otimes A_n^{n_n})\psi_1 \otimes \cdots \otimes \psi_n = (A_1^{n_1}\psi_1) \otimes \cdots \otimes (A_n^{n_n}\psi_n), \quad \psi_j \in \mathfrak{D}(A_j^{n_j}). \tag{4.46}$$

Hence for every polynomial $P(\lambda_1, \dots, \lambda_n)$ of degree N we can define

$$P(A_1, \dots, A_n)\psi_1 \otimes \cdots \otimes \psi_n, \quad \psi_j \in \mathfrak{D}(A_j^N), \tag{4.47}$$

and extend this definition to obtain a linear operator on the set

$$\mathfrak{D} = \operatorname{span}\{\psi_1 \otimes \cdots \otimes \psi_n \,|\, \psi_j \in \mathfrak{D}(A_j^N)\}. \tag{4.48}$$

Moreover, if P is real-valued, then the operator $P(A_1, \dots, A_n)$ on $\mathfrak{D}$ is symmetric and we can consider its closure, which will again be denoted by $P(A_1, \dots, A_n)$.

Theorem 4.13. *Suppose A_j, $1 \leq j \leq n$, are self-adjoint operators on $\mathfrak{H}_j$ and let $P(\lambda_1, \dots, \lambda_n)$ be a real-valued polynomial and define $P(A_1, \dots, A_n)$ as above.*

Then $P(A_1, \dots, A_n)$ is self-adjoint and its spectrum is the closure of the range of P on the product of the spectra of the A_j, that is,

$$\sigma(P(A_1, \dots, A_n)) = \overline{P(\sigma(A_1), \dots, \sigma(A_n))}. \tag{4.49}$$

Proof. By the spectral theorem it is no restriction to assume that A_j is multiplication by λ_j on $L^2(\mathbb{R}, d\mu_j)$ and $P(A_1, \dots, A_n)$ is hence multiplication by $P(\lambda_1, \dots, \lambda_n)$ on $L^2(\mathbb{R}^n, d\mu_1 \times \cdots \times d\mu_n)$. Since $\mathfrak{D}$ contains the set of all functions $\psi_1(\lambda_1)\cdots\psi_n(\lambda_n)$ for which $\psi_j \in L^2_0(\mathbb{R}, d\mu_j)$ it follows that the domain of the closure of P contains $L^2_0(\mathbb{R}^n, d\mu_1 \times \cdots \times d\mu_n)$. Hence P is the maximally defined multiplication operator by $P(\lambda_1, \dots, \lambda_n)$, which is self-adjoint.

Now let $\lambda = P(\lambda_1, \dots, \lambda_n)$ with $\lambda_j \in \sigma(A_j)$. Then there exists Weyl sequences $\psi_{j,k} \in \mathfrak{D}(A_j^N)$ with $(A_j - \lambda_j)\psi_{j,k} \to 0$ as $k \to \infty$. Then, $(P - \lambda)\psi_k \to 0$, where $\psi_k = \psi_{1,k} \otimes \cdots \otimes \psi_{1,k}$ and hence $\lambda \in \sigma(P)$. Conversely, if $\lambda \notin \overline{P(\sigma(A_1), \dots, \sigma(A_n))}$, then $|P(\lambda_1, \dots, \lambda_n) - \lambda| \geq \varepsilon$ for a.e. λ_j with respect to μ_j and hence $(P-\lambda)^{-1}$ exists and is bounded, that is $\lambda \in \rho(P)$. □

The two main cases of interest are $A_1 \otimes A_2$, in which case

$$\sigma(A_1 \otimes A_2) = \overline{\sigma(A_1)\sigma(A_2)} = \overline{\{\lambda_1\lambda_2 | \lambda_j \in \sigma(A_j)\}}, \tag{4.50}$$

and $A_1 \otimes \mathbb{I} + \mathbb{I} \otimes A_2$, in which case

$$\sigma(A_1 \otimes \mathbb{I} + \mathbb{I} \otimes A_2) = \overline{\sigma(A_1) + \sigma(A_2)} = \overline{\{\lambda_1 + \lambda_2 | \lambda_j \in \sigma(A_j)\}}. \tag{4.51}$$

Chapter 5

Quantum dynamics

As in the finite dimensional case, the solution the Schrödinger equation

$$\mathrm{i}\frac{d}{dt}\psi(t) = H\psi(t) \tag{5.1}$$

is given by

$$\psi(t) = \exp(-\mathrm{i}tH)\psi(0). \tag{5.2}$$

A detailed investigation of this formula will be our first task. Moreover, in the finite dimensional case the dynamics is understood once the eigenvalues are known and the same is true in our case once we know the spectrum. Note that, like any Hamiltonian system from classical mechanics, our system is not hyperbolic (i.e., the spectrum is not away from the real axis) and hence simple results like, all solutions tend to the equilibrium position cannot be expected.

5.1. The time evolution and Stone's theorem

In this section we want to have a look at the initial value problem associated with the Schrödinger equation (2.11) in the Hilbert space $\mathfrak{H}$. If $\mathfrak{H}$ is one-dimensional (and hence A is a real number), the solution is given by

$$\psi(t) = \mathrm{e}^{-\mathrm{i}tA}\psi(0). \tag{5.3}$$

Our hope is that this formula also applies in the general case and that we can reconstruct a one-parameter unitary group $U(t)$ from its generator A (compare (2.10)) via $U(t) = \exp(-\mathrm{i}tA)$. We first investigate the family of operators $\exp(-\mathrm{i}tA)$.

Theorem 5.1. *Let A be self-adjoint and let $U(t) = \exp(-\mathrm{i}tA)$.*

(i). $U(t)$ is a strongly continuous one-parameter unitary group.

(ii). The limit $\lim_{t\to 0}\frac{1}{t}(U(t)\psi-\psi)$ *exists if and only if* $\psi\in\mathfrak{D}(A)$ *in which case* $\lim_{t\to 0}\frac{1}{t}(U(t)\psi-\psi)=-\mathrm{i}A\psi$.

(iii). $U(t)\mathfrak{D}(A)=\mathfrak{D}(A)$ *and* $AU(t)=U(t)A$.

Proof. The group property (i) follows directly from Theorem 3.1 and the corresponding statements for the function $\exp(-\mathrm{i}t\lambda)$. To prove strong continuity observe that

$$\begin{aligned}\lim_{t\to t_0}\|\mathrm{e}^{-\mathrm{i}tA}\psi-\mathrm{e}^{-\mathrm{i}t_0A}\psi\|^2 &= \lim_{t\to t_0}\int_{\mathbb{R}}|\mathrm{e}^{-\mathrm{i}t\lambda}-\mathrm{e}^{-\mathrm{i}t_0\lambda}|^2 d\mu_\psi(\lambda)\\ &= \int_{\mathbb{R}}\lim_{t\to t_0}|\mathrm{e}^{-\mathrm{i}t\lambda}-\mathrm{e}^{-\mathrm{i}t_0\lambda}|^2 d\mu_\psi(\lambda)=0 \quad (5.4)\end{aligned}$$

by the dominated convergence theorem.

Similarly, if $\psi\in\mathfrak{D}(A)$ we obtain

$$\lim_{t\to 0}\|\frac{1}{t}(\mathrm{e}^{-\mathrm{i}tA}\psi-\psi)+\mathrm{i}A\psi\|^2=\lim_{t\to 0}\int_{\mathbb{R}}|\frac{1}{t}(\mathrm{e}^{-\mathrm{i}t\lambda}-1)+\mathrm{i}\lambda|^2 d\mu_\psi(\lambda)=0 \quad (5.5)$$

since $|\mathrm{e}^{\mathrm{i}t\lambda}-1|\le|t\lambda|$. Now let $\tilde{A}$ be the generator defined as in (2.10). Then $\tilde{A}$ is a symmetric extension of A since we have

$$\langle\varphi,\tilde{A}\psi\rangle=\lim_{t\to 0}\langle\varphi,\frac{\mathrm{i}}{t}(U(t)-1)\psi\rangle=\lim_{t\to 0}\langle\frac{\mathrm{i}}{-t}(U(-t)-1)\varphi,\psi\rangle=\langle\tilde{A}\varphi,\psi\rangle \quad (5.6)$$

and hence $\tilde{A}=A$ by Corollary 2.2. This settles (ii).

To see (iii) replace $\psi\to U(s)\psi$ in (ii). □

For our original problem this implies that formula (5.3) is indeed the solution to the initial value problem of the Schrödinger equation. Moreover,

$$\langle U(t)\psi,AU(t)\psi\rangle=\langle U(t)\psi,U(t)A\psi\rangle=\langle\psi,A\psi\rangle \quad (5.7)$$

shows that the expectations of A are time independent. This corresponds to conservation of energy.

On the other hand, the generator of the time evolution of a quantum mechanical system should always be a self-adjoint operator since it corresponds to an observable (energy). Moreover, there should be a one to one correspondence between the unitary group and its generator. This is ensured by Stone's theorem.

Theorem 5.2 (Stone)**.** *Let* $U(t)$ *be a strongly continuous one-parameter unitary group. Then its generator* A *is self-adjoint and* $U(t)=\exp(-\mathrm{i}tA)$.

Proof. We first show that A is densely defined. Pick $\psi\in\mathfrak{H}$ and set

$$\psi_\tau=\int_0^\tau U(t)\psi dt \quad (5.8)$$

(the integral is defined as in Section 4.1) implying $\lim_{\tau\to 0}\tau^{-1}\psi_\tau = \psi$. Moreover,

$$\begin{aligned}\frac{1}{t}(U(t)\psi_\tau - \psi_\tau) &= \frac{1}{t}\int_t^{t+\tau} U(s)\psi ds - \frac{1}{t}\int_0^\tau U(s)\psi ds\\ &= \frac{1}{t}\int_\tau^{\tau+t} U(s)\psi ds - \frac{1}{t}\int_0^t U(s)\psi ds\\ &= \frac{1}{t}U(\tau)\int_0^t U(s)\psi ds - \frac{1}{t}\int_0^t U(s)\psi ds \to U(\tau)\psi + \psi \qquad (5.9)\end{aligned}$$

as $t\to 0$ shows $\psi_\tau \in \mathfrak{D}(A)$. As in the proof of the previous theorem, we can show that A is symmetric and that $U(t)\mathfrak{D}(A) = \mathfrak{D}(A)$.

Next, let us prove that A is essentially self-adjoint. By Lemma 2.6 it suffices to prove $\mathrm{Ker}(A^* - z^*) = \{0\}$ for $z \in \mathbb{C}\backslash\mathbb{R}$. Suppose $A^*\varphi = z^*\varphi$, then for each $\psi \in \mathfrak{D}(A)$ we have

$$\begin{aligned}\frac{d}{dt}\langle\varphi, U(t)\psi\rangle &= \langle\varphi, -\mathrm{i}AU(t)\psi\rangle = -\mathrm{i}\langle A^*\varphi, U(t)\psi\rangle\\ &= -\mathrm{i}z\langle\varphi, U(t)\psi\rangle \qquad (5.10)\end{aligned}$$

and hence $\langle\varphi, U(t)\psi\rangle = \exp(-\mathrm{i}zt)\langle\varphi,\psi\rangle$. Since the left hand side is bounded for all $t\in\mathbb{R}$ and the exponential on the right hand side is not, we must have $\langle\varphi,\psi\rangle = 0$ implying $\varphi = 0$ since $\mathfrak{D}(A)$ is dense.

So A is essentially self-adjoint and we can introduce $V(t) = \exp(-\mathrm{i}t\overline{A})$. We are done if we can show $U(t) = V(t)$.

Let $\psi \in \mathfrak{D}(A)$ and abbreviate $\psi(t) = (U(t) - V(t))\psi$. Then

$$\lim_{s\to 0}\frac{\psi(t+s) - \psi(t)}{s} = 0 \qquad (5.11)$$

and hence $\frac{d}{dt}\|\psi(t)\|^2 = 0$. Since $\psi(0) = 0$ we have $\psi(t) = 0$ and hence $U(t)$ and $V(t)$ coincide on $\mathfrak{D}(A)$. Furthermore, since $\mathfrak{D}(A)$ is dense we have $U(t) = V(t)$ by continuity. □

As an immediate consequence of the proof we also note the following useful criterion.

Corollary 5.3. *Suppose $\mathfrak{D} \subseteq \mathfrak{D}(A)$ is dense and invariant under $U(t)$. Then A is essentially self-adjoint on $\mathfrak{D}$.*

Proof. As in the above proof it follows $\langle\varphi,\psi\rangle = 0$ for any $\varphi \in \mathrm{Ker}(A^* - z^*)$ and $\psi \in \mathfrak{D}$. □

Note that by Lemma 4.9 two strongly continuous one-parameter groups commute

$$[\mathrm{e}^{-\mathrm{i}tA}, \mathrm{e}^{-\mathrm{i}sB}] = 0 \qquad (5.12)$$

if and only if the generators commute.

Clearly, for a physicist, one of the goals must be to understand the time evolution of a quantum mechanical system. We have seen that the time evolution is generated by a self-adjoint operator, the Hamiltonian, and is given by a linear first order differential equation, the Schrödinger equation. To understand the dynamics of such a first order differential equation, one must understand the spectrum of the generator. Some general tools for this endeavor will be provided in the following sections.

5.2. The RAGE theorem

Now, let us discuss why the decomposition of the spectrum introduced in Section 3.3 is of physical relevance. Let $\|\varphi\| = \|\psi\| = 1$. The vector $\langle\varphi,\psi\rangle\varphi$ is the projection of ψ onto the (one-dimensional) subspace spanned by φ. Hence $|\langle\varphi,\psi\rangle|^2$ can be viewed as the part of ψ which is in the state φ. A first question one might rise is, how does

$$|\langle\varphi, U(t)\psi\rangle|^2 \tag{5.13}$$

behave as $t \to \infty$? By the spectral theorem,

$$\hat{\mu}_{\varphi,\psi}(t) = \langle\varphi, U(t)\psi\rangle = \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}t\lambda} d\mu_{\varphi,\psi}(\lambda) \tag{5.14}$$

is the **Fourier transform** of the measure $\mu_{\varphi,\psi}$. Thus our question is answered by Wieners theorem.

Theorem 5.4 (Wiener)**.** *Let μ be a finite complex Borel measure on $\mathbb{R}$ and let*

$$\hat{\mu}(t) = \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}t\lambda} d\mu(\lambda) \tag{5.15}$$

be its Fourier transform. Then the Cesàro time average of $\hat{\mu}(t)$ has the following limit

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T |\hat{\mu}(t)|^2 dt = \sum_{\lambda\in\mathbb{R}} |\mu(\{\lambda\})|^2, \tag{5.16}$$

where the sum on the right hand side is finite.

Proof. By Fubini we have

$$\begin{aligned} \frac{1}{T}\int_0^T |\hat{\mu}(t)|^2 dt &= \frac{1}{T}\int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}(x-y)t} d\mu(x) d\mu^*(y) dt \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} \left(\frac{1}{T}\int_0^T \mathrm{e}^{-\mathrm{i}(x-y)t} dt\right) d\mu(x) d\mu^*(y). \end{aligned} \tag{5.17}$$

The function in parentheses is bounded by one and converges pointwise to $\chi_{\{0\}}(x-y)$ as $T \to \infty$. Thus, by the dominated convergence theorem, the limit of the above expression is given by

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \chi_{\{0\}}(x-y) d\mu(x) d\mu^*(y) = \int_{\mathbb{R}} \mu(\{y\}) d\mu^*(y) = \sum_{y \in \mathbb{R}} |\mu(\{y\})|^2. \quad (5.18)$$

□

To apply this result to our situation, observe that the subspaces $\mathfrak{H}_{ac}$, $\mathfrak{H}_{sc}$, and $\mathfrak{H}_{pp}$ are invariant with respect to time evolution since $P^{xx}U(t) = \chi_{M_{xx}}(H)\exp(-\mathrm{i}tH) = \exp(-\mathrm{i}tH)\chi_{M_{xx}}(H) = U(t)P^{xx}$, $xx \in \{ac, sc, pp\}$. Moreover, if $\psi \in \mathfrak{H}_{xx}$ we have $P^{xx}\psi = \psi$ which shows $\langle \varphi, f(A)\psi\rangle = \langle \varphi, P^{xx} f(A)\psi\rangle = \langle P^{xx}\varphi, f(A)\psi\rangle$ implying $d\mu_{\varphi,\psi} = d\mu_{P^{xx}\varphi,\psi}$. Thus if μ_ψ is *ac*, *sc*, or *pp*, so is $\mu_{\varphi,\psi}$ for every $\varphi \in \mathfrak{H}$.

That is, if $\psi \in \mathfrak{H}_c = \mathfrak{H}_{ac} \oplus \mathfrak{H}_{sc}$, then the Cesàro mean of $\langle \varphi, U(t)\psi\rangle$ tends to zero. In other words, the average of the probability of finding the system in any prescribed state tends to zero if we start in the continuous subspace $\mathfrak{H}_c$ of A.

If $\psi \in \mathfrak{H}_{ac}$, then $d\mu_{\varphi,\psi}$ is absolutely continuous with respect to Lebesgue measure and thus $\hat{\mu}_{\varphi,\psi}(t)$ is continuous and tends to zero as $|t| \to \infty$. In fact, this follows from the Riemann-Lebesgue lemma (see Lemma 6.5 below).

Now we want to draw some additional consequences from Wiener's theorem. This will eventually yield a dynamical characterization of the continuous and pure point spectrum due to Ruelle, Amrein, Gorgescu, and Enß. But first we need a few definitions.

An operator $K \in \mathfrak{L}(\mathfrak{H})$ is called **finite rank** if its range is finite dimensional. The dimension $n = \dim \mathrm{Ran}(K)$ is called the **rank** of K. If $\{\psi_j\}_{j=1}^n$ is an orthonormal basis for $\mathrm{Ran}(K)$ we have

$$K\psi = \sum_{j=1}^{n} \langle \psi_j, K\psi\rangle \psi_j = \sum_{j=1}^{n} \langle \varphi_j, \psi\rangle \psi_j, \quad (5.19)$$

where $\varphi_j = K^*\psi_j$. The elements φ_j are linearly independent since $\mathrm{Ran}(K) = \mathrm{Ker}(K^*)^\perp$. Hence every finite rank operator is of the form (5.19). In addition, the adjoint of K is also finite rank and given by

$$K^*\psi = \sum_{j=1}^{n} \langle \psi_j, \psi\rangle \varphi_j. \quad (5.20)$$

The closure of the set of all finite rank operators in $\mathfrak{L}(\mathfrak{H})$ is called the set of **compact operators** $\mathfrak{C}(\mathfrak{H})$.

Lemma 5.5. *The set of all compact operators $\mathfrak{C}(\mathfrak{H})$ is a closed $*$-ideal in $\mathfrak{L}(\mathfrak{H})$.*

Proof. The set of finite rank operators is clearly a $*$-ideal and hence so is its closure $\mathfrak{C}(\mathfrak{H})$. □

There is also a weaker version of compactness which is useful for us. The operator K is called **relatively compact** with respect to A if

$$KR_A(z) \in \mathfrak{C}(\mathfrak{H}) \tag{5.21}$$

for one $z \in \rho(A)$. By the first resolvent identity this then follows for all $z \in \rho(A)$. In particular we have $\mathfrak{D}(A) \subseteq \mathfrak{D}(K)$.

Now let us return to our original problem.

Theorem 5.6. *Let A be self-adjoint and suppose K is relatively compact. Then*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \|K\mathrm{e}^{-\mathrm{i}tA}P^c\psi\|^2 dt = 0 \qquad \textit{and} \qquad \lim_{t\to\infty} \|K\mathrm{e}^{-\mathrm{i}tA}P^{ac}\psi\| = 0 \tag{5.22}$$

for every $\psi \in \mathfrak{D}(A)$. In particular, if K is also bounded, then the result holds for any $\psi \in \mathfrak{H}$.

Proof. Let $\psi \in \mathfrak{H}_c$ respectively $\psi \in \mathfrak{H}_{ac}$ and drop the projectors. Then, if K is a rank one operator (i.e., $K = \langle\varphi_1, .\rangle\varphi_2$), the claim follows from Wiener's theorem respectively the Riemann-Lebesgue lemma. Hence it holds for any finite rank operator K.

If K is compact, there is a sequence K_n of finite rank operators such that $\|K - K_n\| \le 1/n$ and hence

$$\|K\mathrm{e}^{-\mathrm{i}tA}\psi\| \le \|K_n\mathrm{e}^{-\mathrm{i}tA}\psi\| + \frac{1}{n}\|\psi\|. \tag{5.23}$$

Thus the claim holds for any compact operator K.

If $\psi \in \mathfrak{D}(A)$ we can set $\psi = (A - \mathrm{i})^{-1}\varphi$, where $\varphi \in \mathfrak{H}_c$ if and only if $\psi \in \mathfrak{H}_c$ (since $\mathfrak{H}_c$ reduces A). Since $K(A + \mathrm{i})^{-1}$ is compact by assumption, the claim can be reduced to the previous situation. If, in addition, K is bounded, we can find a sequence $\psi_n \in \mathfrak{D}(A)$ such that $\|\psi - \psi_n\| \le 1/n$ and hence

$$\|K\mathrm{e}^{-\mathrm{i}tA}\psi\| \le \|K\mathrm{e}^{-\mathrm{i}tA}\psi_n\| + \frac{1}{n}\|K\|, \tag{5.24}$$

concluding the proof. □

With the help of this result we can now prove an abstract version of the RAGE theorem.

Theorem 5.7 (RAGE). *Let A be self-adjoint. Suppose $K_n \in \mathfrak{L}(\mathfrak{H})$ is a sequence of relatively compact operators which converges strongly to the identity. Then*

$$\begin{aligned}\mathfrak{H}_c &= \{\psi \in \mathfrak{H} | \lim_{n\to\infty} \lim_{T\to\infty} \frac{1}{T}\int_0^T \|K_n \mathrm{e}^{-\mathrm{i}tA}\psi\| dt = 0\},\\ \mathfrak{H}_{pp} &= \{\psi \in \mathfrak{H} | \lim_{n\to\infty} \sup_{t\geq 0} \|(\mathbb{I} - K_n)\mathrm{e}^{-\mathrm{i}tA}\psi\| = 0\}.\end{aligned} \tag{5.25}$$

Proof. Abbreviate $\psi(t) = \exp(-\mathrm{i}tA)\psi$. We begin with the first equation.

Let $\psi \in \mathfrak{H}_c$, then

$$\frac{1}{T}\int_0^T \|K_n\psi(t)\| dt \leq \left(\frac{1}{T}\int_0^T \|K_n\psi(t)\|^2 dt\right)^{1/2} \to 0 \tag{5.26}$$

by Cauchy-Schwarz and the previous theorem. Conversely, if $\psi \notin \mathfrak{H}_c$ we can write $\psi = \psi^c + \psi^{pp}$. By our previous estimate it suffices to show $\|K_n\psi^{pp}(t)\| \geq \varepsilon > 0$ for n large. In fact, we even claim

$$\lim_{n\to\infty} K_n\psi^{pp}(t) = \psi^{pp}(t) \tag{5.27}$$

uniformly in t. By the spectral theorem, we can write $\psi^{pp}(t) = \sum_j \alpha_j(t)\psi_j$, where the ψ_j are orthonormal eigenfunctions and $\alpha_j(t) = \exp(-\mathrm{i}t\lambda_j)\alpha_j$. Truncate this expansion after N terms, then this part converges to the desired limit by strong convergence of K_n. Moreover, by the uniform boundedness principle, we can find a positive constant M such that $\|K_n\| \leq M$ and hence the error can be made arbitrarily small by choosing N large.

Now let us turn to the second equation. If $\psi \in \mathfrak{H}_{pp}$ the claim follows by (5.27). Conversely, if $\psi \notin \mathfrak{H}_{pp}$ we can write $\psi = \psi^c + \psi^{pp}$ and by our previous estimate it suffices to show that $\|(\mathbb{I} - K_n)\psi^c(t)\|$ does not tend to 0 as $n \to \infty$. If it would, we would have

$$\begin{aligned} 0 &= \lim_{T\to\infty}\frac{1}{T}\int_0^T \|(\mathbb{I} - K_n)\psi^c(t)\|^2 dt\\ &\geq \|\psi^c(t)\|^2 - \lim_{T\to\infty}\frac{1}{T}\int_0^T \|K_n\psi^c(t)\|^2 dt = \|\psi^c(t)\|^2,\end{aligned} \tag{5.28}$$

a contradiction. □

In summary, regularity properties of spectral measures are related to the long time behavior of the corresponding quantum mechanical system. However, a more detailed investigation of this topic is beyond the scope of this manuscript. For a survey containing several recent results see [7].

It is often convenient to treat the observables as time dependent rather than the states. We set

$$K(t) = \mathrm{e}^{\mathrm{i}tA} K \mathrm{e}^{-\mathrm{i}tA} \tag{5.29}$$

and note

$$\langle\psi(t), K\psi(t)\rangle = \langle\psi, K(t)\psi\rangle, \qquad \psi(t) = \mathrm{e}^{-\mathrm{i}tA}\psi. \tag{5.30}$$

This point of view is often referred to as **Heisenberg picture** in the physics literature. If K is unbounded we will assume $\mathfrak{D}(A) \subseteq \mathfrak{D}(K)$ such that the above equations make sense at least for $\psi \in \mathfrak{D}(A)$. The main interest is the behavior of $K(t)$ for large t. The strong limits are called **asymptotic observables** if they exist.

Theorem 5.8. *Suppose A is self-adjoint and K is relatively compact. Then*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{e}^{\mathrm{i}tA} K \mathrm{e}^{-\mathrm{i}tA}\psi dt = \sum_{\lambda\in\sigma_p(A)} P_A(\{\lambda\}) K P_A(\{\lambda\})\psi, \quad \psi \in \mathfrak{D}(A). \tag{5.31}$$

If K is in addition bounded, the result holds for any $\psi \in \mathfrak{H}$.

Proof. We will assume that K is bounded. To obtain the general result, use the same trick as before and replace K by $KR_A(z)$. Write $\psi = \psi^c + \psi^{pp}$. Then

$$\lim_{T\to\infty} \frac{1}{T}\| \int_0^T K(t)\psi^c dt\| \le \lim_{T\to\infty} \frac{1}{T}\int_0^T \|K(t)\psi^c dt\| = 0 \tag{5.32}$$

by Theorem 5.6. As in the proof of the previous theorem we can write $\psi^{pp} = \sum_j \alpha_j \psi_j$ and hence

$$\sum_j \alpha_j \frac{1}{T}\int_0^T K(t)\psi_j dt = \sum_j \alpha_j \left(\frac{1}{T}\int_0^T \mathrm{e}^{\mathrm{i}t(A-\lambda_j)} dt\right) K\psi_j. \tag{5.33}$$

As in the proof of Wiener's theorem, we see that the operator in parenthesis tends to $P_A(\{\lambda_j\})$ strongly as $T \to \infty$. Since this operator is also bounded by 1 for all T, we can interchange the limit with the summation and the claim follows. □

We also note the following corollary.

Corollary 5.9. *Under the same assumptions as in the RAGE theorem we have*

$$\lim_{n\to\infty}\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{e}^{\mathrm{i}tA} K_n \mathrm{e}^{-\mathrm{i}tA}\psi dt = P^{pp}\psi \tag{5.34}$$

respectively

$$\lim_{n\to\infty}\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{e}^{-\mathrm{i}tA}(\mathbb{I} - K_n)\mathrm{e}^{-\mathrm{i}tA}\psi dt = P^c\psi. \tag{5.35}$$

Part 2

Schrödinger Operators

Chapter 6

The free Schrödinger operator

6.1. The Fourier transform

We first review some basic facts concerning the **Fourier transform** which will be needed in the following section.

Let $C^\infty(\mathbb{R}^n)$ be the set of all complex-valued functions which have partial derivatives of arbitrary order. For $f \in C^\infty(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}_0^n$ we set

$$\partial_\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}, \quad x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \quad |\alpha| = \alpha_1 + \cdots + \alpha_n. \tag{6.1}$$

An element $\alpha \in \mathbb{N}_0^n$ is called **multi-index** and $|\alpha|$ is called its **order**. Recall the Schwarz space

$$\mathcal{S}(\mathbb{R}^n) = \{f \in C^\infty(\mathbb{R}^n) | \sup_x |x^\alpha (\partial_\beta f)(x)| < \infty,\ \alpha, \beta \in \mathbb{N}_0^n\} \tag{6.2}$$

which is dense in $L^2(\mathbb{R}^n)$. For $f \in \mathcal{S}(\mathbb{R}^n)$ we define

$$\mathcal{F}(f)(p) \equiv \hat{f}(p) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}px} f(x) d^n x. \tag{6.3}$$

Then it is an exercise in partial integration to prove

Lemma 6.1. *For any multi-index $\alpha \in \mathbb{N}_0^n$ and any $f \in \mathcal{S}(\mathbb{R}^n)$ we have*

$$(\partial_\alpha f)^\wedge(p) = (\mathrm{i}p)^\alpha \hat{f}(p), \qquad (x^\alpha f(x))^\wedge(p) = \mathrm{i}^{|\alpha|} \partial_\alpha \hat{f}(p). \tag{6.4}$$

Hence we will sometimes write $pf(x)$ for $-\mathrm{i}\partial f(x)$, where $\partial = (\partial_1, \dots, \partial_n)$ is the **gradient**.

In particular $\mathcal{F}$ maps $\mathcal{S}(\mathbb{R}^n)$ into itself. Another useful property is the convolution formula.

Lemma 6.2. *The Fourier transform of the* **convolution**

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y) g(x-y) d^n y = \int_{\mathbb{R}^n} f(x-y) g(y) d^n y \tag{6.5}$$

of two functions $f, g \in \mathcal{S}(\mathbb{R}^n)$ is given by

$$(f * g)^\wedge(p) = (2\pi)^{n/2} \hat{f}(p) \hat{g}(p). \tag{6.6}$$

Proof. We compute

$$\begin{aligned} (f * g)^\wedge(p) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}px} \int_{\mathbb{R}^n} f(y) g(x-y) d^n y \, d^n x \\ &= \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}py} f(y) \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}p(x-y)} g(x-y) d^n x \, d^n y \\ &= \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}py} f(y) \hat{g}(p) d^n y = (2\pi)^{n/2} \hat{f}(p) \hat{g}(p), \end{aligned} \tag{6.7}$$

where we have used Fubini's theorem. □

Next, we want to compute the inverse of the Fourier transform. For this the following lemma will be needed.

Lemma 6.3. *We have $\mathrm{e}^{-zx^2/2} \in \mathcal{S}(\mathbb{R}^n)$ for $\mathrm{Re}(z) > 0$ and*

$$\mathcal{F}(\mathrm{e}^{-zx^2/2})(p) = \frac{1}{z^{n/2}} \mathrm{e}^{-p^2/(2z)}. \tag{6.8}$$

Here $z^{n/2}$ has to be understood as $(\sqrt{z})^n$, where the branch cut of the root is chosen along the negative real axis.

Proof. Due to product structure of the exponential, one can treat each coordinate separately, reducing the problem to the case $n = 1$.

Let $\phi_z(x) = \exp(-zx^2/2)$. Then $\phi_z'(x) + zx\phi_z(x) = 0$ and hence $\mathrm{i}(p\hat{\phi}_z(p) + z\hat{\phi}_z'(p)) = 0$. Thus $\hat{\phi}_z(p) = c\phi_{1/z}(p)$ and

$$c = \hat{\phi}_z(0) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp(-zx^2/2) dx = \frac{1}{\sqrt{z}} \tag{6.9}$$

at least for $z > 0$. However, since the integral is holomorphic for $\mathrm{Re}(z) > 0$, this holds for all z with $\mathrm{Re}(z) > 0$ if we choose the branch cut of the root along the negative real axis. □

Now we can show

Lemma 6.4. *The Fourier transform $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is a bijection. Its inverse is given by*

$$\mathcal{F}^{-1}(g)(x) \equiv \check{g}(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}px} g(p) d^n p. \tag{6.10}$$

We have $\mathcal{F}^2(f)(x) = f(-x)$ and thus $\mathcal{F}^4 = \mathbb{I}$.

Proof. It suffices to show $\mathcal{F}^2(f)(x) = f(-x)$. Let $\phi(x) = \exp(-x^2/2)$. Then, by the dominated convergence theorem we have

$$\mathcal{F}(\hat{f}) = \lim_{\varepsilon \downarrow 0} \mathcal{F}(\phi(\varepsilon p)\hat{f}(p)). \tag{6.11}$$

Using the convolution formula we see that the last term is equal to

$$\begin{aligned} &\lim_{\varepsilon \downarrow 0} \frac{1}{(2\pi\varepsilon^2)^{n/2}} \int_{\mathbb{R}^n} \phi(y/\varepsilon) f(y-x) d^n y \\ &= \lim_{\varepsilon \downarrow 0} \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \phi(y) f(\varepsilon y - x) d^n y \\ &= \frac{f(-x)}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \phi(y) d^n y = f(-x) \end{aligned} \tag{6.12}$$

concluding the proof. □

From Fubini's theorem it follows

$$\begin{aligned} \int_{\mathbb{R}^n} \hat{g}(p)^* f(p) d^n p &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} g(p)^* f(x) \mathrm{e}^{\mathrm{i}px} d^n p \, d^n x \\ &= \int_{\mathbb{R}^n} g(x)^* \hat{f}(x) d^n x \end{aligned} \tag{6.13}$$

and in particular, we see $\mathcal{F}^{-1} g = \mathcal{F}^* g$, $g \in \mathcal{S}(\mathbb{R}^n)$, which implies **Parseval's identity**

$$\int_{\mathbb{R}^n} |f(x)|^2 d^n x = \int_{\mathbb{R}^n} |\hat{f}(p)|^2 d^n p \tag{6.14}$$

for $f \in \mathcal{S}(\mathbb{R}^n)$. Moreover, $\mathcal{F}$ extends to a unitary operator $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$. Its spectrum satisfies

$$\sigma(\mathcal{F}) \subseteq \{z \in \mathbb{C} | z^4 = 1\}. \tag{6.15}$$

In fact, if ψ_n is a Weyl sequence, then $(\mathcal{F}^2 + z^2)(\mathcal{F} + z)(\mathcal{F} - z)\psi_n = (\mathcal{F}^4 - z^4)\psi_n = (1 - z^4)\psi_n \to 0$ implies $z^4 = 1$. We will show in (7.42) that equality holds.

Lemma 6.1 also allows us to extend differentiation to a larger class. Let us introduce the Sobolev space

$$H^r(\mathbb{R}^n) = \{f \in L^2(\mathbb{R}^n) | |p|^r \hat{f}(p) \in L^2(\mathbb{R}^n)\}. \tag{6.16}$$

We will abbreviate

$$\partial_\alpha f = ((\mathrm{i}p)^\alpha \hat{f}(p))^\vee, \qquad f \in H^r(\mathbb{R}^n),\ |\alpha| \le r \tag{6.17}$$

which implies

$$\int_{\mathbb{R}^n} g(x)(\partial_\alpha f)(x) d^n x = (-1)^\alpha \int_{\mathbb{R}^n} (\partial_\alpha g)(x) f(x) d^n x, \tag{6.18}$$

for $f \in H^r(\mathbb{R}^n)$ and $g \in \mathcal{S}(\mathbb{R}^n)$. That is, $\partial_\alpha f$ is the derivative of f in the sense of distributions.

Finally, we have the **Riemann-Lebesgue lemma**.

Lemma 6.5 (Riemann-Lebesgue)**.** *Let $C_\infty(\mathbb{R}^n)$ denote the Banach space of all continuous functions $f : \mathbb{R}^n \to \mathbb{C}$ which vanish at ∞ equipped with the sup norm. Then the Fourier transform is a bounded map from $L^1(\mathbb{R}^n)$ into $C_\infty(\mathbb{R}^n)$ satisfying*

$$\|\hat{f}\|_\infty \le (2\pi)^{-n/2} \|f\|_1. \tag{6.19}$$

Proof. Clearly we have $\hat{f} \in C_\infty(\mathbb{R}^n)$ if $f \in \mathcal{S}(\mathbb{R}^n)$. Moreover, the estimate

$$\sup_p |\hat{f}(p)| \le \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} |\mathrm{e}^{-\mathrm{i}px} f(x)| d^n x = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} |f(x)| d^n x. \tag{6.20}$$

shows $\hat{f} \in C_\infty(\mathbb{R}^n)$ for arbitrary $f \in L^1(\mathbb{R}^n)$ since $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n)$. □

6.2. The free Schrödinger operator

In Section 2.1 we have seen that the Hilbert space corresponding to one particle in $\mathbb{R}^3$ is $L^2(\mathbb{R}^3)$. More generally, the Hilbert space for N particles in $\mathbb{R}^d$ is $L^2(\mathbb{R}^n)$, $n = Nd$. The corresponding non relativistic Hamilton operator, if the particles do not interact, is given by

$$H_0 = -\Delta, \tag{6.21}$$

where Δ is the Laplace operator

$$\Delta = \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}. \tag{6.22}$$

Our first task is to find a good domain such that H_0 is a self-adjoint operator.

By Lemma 6.1 we have that

$$-\Delta\psi(x) = (p^2 \hat{\psi}(p))^\vee(x), \qquad \psi \in H^2(\mathbb{R}^n), \tag{6.23}$$

and hence the operator

$$H_0\psi = -\Delta\psi, \qquad \mathfrak{D}(H_0) = H^2(\mathbb{R}^n), \tag{6.24}$$

is unitarily equivalent to the maximally defined multiplication operator

$$(\mathcal{F} H_0 \mathcal{F}^{-1})\varphi(p) = p^2\varphi(p), \qquad \mathfrak{D}(p^2) = \{\varphi \in L^2(\mathbb{R}^n) | p^2\varphi(p) \in L^2(\mathbb{R}^n)\}. \tag{6.25}$$

Theorem 6.6. *The free Schrödinger operator H_0 is self-adjoint and its spectrum is characterized by*

$$\sigma(H_0) = \sigma_{ac}(H_0) = [0, \infty), \qquad \sigma_{sc}(H_0) = \sigma_{pp}(H_0) = \emptyset. \tag{6.26}$$

Proof. It suffices to show that $d\mu_\psi$ is purely absolutely continuous for every ψ. First observe that

$$\langle \psi, R_{H_0}(z)\psi\rangle = \langle \hat{\psi}, R_{p^2}(z)\hat{\psi}\rangle = \int_{\mathbb{R}^n} \frac{|\hat{\psi}(p)|^2}{p^2 - z} d^n p = \int_{\mathbb{R}} \frac{1}{r^2 - z} d\tilde{\mu}_\psi(r), \tag{6.27}$$

where

$$d\tilde{\mu}_\psi(r) = \chi_{[0,\infty)}(r) r^{n-1} \left(\int_{S^{n-1}} |\hat{\psi}(r\omega)|^2 d^{n-1}\omega \right) dr. \tag{6.28}$$

Hence, after a change of coordinates, we have

$$\langle \psi, R_{H_0}(z)\psi\rangle = \int_{\mathbb{R}} \frac{1}{\lambda - z} d\mu_\psi(\lambda), \tag{6.29}$$

where

$$d\mu_\psi(\lambda) = \frac{1}{2}\chi_{[0,\infty)}(\lambda)\lambda^{n/2-1} \left(\int_{S^{n-1}} |\hat{\psi}(\sqrt{\lambda}\omega)|^2 d^{n-1}\omega \right) d\lambda, \tag{6.30}$$

proving the claim. □

Finally, we note that $C_0^\infty(\mathbb{R}^n)$ is a core for H_0.

Lemma 6.7. *The set $C_0^\infty(\mathbb{R}^n) = \{f \in \mathcal{S}(\mathbb{R}^n) | \mathrm{supp}(f)$ is compact$\}$ is a core for H_0.*

Proof. Let $A = H_0|_{C_0^\infty(\mathbb{R}^n)}$. It suffices to show $A^* \subseteq H_0$, implying $H_0 = H_0^* \subseteq A^{**} = \overline{A}$ and hence $H_0 = \overline{A}$. Let $\psi \in \mathfrak{D}(A^*)$, then there is a θ such that $\langle -\Delta\varphi, \psi\rangle = \langle \varphi, \theta\rangle$ for all $\varphi \in C_0^\infty(\mathbb{R}^n)$. Taking the Fourier transform we see $\langle p^2\hat{\varphi}, \hat{\psi}\rangle = \langle \hat{\varphi}, \hat{\theta}\rangle$ from which we expect $p^2\hat{\psi} = \hat{\theta}$. But since we don't know $p^2\hat{\psi} \in L^2(\mathbb{R}^n)$, a more careful argument is needed. Switching to $1 + p^2$ we see

$$\langle (1 + p^2)\hat{\varphi}, \hat{\psi} - \frac{1}{1 + p^2}(\hat{\psi} + \hat{\theta})\rangle, \qquad \varphi \in C_0^\infty(\mathbb{R}^n), \tag{6.31}$$

implying $\hat{\psi} = \frac{1}{1+p^2}(\hat{\psi} + \hat{\theta}) = 0$. Hence $p^2\hat{\psi} = \hat{\theta} \in L^2(\mathbb{R}^n)$ and thus $\psi \in \mathfrak{D}(H_0)$. □

Note also that the quadratic form of H_0 is given by

$$q_{H_0}(\psi) = \sum_{j=1}^{n} \int_{\mathbb{R}^n} |\partial_j \psi(x)|^2 d^n x, \quad \psi \in \mathfrak{Q}(H_0) = H^1(\mathbb{R}^n). \tag{6.32}$$

6.3. The time evolution in the free case

Now let us look at the time evolution. We have

$$\mathrm{e}^{-\mathrm{i}tH_0}\psi(x) = \mathcal{F}^{-1}\mathrm{e}^{-\mathrm{i}tp^2}\hat{\psi}(p). \tag{6.33}$$

The right hand side is a product and hence our operator should be expressible as an integral operator via the convolution formula. However, since $\mathrm{e}^{-\mathrm{i}tp^2}$ is not in L^2, a more careful analysis is needed.

Consider

$$f_\varepsilon(p^2) = \mathrm{e}^{-(\mathrm{i}t+\varepsilon)p^2}, \quad \varepsilon > 0. \tag{6.34}$$

Then $f_\varepsilon(H_0)\psi \to \mathrm{e}^{-\mathrm{i}tH_0}\psi$ by Theorem 3.1. Moreover, by Lemma 6.3 and the convolution formula we have

$$f_\varepsilon(H_0)\psi(x) = \frac{1}{(4\pi(\mathrm{i}t+\varepsilon))^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{-\frac{|x-y|^2}{4(\mathrm{i}t+\varepsilon)}} \psi(y) d^n y \tag{6.35}$$

and hence

$$\mathrm{e}^{-\mathrm{i}tH_0}\psi(x) = \frac{1}{(4\pi \mathrm{i}t)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}\frac{|x-y|^2}{4t}} \psi(y) d^n y \tag{6.36}$$

for $t \neq 0$ and $\psi \in L^1 \cap L^2$. For general $\psi \in L^2$ the integral has to be understood as a limit.

Using this explicit form, it is not hard to draw some immediate consequences. For example, if $\psi \in L^2(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$, then $\psi(t) \in C(\mathbb{R}^n)$ for $t \neq 0$ (use dominated convergence and continuity of the exponential) and satisfies

$$\|\psi(t)\|_\infty \le \frac{1}{|4\pi t|^{n/2}} \|\psi(0)\|_1 \tag{6.37}$$

by the Riemann-Lebesgue lemma. Thus we have spreading of wave functions in this case. Moreover, it is even possible to determine the asymptotic form of the wave function for large t as follows. Observe

$$\begin{aligned} \mathrm{e}^{-\mathrm{i}tH_0}\psi(x) &= \frac{\mathrm{e}^{\mathrm{i}\frac{x^2}{4t}}}{(4\pi \mathrm{i}t)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}\frac{y^2}{4t}} \psi(y) \mathrm{e}^{\mathrm{i}\frac{xy}{2t}} d^n y \\ &= \left(\frac{1}{2\mathrm{i}t}\right)^{n/2} \mathrm{e}^{\mathrm{i}\frac{x^2}{4t}} \left(\mathrm{e}^{\mathrm{i}\frac{y^2}{4t}}\psi(y)\right)^{\wedge}\left(\frac{x}{2t}\right). \end{aligned} \tag{6.38}$$

Moreover, since $\exp(\mathrm{i}\frac{y^2}{4t})\psi(y) \to \psi(y)$ in L^2 as $|t| \to \infty$ (dominated convergence) we obtain

Lemma 6.8. *For any* $\psi \in L^2(\mathbb{R}^n)$ *we have*

$$\mathrm{e}^{-\mathrm{i}tH_0}\psi(x) - \left(\frac{1}{2\mathrm{i}t}\right)^{n/2} \mathrm{e}^{\mathrm{i}\frac{x^2}{4t}}\hat{\psi}(\frac{x}{2t}) \to 0 \tag{6.39}$$

in L^2 *as* $|t| \to \infty$.

Next we want to apply the RAGE theorem in order to show that for any initial condition, a particle will escape to infinity. But first we need some good criteria when an operator is compact.

Of particular interest for us is the case of integral operators

$$K\psi(x) = \int_{\mathbb{R}^n} K(x,y)\psi(y)d^n y, \tag{6.40}$$

where $K(x,y) \in L^2(\mathbb{R}^n \times \mathbb{R}^n)$. Such an operator is called **Hilbert-Schmidt operator**. Using Cauchy-Schwarz,

$$\begin{aligned}\int_{\mathbb{R}^n} |K\psi(x)|^2 d^n x &= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} |K(x,y)\psi(y)| d^n y \right|^2 d^n x \\ &\le \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |K(x,y)|^2 d^n y \right) \left(\int_{\mathbb{R}^n} |\psi(y)|^2 d^n y \right) d^n x \\ &= \left(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |K(x,y)|^2 d^n y\, d^n x \right) \left(\int_{\mathbb{R}^n} |\psi(y)|^2 d^n y \right) \end{aligned} \tag{6.41}$$

we see that K is bounded. Next, pick an orthonormal basis $\varphi_j(x)$ for $L^2(\mathbb{R}^n)$. Then, by Lemma 1.8, $\varphi_i(x)\varphi_j(y)$ is an orthonormal basis for $L^2(\mathbb{R}^n \times \mathbb{R}^n)$ and

$$K(x,y) = \sum_{i,j} c_{i,j}\varphi_i(x)\varphi_j(y), \quad c_{i,j} = \langle \varphi_i, K\varphi_j \rangle, \tag{6.42}$$

where

$$\sum_{i,j} |c_{i,j}|^2 = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |K(x,y)|^2 d^n y\, d^n x < \infty. \tag{6.43}$$

In particular,

$$K\psi(x) = \sum_{i,j} c_{i,j} \langle \varphi_j, \psi \rangle \varphi_i(x) \tag{6.44}$$

shows that K can be approximated by finite rank operators (take finitely many terms in the sum) and is hence compact.

Now we can prove

Lemma 6.9. *Let* $g(x)$ *be the multiplication operator by* g *and let* $f(p)$ *be the operator given by* $f(p)\psi(x) = \mathcal{F}^{-1}(f(p)\hat{\psi}(p))(x)$. *Denote by* $L^\infty_\infty(\mathbb{R}^n)$ *the bounded Borel functions which vanish at infinity. Then*

$$f(p)g(x) \quad \text{and} \quad g(x)f(p) \tag{6.45}$$

are compact if $f, g \in L^\infty_\infty(\mathbb{R}^n)$ and (extend to) Hilbert-Schmidt operators if $f, g \in L^2(\mathbb{R}^n)$.

Proof. By symmetry it suffices to consider $g(x)f(p)$. Let $f, g \in L^2$, then

$$g(x)f(p)\psi(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} g(x)\check{f}(x-y)\psi(y)d^n y \tag{6.46}$$

shows that $g(x)f(p)$ is Hilbert-Schmidt since $g(x)\check{f}(x-y) \in L^2(\mathbb{R}^n \times \mathbb{R}^n)$.

If f, g are bounded then the functions $f_R(p) = \chi_{\{p|p^2 \le R\}}(p)f(p)$ and $g_R(x) = \chi_{\{x|x^2 \le R\}}(x)g(x)$ are in L^2. Thus $g_R(x)f_R(p)$ is compact and tends to $g(x)f(p)$ in norm since f, g vanish at infinity. □

In particular, this lemma implies that

$$\chi_\Omega(H_0 + \mathrm{i})^{-1} \tag{6.47}$$

is compact if $\Omega \subseteq \mathbb{R}^n$ is bounded and hence

$$\lim_{t\to\infty} \|\chi_\Omega \mathrm{e}^{-\mathrm{i}tH_0}\psi\|^2 = 0 \tag{6.48}$$

for any $\psi \in L^2(\mathbb{R}^n)$ and any bounded subset Ω of $\mathbb{R}^n$. In other words, the particle will eventually escape to infinity since the probability of finding the particle in any bounded set tends to zero. (If $\psi \in L^1(\mathbb{R}^n)$ this of course also follows from (6.37).)

6.4. The resolvent and Green's function

Now let us compute the resolvent of H_0. We will try to use a similar approach as for the time evolution in the previous section. However, since it is highly nontrivial to compute the inverse Fourier transform of $\exp(-\varepsilon p^2)(p^2 - z)^{-1}$ directly, we will use a small ruse.

Note that

$$R_{H_0}(z) = \int_0^\infty \mathrm{e}^{zt}\mathrm{e}^{-tH_0}dt, \qquad \mathrm{Re}(z) < 0 \tag{6.49}$$

by Lemma 4.1. Moreover,

$$\mathrm{e}^{-tH_0}\psi(x) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{-\frac{|x-y|^2}{4t}}\psi(y)d^n y, \quad t > 0, \tag{6.50}$$

by the same analysis as in the previous section. Hence, by Fubini, we have

$$R_{H_0}(z)\psi(x) = \int_{\mathbb{R}^n} G_0(z, |x-y|)\psi(y)d^n y, \tag{6.51}$$

where

$$G_0(z, r) = \int_0^\infty \frac{1}{(4\pi t)^{n/2}} \mathrm{e}^{-\frac{r^2}{4t}+zt}dt, \quad r > 0, \ \mathrm{Re}(z) < 0. \tag{6.52}$$

The function $G_0(z, r)$ is called **Green's function** of H_0. The integral can be evaluated in terms of modified Bessel functions of the second kind

$$G_0(z, r) = \frac{1}{2\pi} \left(\frac{-z}{4\pi^2 r^2} \right)^{\frac{n-2}{4}} K_{\frac{n}{2}-1}(\sqrt{-z}r). \tag{6.53}$$

The functions $K_\nu(x)$ satisfy the following differential equation

$$\left(\frac{d^2}{dx^2} + \frac{1}{x}\frac{d}{dx} - 1 - \frac{\nu^2}{x^2} \right) K_\nu(x) = 0 \tag{6.54}$$

and have the following asymptotics

$$K_\nu(x) = \begin{cases} \frac{\Gamma(\nu)}{2} \left(\frac{x}{2}\right)^{-\nu} + O(x^{-\nu+1}) & \nu \neq 0 \\ -\ln(\frac{x}{2}) + O(1) & \nu = 0 \end{cases} \tag{6.55}$$

for $|x| \to 0$ and

$$K_\nu(x) = \sqrt{\frac{\pi}{2x}} \mathrm{e}^{-x}(1 + O(x^{-1})) \tag{6.56}$$

for $|x| \to \infty$. For more information see for example [**18**]. In particular, $G_0(z, r)$ has an analytic continuation for $z \in \mathbb{C}\backslash[0, \infty) = \rho(H_0)$. Hence we can define the right hand side of (6.51) for all $z \in \rho(H_0)$ such that

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \varphi(x) G_0(z, |x - y|) \psi(y) d^n y d^n x \tag{6.57}$$

is analytic for $z \in \rho(H_0)$ and $\varphi, \psi \in \mathcal{S}(\mathbb{R}^n)$ (by Morea's theorem). Since it is equal to $\langle \varphi, R_{H_0}(z)\psi \rangle$ for $\mathrm{Re}(z) < 0$ it is equal to this function for all $z \in \rho(H_0)$, since both functions are analytic in this domain. In particular, (6.51) holds for all $z \in \rho(H_0)$.

If n is odd, we have the case of spherical Bessel functions which can be expressed in terms of elementary functions. For example, we have

$$G_0(z, r) = \frac{1}{2\sqrt{-z}} \mathrm{e}^{-\sqrt{-z}\, r}, \qquad n = 1, \tag{6.58}$$

and

$$G_0(z, r) = \frac{1}{4\pi r} \mathrm{e}^{-\sqrt{-z}\, r}, \qquad n = 3. \tag{6.59}$$

Chapter 7

Algebraic methods

7.1. Position and momentum

Apart from the Hamiltonian H_0, which corresponds to the kinetic energy, there are several other important observables associated with a single particle in three dimensions. Using commutation relation between these observables, many important consequences about these observables can be derived.

First consider the one-parameter unitary group

$$(U_j(t)\psi)(x) = \mathrm{e}^{-\mathrm{i}tx_j}\psi(x), \qquad 1 \le j \le 3. \tag{7.1}$$

For $\psi \in \mathcal{S}(\mathbb{R}^3)$ we compute

$$\lim_{t\to 0} \mathrm{i}\frac{\mathrm{e}^{-\mathrm{i}tx_j}\psi(x) - \psi(x)}{t} = x_j\psi(x) \tag{7.2}$$

and hence the generator is the multiplication operator by the j-th coordinate function. By Corollary 5.3 it is essentially self-adjoint on $\psi \in \mathcal{S}(\mathbb{R}^3)$. It is custom to combine all three operators to one vector valued operator x, which is known as **position operator**. Moreover, it is not hard to see that the spectrum of x_j is purely absolutely continuous and given by $\sigma(x_j) = \mathbb{R}$. In fact, let $\varphi(x)$ be an orthonormal basis for $L^2(\mathbb{R})$. Then $\varphi_i(x_1)\varphi_j(x_2)\varphi_k(x_3)$ is an orthonormal basis for $L^2(\mathbb{R}^3)$ and x_1 can be written as a orthogonal sum of operators restricted to the subspaces spanned by $\varphi_j(x_2)\varphi_k(x_3)$. Each subspace is unitarily equivalent to $L^2(\mathbb{R})$ and x_1 is given by multiplication with the identity. Hence the claim follows (or use Theorem 4.13).

Next, consider the one-parameter unitary group of translations

$$(U_j(t)\psi)(x) = \psi(x - te_j), \qquad 1 \le j \le 3, \tag{7.3}$$

where e_j is the unit vector in the j-th coordinate direction. For $\psi \in \mathcal{S}(\mathbb{R}^3)$ we compute

$$\lim_{t\to 0} \mathrm{i}\frac{\psi(x - te_j) - \psi(x)}{t} = \frac{1}{\mathrm{i}} \frac{\partial}{\partial x_j} \psi(x) \tag{7.4}$$

and hence the generator is $p_j = \frac{1}{\mathrm{i}} \frac{\partial}{\partial x_j}$. Again it is essentially self-adjoint on $\psi \in \mathcal{S}(\mathbb{R}^3)$. Moreover, since it is unitarily equivalent to x_j by virtue of the Fourier transform we conclude that the spectrum of p_j is again purely absolutely continuous and given by $\sigma(p_j) = \mathbb{R}$. The operator p is known as **momentum operator**. Note that since

$$[H_0, p_j]\psi(x) = 0, \qquad \psi \in \mathcal{S}(\mathbb{R}^3) \tag{7.5}$$

we have

$$\frac{d}{dt}\langle \psi(t), p_j \psi(t)\rangle = 0, \qquad \psi(t) = \mathrm{e}^{-\mathrm{i}tH_0}\psi(0) \in \mathcal{S}(\mathbb{R}^3), \tag{7.6}$$

that is, the momentum is a conserved quantity for the free motion. Similarly one has

$$[p_j, x_k]\psi(x) = \delta_{jk}\psi(x), \qquad \psi \in \mathcal{S}(\mathbb{R}^3), \tag{7.7}$$

which is known as the **Weyl relation**.

7.2. Angular momentum

Now consider the one-parameter unitary group of rotations

$$(U_j(t)\psi)(x) = \psi(M_j(t)x), \qquad 1 \le j \le 3, \tag{7.8}$$

where $M_j(t)$ is the matrix of rotation around e_j by an angle of t. For $\psi \in \mathcal{S}(\mathbb{R}^3)$ we compute

$$\lim_{t\to 0} \mathrm{i}\frac{\psi(M_i(t)x) - \psi(x)}{t} = \sum_{j,k=1}^{3} \varepsilon_{ijk} x_j p_k \psi(x), \tag{7.9}$$

where

$$\varepsilon_{ijk} = \begin{cases} 1 & \text{if } ijk \text{ is an even permutation of } 123 \\ -1 & \text{if } ijk \text{ is an odd permutation of } 123 \\ 0 & \text{else} \end{cases}. \tag{7.10}$$

Again one combines the three components to one vector valued operator $L = x \wedge p$, which is known as **angular momentum operator**. Since $\mathrm{e}^{\mathrm{i}2\pi L_j} = \mathbb{I}$, we see that the spectrum is a subset of $\mathbb{Z}$. In particular, the continuous spectrum is empty. We will show below that we have $\sigma(L_j) = \mathbb{Z}$. Note that since

$$[H_0, L_j]\psi(x) = 0, \qquad \psi \in \mathcal{S}(\mathbb{R}^3), \tag{7.11}$$

we have again

$$\frac{d}{dt}\langle \psi(t), L_j \psi(t)\rangle = 0, \qquad \psi(t) = \mathrm{e}^{-\mathrm{i}tH_0}\psi(0) \in \mathcal{S}(\mathbb{R}^3), \tag{7.12}$$

that is, the angular momentum is a conserved quantity for the free motion as well.

Moreover, we even have

$$[L_i, K_j]\psi(x) = \mathrm{i} \sum_{k=1}^{3} \varepsilon_{ijk} K_k \psi(x), \qquad \psi \in \mathcal{S}(\mathbb{R}^3), K_j \in \{L_j, p_j, x_j\}, \quad (7.13)$$

and these algebraic commutation relations are often used to derive information on the point spectra of these operators. In this respect the following domain

$$\mathfrak{D} = \mathrm{span}\{x^\alpha \mathrm{e}^{-\frac{x^2}{2}} \,|\, \alpha \in \mathbb{N}_0^n\} \subset \mathcal{S}(\mathbb{R}^n) \qquad (7.14)$$

is often used. It has the nice property that the finite dimensional subspaces

$$\mathfrak{D}_k = \mathrm{span}\{x^\alpha \mathrm{e}^{-\frac{x^2}{2}} \,|\, |\alpha| \le k\} \qquad (7.15)$$

are invariant under L_j (and hence they reduce L_j).

Lemma 7.1. *The subspace $\mathfrak{D} \subset L^2(\mathbb{R}^n)$ defined in (7.14) is dense.*

Proof. By Lemma 1.8 it suffices to consider the case $n = 1$. Suppose $\langle \varphi, \psi \rangle = 0$ for every $\psi \in \mathfrak{D}$. Then

$$\frac{1}{\sqrt{2\pi}} \int \overline{\varphi(x)} \mathrm{e}^{-\frac{x^2}{2}} \sum_{j=1}^{k} \frac{(\mathrm{i}tx)^k}{j!} = 0 \qquad (7.16)$$

for any finite k and hence also in the limit $k \to \infty$ by the dominated convergence theorem. But the limit is the Fourier transform of $\overline{\varphi(x)}\mathrm{e}^{-\frac{x^2}{2}}$, which shows that this function is zero. Hence $\varphi(x) = 0$. □

Since it is invariant under the unitary groups generated by L_j, the operators L_j are essentially self-adjoint on $\mathfrak{D}$ by Corollary 5.3.

Introducing $L^2 = L_1^2 + L_2^2 + L_3^2$ it is straightforward to check

$$[L^2, L_j]\psi(x) = 0, \qquad \psi \in \mathcal{S}(\mathbb{R}^3). \qquad (7.17)$$

Moreover, $\mathfrak{D}_k$ is invariant under L^2 and L_3 and hence $\mathfrak{D}_k$ reduces L^2 and L_3. In particular, L^2 and L_3 are given by finite matrices on $\mathfrak{D}_k$. Now let $\mathfrak{H}_m = \mathrm{Ker}(L_3 - m)$ and denote by P_k the projector onto $\mathfrak{D}_k$. Since L^2 and L_3 commute on $\mathfrak{D}_k$, the space $P_k\mathfrak{H}_m$ is invariant under L^2 which shows that we can choose an orthonormal basis consisting of eigenfunctions of L^2 for $P_k\mathfrak{H}_m$. Increasing k we get an orthonormal set of simultaneous eigenfunctions whose span is equal to $\mathfrak{D}$. Hence there is an orthonormal basis of simultaneous eigenfunctions of L^2 and L_3.

Now let us try to draw some further consequences by using the commutation relations (7.13). (All commutation relations below hold for $\psi \in \mathcal{S}(\mathbb{R}^3)$.) Denote by $\mathfrak{H}_{l,m}$ the set of all functions in $\mathfrak{D}$ satisfying

$$L_3\psi = m\psi, \qquad L^2\psi = l(l+1)\psi. \tag{7.18}$$

By $L^2 \geq 0$ and $\sigma(L_3) \subseteq \mathbb{Z}$ we can restrict our attention to the case $l \geq 0$ and $m \in \mathbb{Z}$.

First introduce two new operators

$$L_\pm = L_1 \pm \mathrm{i}L_2, \qquad [L_3, L_\pm] = \pm L_\pm. \tag{7.19}$$

Then, for every $\psi \in \mathfrak{H}_{l,m}$ we have

$$L_3(L_\pm\psi) = (m \pm 1)(L_\pm\psi), \qquad L^2(L_\pm\psi) = l(l+1)(L_\pm\psi), \tag{7.20}$$

that is, $L_\pm\mathfrak{H}_{l,m} \to \mathfrak{H}_{l,m\pm1}$. Moreover, since

$$L^2 = L_3^2 \pm L_3 + L_\mp L_\pm \tag{7.21}$$

we obtain

$$\|L_\pm\psi\|^2 = \langle\psi, L_\mp L_\pm\psi\rangle = (l(l+1) - m(m \pm 1))\|\psi\| \tag{7.22}$$

for every $\psi \in \mathfrak{H}_{l,m}$. If $\psi \neq 0$ we must have $l(l+1) - m(m \pm 1) \geq 0$ which shows $\mathfrak{H}_{l,m} = \{0\}$ for $|m| > l$. Moreover, $L_\pm\mathfrak{H}_{l,m} \to \mathfrak{H}_{l,m\pm1}$ is injective unless $|m| = l$. Hence we must have $\mathfrak{H}_{l,m} = \{0\}$ for $l \notin \mathbb{N}_0$.

Up to this point we know $\sigma(L^2) \subseteq \{l(l+1)|l \in \mathbb{N}_0\}$, $\sigma(L_3) \subseteq \mathbb{Z}$. In order to show that equality holds in both cases, we need to show that $\mathfrak{H}_{l,m} \neq \{0\}$ for $l \in \mathbb{N}_0$, $m = -l, -l+1, \ldots, l-1, l$. First of all we observe

$$\psi_{0,0}(x) = \frac{1}{\pi^{3/2}}\mathrm{e}^{-\frac{x^2}{2}} \in \mathfrak{H}_{0,0}. \tag{7.23}$$

Next, we note that (7.13) implies

$$\begin{aligned}
&[L_3, x_\pm] = \pm x_\pm, \qquad x_\pm = x_1 \pm \mathrm{i}x_2,\\
&[L_\pm, x_\pm] = 0, \quad [L_\pm, x_\mp] = \pm 2x_3,\\
&[L^2, x_\pm] = 2x_\pm(1 \pm L_3) \mp 2x_3L_\pm.
\end{aligned} \tag{7.24}$$

Hence if $\psi \in \mathfrak{H}_{l,l}$, then $(x_1 \pm \mathrm{i}x_2)\psi \in \mathfrak{H}_{l\pm1,l\pm1}$. And thus

$$\psi_{l,l}(x) = \frac{1}{\sqrt{l!}}(x_1 \pm \mathrm{i}x_2)^l\psi_{0,0}(x) \in \mathfrak{H}_{l,l}, \tag{7.25}$$

respectively

$$\psi_{l,m}(x) = \sqrt{\frac{(l+m)!}{(l-m)!(2l)!}}L_-^{l-m}\psi_{l,l}(x) \in \mathfrak{H}_{l,m}. \tag{7.26}$$

The constants are chosen such that $\|\psi_{l,m}\| = 1$.

In summary,

Theorem 7.2. *There exists an orthonormal basis of simultaneous eigenvectors for the operators* L^2 *and* L_j*. Moreover, their spectra are given by*

$$\sigma(L^2) = \{l(l+1) | l \in \mathbb{N}_0\}, \qquad \sigma(L_3) = \mathbb{Z}. \tag{7.27}$$

We will rederive this result using different methods in Section 9.2.

7.3. The harmonic oscillator

Finally, let us consider another important model whose algebraic structure is similar to those of the angular momentum, the **harmonic oscillator**

$$H = H_0 + \omega^2 x^2, \qquad \omega > 0. \tag{7.28}$$

As domain we will choose

$$\mathfrak{D}(H) = \mathfrak{D} = \operatorname{span}\{x^\alpha \mathrm{e}^{-\frac{x^2}{2}} \,|\, \alpha \in \mathbb{N}_0^3\} \subseteq L^2(\mathbb{R}^3) \tag{7.29}$$

from our previous section.

We will first consider the one-dimensional case. Introducing

$$A_\pm = \frac{1}{\sqrt{2}}\left(\sqrt{\omega}x \mp \frac{1}{\sqrt{\omega}}\frac{d}{dx}\right) \tag{7.30}$$

we have

$$[A_-, A_+] = 1 \tag{7.31}$$

and

$$H = \omega(2N+1), \quad N = A_+ A_-, \tag{7.32}$$

for any function in $\mathfrak{D}$.

Moreover, since

$$[N, A_\pm] = \pm A_\pm, \tag{7.33}$$

we see that $N\psi = n\psi$ implies $NA_\pm\psi = (n \pm 1)A_\pm\psi$. Moreover, $\|A_+\psi\|^2 = \langle\psi, A_- A_+\psi\rangle = (n+1)\|\psi\|^2$ respectively $\|A_-\psi\|^2 = n\|\psi\|^2$ in this case and hence we conclude that $\sigma(N) \subseteq \mathbb{N}_0$

If $N\psi_0 = 0$, then we must have $A_-\psi = 0$ and the normalized solution of this last equation is given by

$$\psi_0(x) = \left(\frac{\omega}{\pi}\right)^{1/4} \mathrm{e}^{-\frac{\omega x^2}{2}} \in \mathfrak{D}. \tag{7.34}$$

Hence

$$\psi_n(x) = \frac{1}{\sqrt{n!}} A_+^n \psi_0(x) \tag{7.35}$$

is a normalized eigenfunction of N corresponding to the eigenvalue n. Moreover, since

$$\psi_n(x) = \frac{1}{\sqrt{n!}}\left(\frac{\omega}{4\pi}\right)^{1/4} H_n(\frac{x}{\sqrt{\omega}})\mathrm{e}^{-\frac{\omega x^2}{2}} \tag{7.36}$$

where $H_n(x)$ is a polynomial of degree n given by

$$H_n(x) = \mathrm{e}^{\frac{x^2}{2}} \left(x - \frac{d}{dx} \right)^n \mathrm{e}^{-\frac{x^2}{2}} = (-1)^n \mathrm{e}^{x^2} \frac{d^n}{dx^n} \mathrm{e}^{-x^2}, \qquad (7.37)$$

we conclude $\mathrm{span}\{\psi_n\} = \mathfrak{D}$. The polynomials $H_n(x)$ are called **Hermite polynomials**.

In summary,

Theorem 7.3. *The harmonic oscillator H is essentially self adjoint on $\mathfrak{D}$ and has an orthonormal basis of eigenfunctions*

$$\psi_{n_1,n_2,n_3}(x) = \psi_{n_1}(x_1)\psi_{n_2}(x_2)\psi_{n_3}(x_3), \qquad (7.38)$$

with $\psi_{n_j}(x_j)$ from (7.36). The spectrum is given by

$$\sigma(H) = \{(2n+3)\omega | n \in \mathbb{N}_0\}. \qquad (7.39)$$

Finally, there is also a close connection with the Fourier transformation. without restriction we choose $\omega = 1$ and consider only one dimension. Then it easy to verify that H commutes with the Fourier transformation

$$\mathcal{F}H = H\mathcal{F} \qquad (7.40)$$

and hence $\{\psi_n\}$ is an orthonormal basis of eigenfunctions for $\mathcal{F}$ as well. Moreover, by $\mathcal{F}A_\pm = \mp\mathrm{i}A_\pm\mathcal{F}$ we even infer

$$\mathcal{F}\psi_n = (-\mathrm{i})^n \psi_n \qquad (7.41)$$

and hence

$$\sigma(\mathcal{F}) = \{z \in \mathbb{C} | z^4 = 1\}. \qquad (7.42)$$

Chapter 8

Self-adjointness of Schrödinger operators

The Hamiltonian of a quantum mechanical system is usually the sum of the kinetic energy H_0 (free Schrödinger operator) plus an operator V corresponding to the potential energy. Since we already know a lot about H_0, we will try to consider V as a perturbation of H_0. This will only work if V is *small* with respect to H_0. Hence we study perturbations of self-adjoint operators first.

8.1. Relatively bounded operators and the Kato-Rellich theorem

An operator B is called A **bounded** or **relatively bounded** with respect to A if $\mathfrak{D}(A) \subseteq \mathfrak{D}(B)$ and if there are constants $a, b \geq 0$ such that

$$\|B\psi\| \leq a\|A\psi\| + b\|\psi\|, \qquad \psi \in \mathfrak{D}(A). \tag{8.1}$$

The infimum of all such constants is called the A-bound of B.

The triangle inequality implies

Lemma 8.1. *Suppose B_j, $j = 1, 2$, are A bounded with respective A-bounds a_i, $i = 1, 2$. Then $\alpha_1 B_1 + \alpha_2 B_2$ is also A bounded with A-bound less than $|\alpha_1|a_1 + |\alpha_2|a_2$. In particular, the set of all A bounded operators forms a linear space.*

We are mainly interested in the situation where A is self-adjoint and B is symmetric. Hence we will restrict our attention to this case.

Lemma 8.2. *Suppose A is self-adjoint and B is closed with $\mathfrak{D}(A) \subseteq \mathfrak{D}(B)$. Then B is A bounded.*

Proof. The set $\mathfrak{D}(A)$ is a Hilbert space with respect to the scalar product $\langle \varphi, \psi \rangle_A = \langle (1+|A|)\varphi, (1+|A|)\psi \rangle$. Moreover, since $\|\psi\|_A \le \|A\psi\| + \|\psi\|$, the restriction of B to $\mathfrak{D}(A)$ is still closed (with respect to this new norm) and hence there is a constant $a \ge 0$ such that $\|B\psi\| \le a\|\psi\|_A$ by the closed graph theorem (Theorem 2.7). □

Lemma 8.3. *Suppose A is self-adjoint and $\mathfrak{D}(A) \subseteq \mathfrak{D}(B)$. Then B is A bounded if and only if $BR_A(z)$ is bounded for some $z \in \rho(A)$. The A-bound of B is given by*

$$\lim_{\lambda\to\infty} \|BR_A(\pm \mathrm{i}\lambda)\|. \tag{8.2}$$

If A is bounded from below, we can also replace $\pm\mathrm{i}\lambda$ by $-\lambda$.

Proof. If $BR_A(z)$ is bounded for one $z \in \rho(A)$ it is bounded for all $z \in \rho(A)$ by the first resolvent formula. Hence it suffices to consider only $z = \pm\mathrm{i}\lambda$, $\lambda > 0$. Let $\varphi = R_A(\pm\mathrm{i}\lambda)\psi$ and let a_∞ be the A-bound of B. If B is A bounded, then (use the spectral theorem to estimate the norms)

$$\|BR_A(\pm\mathrm{i}\lambda)\psi\| \le a\|A\varphi\| + b\|\varphi\| \le (a + \frac{b}{\lambda})\|\psi\|. \tag{8.3}$$

Conversely, if $a_\lambda = \|BR_A(\pm\mathrm{i}\lambda)\| < \infty$, we have

$$\|B\varphi\| \le a_\lambda\|\psi\| \le a_\lambda(\|A\varphi\| + \lambda\|\psi\|). \tag{8.4}$$

In addition, this shows $a_0 \le a_\lambda \le a + b/\lambda$ and hence all limiting values of a_λ must lie in $[a_0, a]$. Since a can be chosen arbitrarily close to a_0 we are done.

The case where a is bounded from below is similar. □

Now we will show the basic perturbation result due to Kato and Rellich.

Theorem 8.4 (Kato-Rellich)**.** *Suppose A is (essentially) self-adjoint and B is symmetric with A-bound less then one. Then $A+B$, $\mathfrak{D}(A+B) = \mathfrak{D}(A)$, is (essentially) self-adjoint. If A is essentially self-adjoint we have $\mathfrak{D}(\overline{A}) \subseteq \mathfrak{D}(\overline{B})$ and $\overline{A} + \overline{B} = \overline{A+B}$.*

If A is bounded from below by γ, then $A+B$ is bounded from below by $\min(\gamma, b/(a-1))$.

Proof. Since $\mathfrak{D}(\overline{A}) \subseteq \mathfrak{D}(\overline{B})$ and $\mathfrak{D}(\overline{A}) \subseteq \mathfrak{D}(\overline{A+B})$ by (8.1) we can assume that A is closed (i.e., self-adjoint). It suffices to show that $\mathrm{Ran}(A+B\pm\mathrm{i}\lambda) = \mathfrak{H}$. By the above lemma we can find a $\lambda > 0$ such that $\|BR_A(\pm\mathrm{i}\lambda)\| < 1$. Hence $1 \in \rho(BR_A(\pm\mathrm{i}\lambda))$ and thus $\mathbb{I} + BR_A(\pm\mathrm{i}\lambda)$ is onto. Thus

$$(A + B \pm \mathrm{i}\lambda) = (\mathbb{I} + BR_A(\pm\mathrm{i}\lambda))(A \pm \mathrm{i}\lambda) \tag{8.5}$$

is onto and the proof of the first part is complete.

If A is bounded from below we can replace $\pm i\lambda$ by $-\lambda$ and the above equation shows that R_{A+B} exists for λ sufficiently large. By the proof of the previous lemma we can choose $-\lambda < \min(\gamma, b/(a-1))$. □

Finally, let us show that there is also a connection between the resolvents.

Lemma 8.5. *Suppose A and B are closed and $\mathfrak{D}(A) \subseteq \mathfrak{D}(B)$. Then we have the* **second resolvent formula**

$$R_{A+B}(z) - R_A(z) = -R_A(z) B R_{A+B}(z) = -R_{A+B}(z) B R_A(z) \tag{8.6}$$

for $z \in \rho(A) \cap \rho(A+B)$.

Proof. We compute

$$R_{A+B}(z) + R_A(z) B R_{A+B}(z) = R_A(z)(A+B-z) R_{A+B}(z) = R_A(z). \tag{8.7}$$

The second identity is similar. □

8.2. More on compact operators

Recall from Section 5.2 that we have introduced the set of compact operators $\mathfrak{C}(\mathfrak{H})$ as the closure of the set of all finite rank operators in $\mathfrak{L}(\mathfrak{H})$. Before we can proceed, we need to establish some further results for such operators. We begin by investigating the spectrum of self-adjoint compact operators and show that the spectral theorem takes a particular simple form in this case.

We introduce some notation first. The **discrete spectrum** $\sigma_d(A)$ is the set of all eigenvalues which are discrete points of the spectrum and whose corresponding eigenspace is finite dimensional. The complement of the discrete spectrum is called the **essential spectrum** $\sigma_{ess}(A) = \sigma(A) \backslash \sigma_d(A)$. If A is self-adjoint we might equivalently set

$$\sigma_d(A) = \{\lambda \in \sigma_p(A) | \operatorname{rank}(P_A((\lambda-\varepsilon, \lambda+\varepsilon))) < \infty \text{ for some } \varepsilon > 0\}. \tag{8.8}$$

respectively

$$\sigma_{ess}(A) \{\lambda \in \mathbb{R} | \operatorname{rank}(P_A((\lambda-\varepsilon, \lambda+\varepsilon))) = \infty \text{ for all } \varepsilon > 0\}. \tag{8.9}$$

Theorem 8.6 (Spectral theorem for compact operators)**.** *Suppose the operator K is self-adjoint and compact. Then the spectrum of K consists of an at most countable number of eigenvalues which can only cluster at 0. Moreover, the eigenspace to each nonzero eigenvalue is finite dimensional. In other words,*

$$\sigma_{ess}(K) \subseteq \{0\}, \tag{8.10}$$

where equality holds if and only if $\mathfrak{H}$ is infinite dimensional.

In addition, we have

$$K = \sum_{\lambda \in \sigma(K)} \lambda P_K(\{\lambda\}). \tag{8.11}$$

Proof. It suffices to show $\mathrm{rank}(P_K((\lambda - \varepsilon, \lambda + \varepsilon))) < \infty$ for $0 < \varepsilon < |\lambda|$.

Let K_n be a sequence of finite rank operators such that $\|K - K_n\| \le 1/n$. If $\mathrm{Ran} P_K((\lambda - \varepsilon, \lambda + \varepsilon))$ is infinite dimensional we can find a vector ψ_n in this range such that $\|\psi_n\| = 1$ and $K_n \psi_n = 0$. But this yields a contradiction

$$\frac{1}{n} \ge |\langle \psi_n, (K - K_n)\psi_n\rangle| = |\langle \psi_n, K\psi_n\rangle| \ge |\lambda| - \varepsilon > 0 \tag{8.12}$$

by (4.2). □

As a consequence we obtain the canonical form of a general compact operator.

Theorem 8.7 (Canonical form of compact operators)**.** *Let K be compact. There exists orthonormal sets $\{\psi_j\}$, $\{\varphi_j\}$ and positive numbers $\{\lambda_j\}$ such that*

$$K = \sum_j \lambda_j \langle \varphi_j, .\rangle \psi_j, \qquad K^* = \sum_j \lambda_j \langle \psi_j, .\rangle \varphi_j. \tag{8.13}$$

Note $K\varphi_j = \lambda_j \psi_j$ and $K^*\psi_j = \lambda_j \varphi_j$ and hence $K^*K\varphi_j = \lambda_j^2 \varphi_j$ and $KK^*\psi_j = \lambda_j^2 \psi_j$.

The numbers $\lambda_j^2 > 0$ are the nonzero eigenvalues of KK^* respectively K^*K and are called **singular values** *of K.*

Proof. Let $\{\varphi_j\}$ be an orthonormal basis of eigenvectors for $P_{K^*K}((0,\infty))\mathfrak{H}$ and let λ_j^2 be the eigenvalue corresponding to φ_j. Then, for any $\psi \in \mathfrak{H}$ we can write

$$\psi = \sum_j \langle \varphi_j, \psi\rangle \varphi_j + \tilde{\psi} \tag{8.14}$$

with $\tilde{\psi} \in \mathrm{Ker}(K^*K)$. Then

$$K\psi = \sum_j \lambda_j \langle \varphi_j, \psi\rangle \psi_j, \tag{8.15}$$

where $\psi_j = \lambda_j^{-1} K\varphi_j$, since $\|K\tilde{\psi}\| = \langle \tilde{\psi}, K^*K\tilde{\psi}\rangle = 0$. That $\{\psi_j\}$ are orthonormal is a straightforward calculation and the formula for K^* can be proven similar. □

Finally, let me remark that there are a number of other equivalent definitions for compact operators.

Lemma 8.8. *For $K \in \mathfrak{L}(\mathfrak{H})$ the following statements are equivalent:*

(1) *K is compact.*

(1') *K^* is compact.*

(2) *$A_n \in \mathfrak{L}(\mathfrak{H})$ is normal and $A_n \to A$ strongly implies $KA_n \to KA$.*

(2') *$A_n \in \mathfrak{L}(\mathfrak{H})$ and $A_n \to A$ strongly implies $A_n K \to AK$.*

(3) *$\psi_n \to \psi$ weakly implies $K\psi_n \to K\psi$ in norm.*

(4) *ψ_n bounded implies that $K\psi_n$ has a (norm) convergent subsequence.*

Proof. (1) $\Leftrightarrow$ (1'). This is immediate from Theorem 8.7.

(1) $\Rightarrow$ (2). Translating $A_n \to A_n - A$ it is no restriction to assume $A = 0$. Since $\|A_n\| \le M$ by the uniform boundedness principle, it suffices to consider the case where K is finite rank. Then (by (8.13))

$$\|KA_n\|^2 = \sup_{\|\psi\|=1} \sum_{j=1}^n \lambda_j |\langle \varphi_j, A_n \psi\rangle|^2 \le \sum_{j=1}^n \lambda_j \|A_n \varphi_j\|^2 \to 0 \tag{8.16}$$

since $\|A_n^* \varphi_j\| = \|A_n \varphi_j\|$.

(2) $\Rightarrow$ (3). Again, replace $\psi_n \to \psi_n - \psi$ and assume $\psi = 0$. Choose $A_n = \langle K\psi_n, .\rangle \psi_n$, then $\|KA_n\| = \|K\psi_n\|^2 \to 0$.

(3) $\Rightarrow$ (4). If ψ_n is bounded it has a weakly convergent subsequence. Now apply (3) to this subsequence.

(4) $\Rightarrow$ (1). Let φ_j be an orthonormal basis and set

$$K_n = \sum_{j=1}^n \langle \varphi_j, .\rangle K\varphi_j. \tag{8.17}$$

Then

$$\lambda_n = \|K - K_n\| = \sup_{\psi \in \mathrm{span}\{\varphi_j\}_{j=n}^\infty, \|\psi\|=1} \|K\psi\| \tag{8.18}$$

is a decreasing sequence tending to a limit $\varepsilon \ge 0$. Moreover, we can find a sequence of unit vectors $\psi_n \in \mathrm{span}\{\varphi_j\}_{j=n}^\infty$ for which $\|K\psi_n\| \ge \varepsilon$. By assumption, $K\psi_n$ has a convergent subsequence which, since ψ_n converges weakly to 0, converges to 0. Hence ε must be 0 and we are done.

(1) $\Rightarrow$ (2') follows with a similar argument as for (1) $\Rightarrow$ (2).

(2') $\Rightarrow$ (1') Using $A_n = \langle \psi_n, .\rangle K^* \psi_n$ shows that (2') implies (3) with K replaced by K^*, which shows that K^* is compact. □

The last condition explains the name compact.

8.3. Relatively compact operators and Weyl's theorem

In the previous section we have seen that the sum of a self-adjoint and a symmetric operator is again self-adjoint if the perturbing operator is *small*. In this section we want to study the influence of perturbations on the spectrum. Our hope is that at least some parts of the spectrum remain invariant.

Let A be self-adjoint. Note that if we add a multiple of the identity to A, we shift the entire spectrum. Hence, in general, we cannot expect a (relatively) bounded perturbation to leave any part of the spectrum invariant. Next, if λ_0 is in the discrete spectrum, we can easily remove this eigenvalue with a finite rank perturbation of arbitrary small norm. In fact, consider

$$A + \varepsilon P_A(\{\lambda_0\}). \tag{8.19}$$

Hence our only hope is that the remainder, namely the essential spectrum, is stable under finite rank perturbations. To show this, we first need a good criterion for a point to be in the essential spectrum of A.

Lemma 8.9 (Weyl criterion). *A point λ is in the essential spectrum of a self-adjoint operator A if and only if there is a sequence ψ_n such that $\|\psi_n\| = 1$, ψ_n converges weakly to 0, and $\|(A - \lambda)\psi_n\| \to 0$. Moreover, the sequence can chosen to be orthonormal. Such a sequence is called* **singular Weyl sequence**.

Proof. Let ψ_n be a singular Weyl sequence for the point λ_0. By Lemma 2.12 we have $\lambda_0 \in \sigma(A)$ and hence it suffices to show $\lambda_0 \notin \sigma_d(A)$. If $\lambda_0 \in \sigma_d(A)$ we can find an $\varepsilon > 0$ such that $P_\varepsilon = P_A((\lambda_0 - \varepsilon, \lambda_0 + \varepsilon))$ is finite rank. Consider $\tilde{\psi}_n = P_\varepsilon \psi_n$. Clearly $\tilde{\psi}_n$ converges weakly to zero and $\|(A - \lambda_0)\tilde{\psi}_n\| \to 0$. Moreover,

$$\begin{aligned} \|\psi_n - \tilde{\psi}_n\|^2 &= \int_{\mathbb{R}\setminus(\lambda-\varepsilon,\lambda+\varepsilon)} d\mu_{\psi_n}(\lambda) \\ &\leq \frac{1}{\varepsilon^2} \int_{\mathbb{R}\setminus(\lambda-\varepsilon,\lambda+\varepsilon)} (\lambda - \lambda_0)^2 d\mu_{\psi_n}(\lambda) \\ &\leq \frac{1}{\varepsilon^2} \|(A - \lambda_0)\psi_n\|^2 \end{aligned} \tag{8.20}$$

and hence $\|\tilde{\psi}_n\| \to 1$. Thus $\varphi_n = \|\tilde{\psi}_n\|^{-1}\tilde{\psi}_n$ is also a singular Weyl sequence. But φ_n is a sequence of unit length vectors which lives in a finite dimensional space and converges to 0 weakly, a contradiction.

Conversely, if $\lambda_0 \in \sigma_{ess}(A)$, consider $P_n = P_A([\lambda - \frac{1}{n}, \lambda - \frac{1}{n+1}) \cup (\lambda + \frac{1}{n+1}, \lambda + \frac{1}{n}])$. Then $\operatorname{rank}(P_{n_j}) > 0$ for an infinite subsequence n_j. Now pick $\psi_j \in \operatorname{Ran} P_{n_j}$. □

Now let K be a self-adjoint compact operator and ψ_n a singular Weyl sequence for A. Then ψ_n converges weakly to zero and hence

$$\|(A + K - \lambda)\psi_n\| \leq \|(A - \lambda)\psi_n\| + \|K\psi_n\| \to 0 \tag{8.21}$$

since $\|(A-\lambda)\psi_n\| \to 0$ by assumption and $\|K\psi_n\| \to 0$ by Lemma 8.8 (3). Hence $\sigma_{ess}(A) \subseteq \sigma_{ess}(A+K)$. Reversing the roles of $A+K$ and A shows $\sigma_{ess}(A+K) = \sigma_{ess}(A)$. Since we have shown that we can remove any point in the discrete spectrum by a self-adjoint finite rank operator we obtain the following equivalent characterization of the essential spectrum.

Lemma 8.10. *The essential spectrum of a self-adjoint operator A is precisely the part which is invariant under rank-one perturbations. In particular,*

$$\sigma_{ess}(A) = \bigcap_{K\in\mathfrak{C}(\mathfrak{H}),K^*=K} \sigma(A+K). \tag{8.22}$$

There is even a larger class of operators under which the essential spectrum is invariant.

Theorem 8.11 (Weyl). *Suppose A and B are self-adjoint operators. If*

$$R_A(z) - R_B(z) \in \mathfrak{C}(\mathfrak{H}) \tag{8.23}$$

for one $z \in \rho(A)\cap\rho(B)$, then

$$\sigma_{ess}(A) = \sigma_{ess}(B). \tag{8.24}$$

Proof. In fact, suppose $\lambda \in \sigma_{ess}(A)$ and let ψ_n be a corresponding singular Weyl sequence. Then $(R_A(z) - \frac{1}{\lambda-z})\psi_n = \frac{R_A(z)}{\lambda-z}(A-\lambda)\psi_n$ and thus $\|(R_A(z) - \frac{1}{\lambda-z})\psi_n\| \to 0$. Moreover, by our assumption we also have $\|(R_B(z) - \frac{1}{\lambda-z})\psi_n\| \to 0$ and thus $\|(B-\lambda)\varphi_n\| \to 0$, where $\varphi_n = R_B(z)\psi_n$. Since $\lim_{n\to\infty}\|\varphi_n\| = \lim_{n\to\infty}\|R_A(z)\psi_n\| = |\lambda - z|^{-1} \neq 0$ we obtain a singular Weyl sequence for B, showing $\lambda \in \sigma_{ess}(B)$. Now interchange the roles of A and B. □

As a first consequence note the following result

Theorem 8.12. *Suppose A is symmetric with equal finite defect indices, then all self-adjoint extensions have the same essential spectrum.*

Proof. By Lemma 2.26 the resolvent difference of two self-adjoint extensions is a finite rank operator if the defect indices are finite. □

In addition, the following result is of interest.

Lemma 8.13. *Suppose*

$$R_A(z) - R_B(z) \in \mathfrak{C}(\mathfrak{H}) \tag{8.25}$$

for one $z \in \rho(A)\cap\rho(B)$, then this holds for all $z \in \rho(A)\cap\rho(B)$. In addition, if A and B are self-adjoint, then

$$f(A) - f(B) \in \mathfrak{C}(\mathfrak{H}) \tag{8.26}$$

for any $f \in C_\infty(\mathbb{R})$.

Proof. If the condition holds for one z it holds for all since we have (using both resolvent formulas)

$$\begin{aligned} R_A(z') - R_B(z') \\ &= (1 - (z - z')R_B(z'))(R_A(z) - R_B(z))(1 - (z - z')R_A(z')). \end{aligned} \tag{8.27}$$

Let A and B be self-adjoint. The set of all functions f for which the claim holds is a closed $*$-subalgebra of $C_\infty(\mathbb{R})$ (with sup norm). Hence the claim follows from Lemma 4.5. □

Remember that we have called K relatively compact with respect to A if $KR_A(z)$ is compact (for one and hence for all z) and note that the the resolvent difference $R_{A+K}(z) - R_A(z)$ is compact if K is relatively compact. In particular, Theorem 8.11 applies if $B = A + K$, where K is relatively compact.

For later use observe that set of all operators which are relatively compact with respect to A forms a linear space (since compact operators do) and relatively compact operators have A-bound zero.

Lemma 8.14. *Let A be self-adjoint and suppose K is relatively compact with respect to A. Then the A-bound of K is zero.*

Proof. Write

$$KR_A(\lambda \mathrm{i}) = (KR_A(\mathrm{i}))((A + \mathrm{i})R_A(\lambda \mathrm{i})) \tag{8.28}$$

and observe that the first operator is compact and the second is normal and converges strongly to 0 (by the spectral theorem). Hence the claim follows from Lemma 8.3 and Lemma 8.8 (2). □

In addition, note the following result which is a straightforward consequence of the second resolvent identity.

Lemma 8.15. *Suppose A is self-adjoint and B is symmetric with A-bound less then one. If K is relatively compact with respect to A then it is also relatively compact with respect to $A + B$.*

Proof. Since B is A bounded with A-bound less than one, we can choose a $z \in \mathbb{C}$ such that $\|BR_A(z)\| < 1$. And hence

$$BR_{A+B}(z) = BR_A(z)(\mathbb{I} + BR_A(z))^{-1} \tag{8.29}$$

shows that B is also $A + B$ bounded and the result follows from

$$KR_{A+B}(z) = KR_A(z)(\mathbb{I} - BR_{A+B}(z)) \tag{8.30}$$

since $KR_A(z)$ is compact and $BR_{A+B}(z)$ is bounded. □

8.4. One-particle Schrödinger operators

Our next goal is to apply these results to Schrödinger operators. The Hamiltonian of one particle in d dimensions is given by

$$H = H_0 + V, \tag{8.31}$$

where $V : \mathbb{R}^d \to \mathbb{R}$ is the potential energy of the particle. We are mainly interested in the case $1 \le d \le 3$ and want to find classes of potentials which are relatively bounded respectively relatively compact. To do this we need a better understanding of the functions in the domain of H_0.

Lemma 8.16. *Suppose $n \le 3$ and $\psi \in H^2(\mathbb{R}^n)$. Then $\psi \in C_\infty(\mathbb{R}^n)$ and for any $a > 0$ there is a $b > 0$ such that*

$$\|\psi\|_\infty \le a\|H_0\psi\| + b\|\psi\|. \tag{8.32}$$

Proof. The important observation is that $(p^2+\gamma^2)^{-1} \in L^2(\mathbb{R}^n)$ if $n \le 3$. Hence, since $(p^2+\gamma^2)\hat{\psi} \in L^2(\mathbb{R}^n)$, the Cauchy-Schwarz inequality

$$\begin{aligned} \|\hat{\psi}\|_1 &\le \|(p^2+\gamma^2)^{-1}(p^2+\gamma^2)\hat{\psi}(p)\|_1 \\ &\le \|(p^2+\gamma^2)^{-1}\| \, \|(p^2+\gamma^2)\hat{\psi}(p)\|. \end{aligned} \tag{8.33}$$

shows $\hat{\psi} \in L^1(\mathbb{R}^n)$. But now everything follows from the Riemann-Lebesgue lemma

$$\begin{aligned} \|\psi\|_\infty &\le (2\pi)^{-n/2}\|(p^2+\gamma^2)^{-1}\|(\|p^2\hat{\psi}(p)\| + \gamma^2\|\hat{\psi}(p)\|) \\ &= (\gamma/2\pi)^{n/2}\|(p^2+1)^{-1}\|(\gamma^{-2}\|H_0\psi\| + \|\psi\|) \end{aligned} \tag{8.34}$$

finishes the proof. □

Now we come to our first result.

Theorem 8.17. *Let V be real-valued and $V \in L^\infty_\infty(\mathbb{R}^n)$ or $n \le 3$ and $V \in L^\infty_\infty(\mathbb{R}^n) + L^2(\mathbb{R}^n)$. Then V is relatively compact with respect to H_0. In particular,*

$$H = H_0 + V, \qquad \mathfrak{D}(H) = H^2(\mathbb{R}^n), \tag{8.35}$$

is self-adjoint, bounded from below and

$$\sigma_{ess}(H) = [0, \infty). \tag{8.36}$$

Proof. Our previous lemma shows $\mathfrak{D}(H_0) \subseteq \mathfrak{D}(V)$ and the rest follows from Lemma 6.9 using $f(p) = (P^2 - z)^{-1}$ and $g(x) = V(x)$. □

Observe that since $C_0^\infty(\mathbb{R}^n) \subseteq \mathfrak{D}(H_0)$, we must have $V \in L^2_{loc}(\mathbb{R}^n)$ if $\mathfrak{D}(V) \subseteq \mathfrak{D}(H_0)$.

8.5. Sturm-Liouville operators

In this section we want to illustrate some of the results obtained thus far by investigating a specific example, the **Sturm-Liouville equations**.

$$\tau f(x) = \frac{1}{r(x)} \left(-\frac{d}{dx} p(x) \frac{d}{dx} f(x) + q(x) f(x) \right), \quad f, pf' \in AC_{loc}(I) \tag{8.37}$$

The case $p = r = 1$ can be viewed as the model of a particle in one dimension in the external potential q. Moreover, the case of a particle in three dimensions can in some situations be reduced to the investigation of Sturm-Liouville equations. In particular, we will see how this works when explicitly solving the hydrogen atom.

The suitable Hilbert space is

$$L^2((a,b), r(x)dx), \qquad \langle f, g\rangle = \int_a^b \overline{f(x)} g(x) r(x) dx, \tag{8.38}$$

where $I = (a, b) \subset \mathbb{R}$ is an arbitrary open interval.

We require

(1) $p \in AC_{loc}(I)$, $p' \in L^2_{loc}(I)$, $p^{-1} \in L^1_{loc}(I)$, real-valued
(2) $q \in L^2_{loc}(I)$, real-valued
(3) $r \in L^1_{loc}(I)$, $r^{-1} \in L^\infty_{loc}(I)$, positive

If a is finite and if $p^{-1}, q, r \in L^1((a, c))$ $(c \in I)$, then the Sturm-Liouville equation (8.37) is called regular at a. Similarly for b. If it is both regular at a and b it is called **regular**.

The maximal domain of definition for τ in $L^2(I, r\,dx)$ is given by

$$\mathfrak{D}(\tau) = \{f \in L^2(I, r\,dx) | f, pf' \in AC_{loc}(I),\ \tau f \in L^2(I, r\,dx)\}. \tag{8.39}$$

Since $C_0^\infty(I) \subset \mathfrak{D}(\tau)$ we infer that $\mathfrak{D}(\tau)$ is dense. (Let me remark that it suffices to require $p^{-1}, q, r \in L^1_{loc}(I)$, however, in this case it is no longer obvious that $\mathfrak{D}(\tau)$ is dense.)

Since we are interested in self-adjoint operators H associated with (8.37), we perform a little calculation. Using integration by parts (twice) we obtain $(a < c < d < b)$:

$$\int_c^d \overline{g} (\tau f)\, r dt = W_d(\overline{g}, f) - W_c(\overline{g}, f) + \int_c^d (\overline{\tau g}) f\, r dt, \quad f, g \in AC_{loc}(I), \tag{8.40}$$

where

$$W_x(f_1, f_2) = \Big(p(f_1 f_2' - f_1' f_2)\Big)(x) \tag{8.41}$$

is called the **modified Wronskian**.

It is straightforward to check that the Wronskian of two solutions of $\tau u = zu$ is constant

$$W_x(u_1, u_2) = W(u_1, u_2), \qquad \tau u_{1,2} = z u_{1,2}. \tag{8.42}$$

Moreover, it is nonzero if and only if u_1 and u_2 are linearly independent (compare Theorem 8.18 below).

If we choose $f, g \in \mathfrak{D}(\tau)$ in (8.40), than we can take the limits $c \to a$ and $d \to b$, which results in

$$\langle g, \tau f \rangle = W_b(\overline{g}, f) - W_a(\overline{g}, f) + \langle \tau g, f \rangle, \qquad f, g \in \mathfrak{D}(\tau). \tag{8.43}$$

Here $W_{a,b}(\overline{g}, f)$ has to be understood as limit.

Finally, we recall the following well-known result from ordinary differential equations.

Theorem 8.18. *Suppose $rg \in L^1_{loc}(I)$, then there exists a unique solution $f, pf' \in AC_{loc}(I)$ of the differential equation*

$$(\tau - z) f = g, \qquad z \in \mathbb{C}, \tag{8.44}$$

satisfying the initial condition

$$f(c) = \alpha, \quad (pf')(c) = \beta, \qquad \alpha, \beta \in \mathbb{C}, \quad c \in I. \tag{8.45}$$

Note that f, pf' can be extended continuously to a regular end point.

Lemma 8.19. *Suppose u_1, u_2 are two solutions of $(\tau - z)u = 0$ with $W(u_1, u_2) = 1$. Then every other solution of (8.44) can be written as ($\alpha, \beta \in \mathbb{C}$)*

$$\begin{aligned} f(x) &= u_1(x)\Big(\alpha + \int_c^x u_2 g r\, dt\Big) + u_2(x)\Big(\beta - \int_c^x u_1 g r\, dt\Big), \\ f'(x) &= u_1'(x)\Big(\alpha + \int_c^x u_2 g r\, dt\Big) + u_2'(x)\Big(\beta - \int_c^x u_1 g r\, dt\Big). \end{aligned} \tag{8.46}$$

Note that the constants α, β coincide with those from Theorem 8.18 if $u_1(c) = p(c)u_2'(c) = 1$ and $p(c)u_1'(c) = u_2(c) = 0$.

Proof. It suffices to show $\tau f - z\, f = g$ (the rest follows from 8.18). Differentiating the first equation of (8.46) gives the second. Next we compute

$$\begin{aligned} (pf')' &= (pu_1')'\Big(\alpha + \int u_2 g r\, dt\Big) + (pu_2')'\Big(\beta - \int u_1 g r\, dt\Big) - W(u_1, u_2) g r \\ &= (q - z)u_1\Big(\alpha + \int u_2 g r\, dt\Big) + (q - z)u_2\Big(\beta - \int u_1 g\, dt\Big) - gr \\ &= (q - z) f - gr \end{aligned} \tag{8.47}$$

which proves the claim. □

Now we want to obtain a symmetric operator and hence we choose

$$A_0 f = \tau f, \qquad \mathfrak{D}(A_0) = \mathfrak{D}(\tau) \cap AC_0(I), \tag{8.48}$$

where $AC_0(I)$ are the functions in $AC(I)$ with compact support. This definition clearly ensures that the Wronskian of two such functions vanishes on the boundary, implying that A_0 is symmetric. Our first task is to compute the closure of A_0 and its adjoint. For this the following elementary fact will be needed.

Lemma 8.20. *Suppose V is a vector space and $l, l_1, \dots, l_n$ are linear functionals (defined on all of V) such that $\bigcap_{j=1}^n \mathrm{Ker}(l_j) \subseteq \mathrm{Ker}(l)$. Then $l = \sum_{j=0}^n \alpha_j l_j$ for some constants $\alpha_j \in \mathbb{C}$.*

Proof. First of all it is no restriction to assume that the functionals l_j are linearly independent. Taking a dual basis $f_k \in V$, that is $l_j(f_k) = 0$ for $j \neq k$ and $l_j(f_j) = 1$. Then $f - \sum_{j=1}^n l_j(f) f_j \in \bigcap_{j=1}^n \mathrm{Ker}(l_j)$ and hence $l(f) - \sum_{j=1}^n l_j(f) l(f_j) = 0$. Thus we can choose $\alpha_j = l(f_j)$. □

Now we are ready to prove

Lemma 8.21. *The closure of A_0 is given by*

$$\overline{A_0} f = \tau f, \quad \mathfrak{D}(\overline{A_0}) = \{f \in \mathfrak{D}(\tau) \,|\, W_a(f,g) = W_b(f,g) = 0,\ \forall g \in \mathfrak{D}(\tau)\}. \tag{8.49}$$

Its adjoint is given by

$$A_0^* f = \tau f, \quad \mathfrak{D}(A_0^*) = \mathfrak{D}(\tau). \tag{8.50}$$

Proof. We start by computing A_0^*. By (8.43) we have $\mathfrak{D}(\tau) \subseteq \mathfrak{D}(A_0^*)$ and it remains to show $\mathfrak{D}(A_0^*) \subseteq \mathfrak{D}(\tau)$. If $h \in \mathfrak{D}(A_0^*)$ we must have

$$\langle h, A_0 f\rangle = \langle k, f\rangle, \qquad \forall f \in \mathfrak{D}(A_0) \tag{8.51}$$

for some $k \in L^2(I, r\,dx)$. Using (8.46) we can find a $\tilde{h}$ such that $\tau \tilde{h} = k$ and from integration by parts we obtain

$$\int_a^b (h(x) - \tilde{h}(x))^* (\tau f)(x) r(x) dx = 0, \qquad \forall f \in \mathfrak{D}(A_0). \tag{8.52}$$

Clearly we expect that $h - \tilde{h}$ will be a solution of the $\tau u = 0$ and to prove this we will invoke Lemma 8.20. Therefore we consider the linear functionals

$$l(f) = \int_a^b (h(x) - \tilde{h}(x))^* g(x) r(x) dx, \quad l_j(f) = \int_a^b u_j(x)^* g(x) r(x) dx, \tag{8.53}$$

on $L^2_0(I, r\,dx)$, where u_j are two solutions of $\tau u = 0$ with $W(u_1, u_2) \neq 0$. We have $\mathrm{Ker}(l_1) \cap \mathrm{Ker}(l_2) \subseteq \mathrm{Ker}(l)$. In fact, if $g \in \mathrm{Ker}(l_1) \cap \mathrm{Ker}(l_2)$, then

$$f(x) = u_1(x) \int_a^x u_2(t) g(t) r(t) dt + u_2(x) \int_x^b u_1(t) g(t) r(t) dt \tag{8.54}$$

is in $\mathfrak{D}(A_0)$ and $g = \tau f \in \mathrm{Ker}(l)$ by (8.53). Now Lemma 8.20 implies

$$\int_a^b (h(x) - \tilde{h}(x) + \alpha_1 u_1(x) + \alpha_2 u_2(x))^* g(x) r(x) dx = 0, \quad \forall g \in L^2_0(I, r dx) \tag{8.55}$$

and hence $h = \tilde{h} + \alpha_1 u_1 + \alpha_2 u_2 \in \mathfrak{D}(\tau)$.

Now we turn to $\overline{A_0}$. Denote the set on the right hand side of (8.49) by $\mathfrak{D}$. Then we have $\mathfrak{D} \subseteq \mathfrak{D}(A_0^{**}) = \overline{A}_0$ by (8.43). Conversely, since $\overline{A_0} \subseteq A_0^*$ we can use (8.43) to conclude

$$W_a(f, h) + W_b(f, h) = 0, \quad f \in \mathfrak{D}(\overline{A_0}), h \in \mathfrak{D}(A_0^*). \tag{8.56}$$

Now replace h by a $\tilde{h} \in \mathfrak{D}(A_0^*)$ which coincides with h near a and vanishes identically near b. Then $W_a(f, h) = W_a(f, \tilde{h}) + W_b(f, \tilde{h}) = 0$. Finally, $W_b(f, h) = -W_a(f, h) = 0$ shows $f \in \mathfrak{D}$. □

This result shows that any self-adjoint extension of A_0 must lie between $\overline{A}_0$ and A_0^*. Moreover, self-adjointness seems to be related to the Wronskian of two functions at the boundary. Hence we collect a few properties first.

Lemma 8.22. *Suppose $v \in \mathfrak{D}(\tau)$ with $W_a(\overline{v}, v) = 0$ and there is a $\hat{f} \in \mathfrak{D}(\tau)$ with $W(\overline{v}, \hat{f})_a \neq 0$. then we have*

$$W_a(v, f) = 0 \quad \Leftrightarrow \quad W_a(v, \overline{f}) = 0 \qquad \forall f \in \mathfrak{D}(\tau) \tag{8.57}$$

and

$$W_a(v, f) = W_a(v, g) = 0 \quad \Rightarrow \quad W_a(\overline{g}, f) = 0 \qquad \forall f, g \in \mathfrak{D}(\tau) \tag{8.58}$$

Proof. For all $f_1, \dots, f_4 \in \mathfrak{D}(\tau)$ we have the Plücker identity

$$W_x(f_1, f_2) W_x(f_3, f_4) + W_x(f_1, f_3) W_x(f_4, f_2) + W_x(f_1, f_4) W_x(f_2, f_3) = 0 \tag{8.59}$$

which remains valid in the limit $x \to a$. Choosing $f_1 = v$, $f_2 = f$, $f_3 = \overline{v}$, $f_4 = \hat{f}$ we infer (8.57). Choosing $f_1 = f$, $f_2 = \overline{g}$, $f_3 = v$, $f_4 = \hat{f}$ we infer (8.58). □

We call τ **limit circle** (l.c.) at a if there is a $v \in \mathfrak{D}(\tau)$ with $W_a(\overline{v}, v) = 0$ such that $W_a(\overline{v}, f) \neq 0$ for at least one $f \in \mathfrak{D}(\tau)$. Otherwise τ is called **limit point** (l.p.) at a. Similarly for b.

Suppose, $W_a(f, v) \neq 0$, then $W_a(f, \mathrm{Re}(v)) \neq 0$ or $W_a(f, \mathrm{Im}(v)) \neq 0$ and hence τ is l.c. at a.

Theorem 8.23. *If τ is l.c. at a, then let $v \in \mathfrak{D}(\tau)$ with $W(\overline{v}, v)_a = 0$ and $W(v, f)_a \neq 0$ for some $f \in \mathfrak{D}(\tau)$. Similarly, if τ is l.c. at b, let w be an analogous function. Then the operator*

$$\begin{array}{rccc} A: & \mathfrak{D}(A) & \to & L^2(I, r\,dx) \\ & f & \mapsto & \tau f \end{array} \tag{8.60}$$

with

$$\mathfrak{D}(A) = \{f \in \mathfrak{D}(\tau) | \begin{array}{l} W(v, f)_a = 0 \text{ if l.c. at } a \\ W(w, f)_b = 0 \text{ if l.c. at } b\} \end{array} \tag{8.61}$$

is self-adjoint.

Proof. Clearly $A \subseteq A^* \subseteq A_0^*$. As in the computation of $\overline{A_0}$ we conclude $W_a(f, g) = 0$ for all $f \in \mathfrak{D}(A)$, $g \in \mathfrak{D}(A^*)$. Moreover, by the lemma we have $W_a(v, g) = 0$ since $W_a(v, f) = 0$. Thus $g \in \mathfrak{D}(A)$. □

The name limit circle respectively limit point stems from the original approach of Weyl, who considered the set of solutions $\tau u = zu$, $z \in \mathbb{C}\backslash\mathbb{R}$ which satisfy $W_c(u^*, u) = 0$. They can be shown to lie on a circle which converges to a circle respectively point as $c \to a$ (or $c \to b$).

Next we want to compute the resolvent of A.

Lemma 8.24. *Suppose $z \in \rho(A)$, then there exists a solution $u_a(z, x)$ which is in $L^2((a, c), r\,dx)$ and which satisfies the boundary condition at a if τ is l.c. at a. Similarly, there exists a solution $u_b(z, x)$ with the analogous properties near b.*

The resolvent of A is given by

$$(A - z)^{-1} g(x) = \int_a^b G(z, x, t) g(t) r(t) dt, \tag{8.62}$$

where

$$G(z, x, t) = \frac{1}{W(u_b(z), u_a(z))} \begin{cases} u_b(z, x) u_a(z, t) & x \geq t \\ u_a(z, x) u_b(z, t) & x \leq t \end{cases}. \tag{8.63}$$

Proof. Let $g \in L_0^2(I, r\,dx)$ and consider $f = (A - z)^{-1} g \in \mathfrak{D}(A)$. Since $\tau f = 0$ near a respectively b, we obtain $u_a(z, x)$ by setting it equal to f near a and using the differential equation to extend it to the rest of I. Similarly we obtain u_b. The only problem is that u_a or u_b might be identically zero. Hence we need to show that this can be avoided by choosing g properly.

Let g be supported in $(c, d) \subset I$. Since $\tau f = g$ we have

$$f(x) = u_1(x) \left(\alpha + \int_a^x u_2 g r\, dt \right) + u_2(x) \left(\beta + \int_x^b u_1 g r\, dt \right). \tag{8.64}$$

Near a ($x < c$) we have $f(x) = \alpha u_1(x) + \tilde{\beta} u_2(x)$ and near b ($x > d$) we have $f(x) = \tilde{\alpha} u_1(x) + \beta u_2(x)$, where $\tilde{\alpha} = \alpha + \int_a^b u_2 g r\, dt$ and $\tilde{\beta} = \beta + \int_a^b u_1 g r\, dt$.

If f vanishes identically near both a and b we must have $\alpha = \beta = \tilde{\alpha} = \tilde{\beta} = 0$ and thus $\alpha = \beta = 0$ and $\int_a^b u_j(t) g(t) r(t) dt = 0$, $j = 1, 2$. This case can be avoided choosing g suitable and hence there is at least one solution, say $u_b(z)$.

Now choose $u_1 = u_b$ and consider the behavior near b. If u_2 is not square integrable on (d, b), we must have $\beta = 0$ since $\beta u_2 = f - \tilde{\alpha} u_b$ is. If u_2 is square integrable, we can find two functions in $\mathfrak{D}(\tau)$ which coincide with u_b and u_2 near b. Since $W(u_b, u_2) = 1$ we see that τ is l.c. at a and hence $0 = W_b(u_b, f) = W_b(u_b, \tilde{\alpha} u_b + \beta u_2) = \beta$. Thus $\beta = 0$ in both cases and we have

$$f(x) = u_b(x) \left(\alpha + \int_a^x u_2 g r \, dt \right) + u_2(x) \int_x^b u_b g r \, dt. \qquad (8.65)$$

Now choosing g such that $\int_a^b u_b g r \, dt \neq 0$ we infer existence of $u_a(z)$. Choosing $u_2 = u_a$ and arguing as before we see $\alpha = 0$ and hence

$$\begin{aligned} f(x) &= u_b(x) \int_a^x u_a(t) g(t) r(t) dt + u_a(x) \int_x^b u_b(t) g(t) r(t) dt \\ &= \int_a^b G(z, x, t) g(t) r(t) dt \end{aligned} \qquad (8.66)$$

for any $g \in L^2_0(I, r\,dx)$. Since this set is dense the claim follows. □

Theorem 8.25 (Weyl alternative)**.** *The operator τ is l.c. at a if and only if for one $z_0 \in \mathbb{C}$ all solutions of $(\tau - z_0)u = 0$ are square integrable near a. This then holds for all $z \in \mathbb{C}$. Similarly for b.*

Proof. If all solutions are square integrable near a, τ is l.c. at a since the Wronskian of two linearly independent solutions does not vanish.

Conversely, take two functions $v, \tilde{v} \in \mathfrak{D}(\tau)$ with $W_a(v, \tilde{v}) \neq 0$. By considering real and imaginary parts it is no restriction th assume that v and $\tilde{v}$ is real-valued. Thus they give rise to two different self-adjoint operators A and $\tilde{A}$ (choose any fixed w for the other endpoint). Let u_a and $\tilde{u_a}$ be the corresponding solutions from above, then $W(u_a, \tilde{u}_a) \neq 0$ (since otherwise $A = \tilde{A}$ by Lemma 8.22) and thus there are two linearly independent solutions which are square integrable near a. Since any other solution can be written as a linear combination of those two, every solution is square integrable near a.

It remains to show that all solutions of $(\tau - z)u = 0$ for all $z \in \mathbb{C}$ are square integrable near a if τ is l.c. at a. In fact, the above argument ensures this for every $z \in \rho(A) \cap \rho(\tilde{A})$, that is, at least for all $z \in \mathbb{C} \backslash \mathbb{R}$.

Suppose $(\tau - z)u = 0.$ and choose two linearly independent solutions u_j, $j = 1, 2$, of $(\tau - z_0)u = 0$ with $W(u_1, u_2) = 1$. Using $(\tau - z_0)u = (z - z_0)u$

and (8.46) we have ($a < c < x < b$)

$$u(x) = \alpha u_1(x) + \beta u_2(x) + (z - z_0)\int_c^x (u_1(x)u_2(t) - u_1(t)u_2(x))u(t)r(t)\,dt. \tag{8.67}$$

Since $u_j \in L^2((c,b), r dx)$ we can find a constant $M \geq 0$ such that

$$\int_c^b |u_{1,2}(t)|^2 r(t)\,dt \leq M. \tag{8.68}$$

Now choose c close to b such that $|z - z_0|M^2 \leq 1/4$. Next, estimating the integral using Cauchy–Schwarz gives

$$\begin{aligned} \Big|\int_c^x (u_1(x)u_2(t) - u_1(t)u_2(x))u(t)r(t)\,dt\Big|^2 \\ \leq \int_c^x |u_1(x)u_2(t) - u_1(t)u_2(x)|^2 r(t)\,dt \int_c^x |u(t)|^2 r(t)\,dt \\ \leq M\Big(|u_1(x)|^2 + |u_2(x)|^2\Big)\int_c^x |u(t)|^2 r(t)\,dt \end{aligned} \tag{8.69}$$

and hence

$$\begin{aligned} \int_c^x |u(t)|^2 r(t)\,dt &\leq (|\alpha|^2 + |\beta|^2)M + 2|z - z_0|M^2 \int_c^x |u(t)|^2 r(t)\,dt \\ &\leq (|\alpha|^2 + |\beta|^2)M + \frac{1}{2}\int_c^x |u(t)|^2 r(t)\,dt. \end{aligned} \tag{8.70}$$

Thus

$$\int_c^x |u(t)|^2 r(t)\,dt \leq 2(|\alpha|^2 + |\beta|^2)M \tag{8.71}$$

and since $u \in AC_{loc}(I)$ we have $u \in L^2((c,b), r\,dx)$ for every $c \in (a,b)$. □

Note that all eigenvalues are simple. If τ is l.p. at one endpoint this is clear, since there is at most one solution of $(\tau - \lambda)u = 0$ which is square integrable near this end point. If τ is l.c. this also follows since the fact that two solutions of $(\tau - \lambda)u = 0$ satisfy the same boundary condition implies that their Wronskian vanishes.

Finally, led us shed some additional light on the number of possible boundary conditions. Suppose τ is l.c. at a and let u_1, u_2 be two solutions of $\tau u = 0$ with $W(u_1, u_2) = 1$. Abbreviate

$$BC_x^j(f) = W_x(u_j, f), \qquad f \in \mathfrak{D}(\tau). \tag{8.72}$$

Let v be as in Theorem 8.23, then, using Lemma 8.22 it is not hard to see that

$$W_a(v,f) = 0 \quad \Leftrightarrow \quad \cos(\alpha)BC_a^1(f) - \sin(\alpha)BC_a^2(f) = 0, \tag{8.73}$$

where $\tan(\alpha) = \frac{BC_a^1(v)}{BC_a^2(v)}$. Hence all possible boundary conditions can be parametrized by $\alpha \in [0, \pi)$. If τ is regular at a and if we choose $u_1(a) = p(a)u_2'(a) = 1$ and $p(a)u_1'(a) = u_2(a) = 0$, then

$$BC_a^1(f) = f(a), \qquad BC_a^2(f) = p(a)f'(a) \tag{8.74}$$

and the boundary condition takes the simple form

$$\cos(\alpha)f(a) - \sin(\alpha)p(a)f'(a) = 0. \tag{8.75}$$

Finally, note that if τ is l.c. at both a and b, then Theorem 8.23 does not give all possible self-adjoint extensions. For example, one could also choose

$$BC_a^1(f) = \mathrm{e}^{\mathrm{i}\alpha} BC_b^1(f), \qquad BC_a^2(f) = \mathrm{e}^{\mathrm{i}\alpha} BC_b^2(f). \tag{8.76}$$

The case $\alpha = 0$ gives rise to periodic boundary conditions in the regular case.

Now we turn to the investigation of the spectrum of A. If τ is l.c. at both endpoints, then the spectrum of A is very simple

Theorem 8.26. *If τ is l.c. at both end points, then the resolvent is a Hilbert–Schmidt operator, that is,*

$$\int_a^b \int_a^b |G(z,x,t)|^2 r(t) dt\, r(x) dx < \infty. \tag{8.77}$$

In particular, the spectrum of any self adjoint extensions is purely discrete and the eigenfunctions (which are simple) form an orthonormal basis.

Proof. This follows from the estimate

$$\begin{aligned} \int_a^b \Big(\int_a^x |u_b(x)u_a(t)|^2 r(t) dt + \int_x^b |u_b(t)u_a(x)|^2 dt \Big) r(x) dx \\ \le 2 \int_a^b |u_a(t)|^2 r(t) dt \int_a^b |u_b(s)|^2 r(s) ds, \end{aligned} \tag{8.78}$$

which shows that the resolvent is Hilbert–Schmidt and hence compact. □

If τ is not l.c. the situation is more complicated and we can only say something about the essential spectrum.

Theorem 8.27. *All self adjoint extensions have the same essential spectrum. Moreover, if A_{ac} and A_{cb} are self-adjoint extensions of τ restricted to (a, c) and (c, b) (for any $c \in I$), then*

$$\sigma_{ess}(A) = \sigma_{ess}(A_{ac}) \cup \sigma_{ess}(A_{cb}). \tag{8.79}$$

Proof. Since $(\tau - \mathrm{i})u = 0$ has two linearly independent solutions, the defect indices are at most two (they are zero if τ is l.p. at both end points, one if τ is l.c. at one and l.p. at the other end point, and two if τ is l.c. at both endpoints). Hence The first claim follows from Theorem 8.12.

For the second claim restrict τ to the functions with compact support in $(a, c) \cup (c, d)$. Then, this operator is the orthogonal sum of the operators $A_{0,ac}$ and $A_{0,cb}$. Hence the same is true for the adjoints and hence the defect indices of $A_{0,ac} \oplus A_{0,cb}$ are at most four. Now note that A and $A_{ac} \oplus A_{cb}$ are both self-adjoint extensions of this operator. Thus the second claim also follows from Theorem 8.12. □

Examples will follow in the next chapter.

Chapter 9

Atomic Schrödinger operators

9.1. The hydrogen atom

We begin with the simple model of a single electron in $\mathbb{R}^3$ moving in the external potential V generated by a nucleus (which is assumed to be fixed at the origin). If one takes only the electrostatic force into account, then V is given by the Coulomb potential and the corresponding Hamiltonian is given by

$$H^{(1)} = -\Delta - \frac{\gamma}{|x|}, \quad \mathfrak{D}(H^{(1)}) = H^2(\mathbb{R}^3). \tag{9.1}$$

If the potential is attracting, that is, if $\gamma > 0$, then it describes the hydrogen atom and is probably the most famous model in quantum mechanics.

As domain we have chosen $\mathfrak{D}(H^{(1)}) = \mathfrak{D}(H_0) \cap \mathfrak{D}(\frac{1}{|x|}) = \mathfrak{D}(H_0)$ and by Theorem 8.17 we conclude that $H^{(1)}$ is self-adjoint. Moreover, Theorem 8.17 also tells us

$$\sigma_{ess}(H^{(1)}) = [0, \infty) \tag{9.2}$$

and that $H^{(1)}$ is bounded from below

$$E_0 = \inf \sigma(H^{(1)}) > -\infty. \tag{9.3}$$

If $\gamma \le 0$ we have $H^{(1)} \ge 0$ and hence $E_0 = 0$, but if $\gamma > 0$, we might have $E_0 < 0$ and there might be some discrete eigenvalues below the essential spectrum.

In order to say more about the eigenvalues of $H^{(1)}$ we will use the fact that both H_0 and $V^{(1)} = -\gamma/|x|$ have a simple behavior with respect to scaling. Consider the **dilation group**

$$U(s)\psi(x) = \mathrm{e}^{-ns/2}\psi(\mathrm{e}^{-s}x), \qquad s \in \mathbb{R}, \tag{9.4}$$

which is a strongly continuous one-parameter unitary group. The generator can be easily computed

$$D\psi(x) = \frac{1}{2}(xp + px)\psi(x) = (xp - \frac{\mathrm{i}n}{2})\psi(x), \qquad \psi \in \mathcal{S}(\mathbb{R}^n). \tag{9.5}$$

Now let us investigate the action of $U(s)$ on $H^{(1)}$

$$H^{(1)}(s) = U(-s)H^{(1)}U(s) = \mathrm{e}^{-2s}H_0 + \mathrm{e}^{-s}V^{(1)}, \quad \mathfrak{D}(H^{(1)}(s)) = \mathfrak{D}(H^{(1)}). \tag{9.6}$$

Now suppose $H\psi = \lambda\psi$, then

$$\langle\psi, [U(s), H]\psi\rangle = \langle U(-s)\psi, H\psi\rangle - \langle H\psi, U(s)\psi\rangle = 0 \tag{9.7}$$

and hence

$$\begin{aligned} 0 &= \lim_{s\to 0} \frac{1}{s}\langle\psi, [U(s), H]\psi\rangle = \lim_{s\to 0}\langle U(-s)\psi, \frac{H - H(s)}{s}\psi\rangle \\ &= \langle\psi, (2H_0 + V^{(1)})\psi\rangle. \end{aligned} \tag{9.8}$$

Thus we have proven the **virial theorem**.

Theorem 9.1. *Suppose $H = H_0 + V$ with $U(-s)VU(s) = \mathrm{e}^{-s}V$. Then any normalized eigenfunction ψ corresponding to an eigenvalue λ satisfies*

$$\lambda = -\langle\psi, H_0\psi\rangle = \frac{1}{2}\langle\psi, V\psi\rangle. \tag{9.9}$$

In particular, all eigenvalues must be negative.

This result even has some further consequences for the point spectrum of $H^{(1)}$.

Corollary 9.2. *Suppose $\gamma > 0$. Then*

$$\sigma_p(H^{(1)}) = \sigma_d(H^{(1)}) = \{E_{j-1}\}_{j\in\mathbb{N}_0}, \qquad E_0 < E_j < E_{j+1} < 0, \tag{9.10}$$

with $\lim_{j\to\infty} E_j = 0$.

Proof. Choose $\psi \in C_0^\infty(\mathbb{R}\backslash\{0\})$ and set $\psi(s) = U(-s)\psi$. Then

$$\langle\psi(s), H^{(1)}\psi(s)\rangle = \mathrm{e}^{-2s}\langle\psi, H_0\psi\rangle + \mathrm{e}^{-s}\langle\psi, V^{(1)}\psi\rangle \tag{9.11}$$

which is negative for s large. Now choose a sequence $s_n \to \infty$ such that we have $\mathrm{supp}(\psi(s_n)) \cap \mathrm{supp}(\psi(s_m)) = \emptyset$ for $n \neq m$. Then Theorem 4.12 (i) shows that $\mathrm{rank}(P_{H^{(1)}}((-\infty, 0))) = \infty$. Since each eigenvalue E_j has finite multiplicity (it lies in the discrete spectrum) there must be an infinite number of eigenvalues which accumulate at 0. □

If $\gamma \le 0$ we have $\sigma_d(H^{(1)}) = \emptyset$ since $H^{(1)} \ge 0$ in this case.

Hence we have gotten a quite complete picture of the spectrum of $H^{(1)}$. Next, we could try to compute the eigenvalues of $H^{(1)}$ (in the case $\gamma > 0$) by solving the corresponding eigenvalue equation, which is given by the partial differential equation

$$-\Delta\psi(x) - \frac{\gamma}{|x|}\psi(x) = \lambda\psi(x). \tag{9.12}$$

For a general potential this is hopeless, but in our case we can use the rotational symmetry of our operator to reduce our partial differential equation to ordinary ones.

First of all, it suggests itself to switch to spherical coordinates $(x_1, x_2, x_3) \to (r, \theta, \varphi)$ which correspond to a unitary transform

$$L^2(\mathbb{R}^3) \to L^2((0,\infty), r^2 dr) \otimes L^2((0,\pi), \sin(\theta) d\theta) \otimes L^2((0,2\pi), d\varphi) \tag{9.13}$$

In these new coordinates (r, θ, φ) our operator reads

$$H^{(1)} = -\frac{1}{r^2}\frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{1}{r^2} L^2 + V(r), \quad V(r) = -\frac{\gamma}{r}, \tag{9.14}$$

where

$$L^2 = L_1^2 + L_2^2 + L_3^2 = -\frac{1}{\sin(\theta)}\frac{\partial}{\partial\theta}\sin(\theta)\frac{\partial}{\partial\theta} - \frac{1}{\sin(\theta)^2}\frac{\partial^2}{\partial\varphi^2}. \tag{9.15}$$

(Recall the angular momentum operators L_j from Section 7.2.)

Making the product ansatz (separation of variables)

$$\psi(r, \theta, \varphi) = R(r)\Theta(\theta)\Phi(\varphi) \tag{9.16}$$

we obtain the following three Sturm-Liouville equations

$$\begin{aligned}
\left(-\frac{1}{r^2}\frac{d}{dr} r^2 \frac{d}{dr} + \frac{l(l+1)}{r^2} + V(r)\right) R(r) &= \lambda R(r) \\
\frac{1}{\sin(\theta)}\left(-\frac{d}{d\theta}\sin(\theta)\frac{d}{d\theta} + \frac{m^2}{\sin(\theta)}\right)\Theta(\theta) &= l(l+1)\Theta(\theta) \\
-\frac{d^2}{d\varphi^2}\Phi(\varphi) &= m^2\Phi(\varphi)
\end{aligned} \tag{9.17}$$

The form chosen for the constants $l(l+1)$ and m^2 is for convenience later on. These equations will be investigated in the following sections.

9.2. Angular momentum

We start by investigating the equation for $\Phi(\varphi)$ which associated with the Stum-Liouville equation

$$\tau\Phi = -\Phi'', \qquad I = (0, 2\pi). \tag{9.18}$$

since we want ψ defined via (9.16) to be in the domain of H_0 (in particular continuous), we choose periodic boundary conditions the Stum-Liouville equation

$$A\Phi = \tau\Phi, \quad \mathfrak{D}(A) = \{\Phi \in L^2(0,\pi) | \begin{array}{l} \Phi \in AC^1[0,\pi], \\ \Phi(0) = \Phi(2\pi), \Phi'(0) = \Phi'(2\pi)\} \end{array}. \tag{9.19}$$

From our analysis in Section 8.5 we immediately obtain

Theorem 9.3. *The operator A defined via (9.18) is self-adjoint. Its spectrum is purely discrete*

$$\sigma(A) = \sigma_d(A) = \{m^2 | m \in \mathbb{Z}\} \tag{9.20}$$

and the corresponding eigenfunctions

$$\Phi_m(\varphi) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{\mathrm{i}m\varphi}, \qquad m \in \mathbb{Z}, \tag{9.21}$$

form an orthonormal basis for $L^2(0, 2\pi)$.

Note that except for the lowest eigenvalue, all eigenvalues are twice degenerate.

We note that this operator is essentially the square of the angular momentum in the third coordinate direction, since in polar coordinates

$$L_3 = \frac{1}{\mathrm{i}} \frac{\partial}{\partial\varphi}. \tag{9.22}$$

Now we turn to the equation for $\Theta(\theta)$

$$\tau_m \Theta(\theta) = \frac{1}{\sin(\theta)} \left(-\frac{d}{d\theta} \sin(\theta) \frac{d}{d\theta} + \frac{m^2}{\sin(\theta)} \right) \Theta(\theta), \quad I = (0,\pi), m \in \mathbb{N}_0. \tag{9.23}$$

For the investigation of the corresponding operator we use the unitary transform

$$L^2((0,\pi), \sin(\theta) d\theta) \to L^2((-1,1), dx), \quad \Theta(\theta) \mapsto f(x) = \Theta(\arccos(x)). \tag{9.24}$$

The operator τ transforms to the somewhat simpler form

$$\tau_m = -\frac{d}{dx}(1-x^2)\frac{d}{dx} - \frac{m^2}{1-x^2}. \tag{9.25}$$

The corresponding eigenvalue equation

$$\tau_m u = l(l-1)u \tag{9.26}$$

is the **associated Legendre equation**. For $l \in \mathbb{N}_0$ it is solved by the **associated Legendre functions**

$$P_{lm}(x) = (1-x)^{m/2} \frac{d^m}{dx^m} P_l(x), \tag{9.27}$$

where

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l}(1-x^2) \tag{9.28}$$

are the **Legendre polynomials**. This is straightforward to check. Moreover, note that $P_l(x)$ are (nonzero) polynomials of degree l. A second, linearly independent solution is given by

$$Q_{lm}(x) = P_{lm}(x) \int_0^x \frac{dt}{(1-t^2)P_{lm}(t)^2}. \tag{9.29}$$

In fact, for every Sturm-Liouville equation $v(x) = \int^x \frac{dt}{p(t)u(t)^2}$ satisfies $\tau v = 0$ whenever $\tau u = 0$. Now fix $l = 0$ and note $P_0(x) = 1$. For $m = 0$ we have $Q_{00} = \mathrm{arctanh}(x) \in L^2$ and so τ_0 is l.c. at both end points. For $m > 0$ we have $Q_{0m} = (x \pm 1)^{-m/2}(C + O(x \pm 1))$ which shows that it is not square integrable. Thus τ_m is l.c. for $m = 0$ and l.p. for $m > 0$ at both endpoints. In order to make sure that the eigenfunctions for $m = 0$ are continuous (such that ψ defined via (9.16) is continuous) we choose the boundary condition generated by $P_0(x) = 1$ in this case

$$A_m f = \tau f, \quad \mathfrak{D}(A_m) = \{f \in L^2(-1,1) |\ \begin{array}{l} f \in AC^1(0,\pi), \tau f \in L^2(-1,1) \\ \lim_{x\to\pm 1}(1-x^2)f'(x) = 0 \end{array}\}. \tag{9.30}$$

Theorem 9.4. *The operator A_m, $m \in \mathbb{N}_0$, defined via (9.30) is self-adjoint. Its spectrum is purely discrete*

$$\sigma(A_m) = \sigma_d(A_m) = \{l(l+1) | l \in \mathbb{N}_0, l \geq m\} \tag{9.31}$$

and the corresponding eigenfunctions

$$u_{lm}(x) = \sqrt{\frac{2l+1}{2}\frac{(l+m)!}{(l-m)!}} P_{lm}(x), \qquad l \in \mathbb{N}_0, l \geq m, \tag{9.32}$$

form an orthonormal basis for $L^2(-1,1)$.

Proof. By Theorem 8.23, A_m is self-adjoint. Moreover, P_{lm} is an eigenfunction corresponding to the eigenvalue $l(l+1)$ and it suffices to show that P_{lm} form a basis. To prove this, it suffices to show that the functions $P_{lm}(x)$ are dense. Since $(1-x^2) > 0$ for $x \in (-1,1)$ it suffices to show that the functions $(1-x^2)^{-m/2} P_{lm}(x)$ are dense. But the span of these functions contains every polynomial. Every continuous function can be approximated by polynomials (in the sup norm and hence in the L^2 norm) and since the continuous functions are dense, so are the polynomials.

The only thing remaining is the normalization of the eigenfunctions, which can be found in any book on special functions. □

Returning to our original setting we conclude that

$$\Theta_{lm} = \sqrt{\frac{2l+1}{2}\frac{(l+m)!}{(l-m)!}} P_{lm}(\cos(\theta)), l = m, m+1, \dots \tag{9.33}$$

form an orthonormal basis for $L^2((0,\pi), \sin(\theta)d\theta)$ for any fixed $m \in \mathbb{N}_0$.

Theorem 9.5. *The operator L^2 on $L^2((0,\pi), \sin(\theta)d\theta) \otimes L^2((0,2\pi))$ has a purely discrete spectrum given*

$$\sigma(L^2) = \{l(l+1) | l \in \mathbb{N}_0\}. \tag{9.34}$$

The **spherical harmonics**

$$Y_{lm}(\theta,\varphi) = \Theta_{l|m|}(\theta)\Phi_m(\varphi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l+|m|)!}{(l-|m|)!}} P_{l|m|}(\cos(\theta))\mathrm{e}^{\mathrm{i}m\varphi}, \quad |m| \le l, \tag{9.35}$$

form an orthonormal basis and satisfy $L^2 Y_{lm} = l(l+1)Y_{lm}$ and $L_3 Y_{lm} = mY_{lm}$.

Proof. Everything follows from our construction, if we can show that Y_{lm} form a basis. But this follows as in the proof of Lemma 1.8. □

Note that transforming Y_{lm} back to cartesian coordinates gives

$$Y_{l,\pm m}(x) = \sqrt{\frac{2l+1}{4\pi}\frac{(l+m)!}{(l-m)!}} \tilde{P}_{lm}(\frac{x_3}{r}) \left(\frac{x_1 \pm \mathrm{i}x_2}{r}\right)^m, \quad r = |x|, \tag{9.36}$$

where $\tilde{P}_{lm}$ is a polynomial of degree $l - m$ given by

$$\tilde{P}_{lm}(x) = (1-x^2)^{-m/2} P_{lm}(x) = \frac{d^{l+m}}{dx^{l+m}}(1-x^2)^l. \tag{9.37}$$

In particular, Y_{lm} are smooth away from the origin and by construction they satisfy

$$-\Delta Y_{lm} = \frac{l(l+1)}{r^2} Y_{lm}. \tag{9.38}$$

9.3. The spectrum of the hydrogen atom

Now we want to use the considerations from the previous section to decompose the Hamiltonian of the hydrogen atom. In fact, we can even admit any potential $V(x) = V(|x|)$ with

$$V(r) \in L^\infty_\infty(\mathbb{R}) + L^2((0,\infty), r^2 dr). \tag{9.39}$$

The important observation is that the spaces

$$\mathfrak{H}_{lm} = \{\psi(x) = R(r)Y_{lm}(\theta,\varphi) | R(r) \in L^2((0,\infty), r^2 dr)\} \tag{9.40}$$

reduce our operator $H = H_0 + V$. Hence

$$H = H_0 + V = \bigoplus_{l,m} \tilde{H}_l, \tag{9.41}$$

where

$$\begin{aligned} \tilde{H}_l R(r) &= \tilde{\tau}_l R(r), \quad \tilde{\tau}_l = -\frac{1}{r^2}\frac{d}{dr}r^2\frac{d}{dr} + \frac{l(l+1)}{r^2} + V(r) \\ \mathfrak{D}(H_l) &\subseteq L^2((0,\infty), r^2 dr). \end{aligned} \tag{9.42}$$

Using the unitary transformation

$$L^2((0,\infty), r^2 dr) \to L^2((0,\infty)), \qquad R(r) \mapsto u(r) = rR(r), \tag{9.43}$$

our operator transforms to

$$\begin{aligned} A_l f &= \tau_l f, \quad \tau_l = -\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} + V(r) \\ \mathfrak{D}(A_l) &\subseteq L^2((0,\infty)). \end{aligned} \tag{9.44}$$

It remains to investigate this operator.

Theorem 9.6. *The domain of the operator A_l is given by*

$$\begin{aligned} \mathfrak{D}(A_l) = \{f \in L^2(I) |\ & f, f' \in AC(I), \tau f \in L^2(I), \\ & \lim_{r\to 0}(f(r) - r f'(r)) = 0 \text{ if } l = 0\}, \end{aligned} \tag{9.45}$$

where $I = (0,\infty)$. Moreover, $\sigma_{ess}(A_l) = [0,\infty)$.

Proof. By construction of A_l we know that it is self-adjoint and satisfies $\sigma_{ess}(A_l) = [0,\infty)$. Hence it remains to compute the domain. We know at least $\mathfrak{D}(A_l) \subseteq \mathfrak{D}(\tau)$ and since $\mathfrak{D}(H) = \mathfrak{D}(H_0)$ it suffices to consider the case $V = 0$. In this case the solutions of $-u''(r) + \frac{l(l+1)}{r^2}u(r) = 0$ are given by $u(r) = \alpha r^{l+1} + \beta r^{-l}$. Thus we are in the l.p. case at ∞ for any $l \in \mathbb{N}_0$. However, at 0 we are in the l.p. case only if $l > 0$, that is, we need an additional boundary condition at 0 if $l = 0$. Since we need $R(r) = \frac{u(r)}{r}$ to be bounded (such that (9.16) is in the domain of H_0), we have to take the boundary condition generated by $u(r) = r$. □

Finally let us turn to some explicit choices for V, where the corresponding differential equation can be explicitly solved. The simplest case is $V = 0$ in this case the solutions of

$$-u''(r) + \frac{l(l+1)}{r^2}u(r) = zu(r) \tag{9.46}$$

are given by the **spherical Bessel** respectively **spherical Neumann** functions

$$u(r) = \alpha\, j_l(\sqrt{z}r) + \beta\, n_l(\sqrt{z}r), \tag{9.47}$$

where

$$j_l(r) = (-r)^l \left(\frac{1}{r}\frac{d}{dr}\right)^l \frac{\sin(r)}{r}. \tag{9.48}$$

In particular,

$$u_a(z,r) = j_l(\sqrt{z}r) \quad \text{and} \quad u_b(z,r) = j_l(\sqrt{z}r) + \mathrm{i}n_l(\sqrt{z}r) \tag{9.49}$$

are the functions which are square integrable and satisfy the boundary condition (if any) near $a = 0$ and $b = \infty$, respectively.

The second case is that of our Coulomb potential

$$V(r) = -\frac{\gamma}{r}, \qquad \gamma > 0, \tag{9.50}$$

where we will try to compute the eigenvalues plus corresponding eigenfunctions. It turns out that they can be expressed in terms of the **Laguerre polynomials**

$$L_j(r) = \mathrm{e}^r \frac{d^j}{dr^j}\mathrm{e}^{-r} r^j \tag{9.51}$$

and the **associated Laguerre polynomials**

$$L_j^k(r) = \frac{d^k}{dr^k} L_j(r). \tag{9.52}$$

Note that L_j^k is a polynomial of degree $j - k$.

Theorem 9.7. *The eigenvalues of $H^{(1)}$ are explicitly given by*

$$E_n = -\left(\frac{\gamma}{2(n+1)}\right)^2, \qquad n \in \mathbb{N}_0. \tag{9.53}$$

An orthonormal basis for the corresponding eigenspace is given by

$$\psi_{nlm}(x) = R_{nl}(r) Y_{lm}(x), \tag{9.54}$$

where

$$R_{nl}(r) = \sqrt{\frac{\gamma^3 (n-l)!}{2n^3((n+l+1)!)^3}} \left(\frac{\gamma r}{n+1}\right)^l \mathrm{e}^{-\frac{\gamma r}{2(n+1)}} L_{n+l+1}^{2l+1}(\frac{\gamma r}{n+1}). \tag{9.55}$$

In particular, the lowest eigenvalue $E_0 = -\frac{\gamma^2}{4}$ is simple and the corresponding eigenfunction $\psi_{000}(x) = \sqrt{\frac{\gamma^3}{4^3\pi}}\mathrm{e}^{-\gamma r/2}$ is positive.

Proof. It is a straightforward calculation to check that R_{nl} are indeed eigenfunctions of A_l corresponding to the eigenvalue $-(\frac{\gamma}{2(n+1)})^2$ and for the norming constants we refer to any book on special functions. The only problem is to show that we have found *all* eigenvalues.

Since all eigenvalues are negative, we need to look at the equation

$$-u''(r) + (\frac{l(l+1)}{r^2} - \frac{\gamma}{r})u(r) = \lambda u(r) \tag{9.56}$$

for $\lambda < 0$. Introducing new variables $x = \sqrt{-\lambda} r$ and $v(x) = x^{l+1} \mathrm{e}^{-x} u(x/\sqrt{-\lambda})$ this equation transforms into

$$x v''(x) + 2(l+1-x)v'(x) + 2n v(x) = 0, \quad n = \frac{\gamma}{2\sqrt{-\lambda}} - (l+1). \tag{9.57}$$

Now let us search for a solution which can be expanded into a convergent power series

$$v(x) = \sum_{j=0}^{\infty} v_j x^j, \quad v_0 = 1. \tag{9.58}$$

The corresponding $u(r)$ is square integrable near 0 and satisfies the boundary condition (if any). Thus we need to find those values of λ for which it is square integrable near $+\infty$.

Substituting the ansatz (9.58) into our differential equation and comparing powers of x gives the following recursion for the coefficients

$$v_{j+1} = \frac{2(j-n)}{(j+1)(j+2(l+1))} v_j \tag{9.59}$$

and thus

$$v_j = \frac{1}{j!} \prod_{k=0}^{j-1} \frac{2(k-n)}{k+2(l+1)}. \tag{9.60}$$

Now there are two cases to distinguish. If $n \in \mathbb{N}_0$, then $v_j = 0$ for $j > n$ and $v(x)$ is a polynomial. In this case $u(r)$ is square integrable and hence an eigenfunction corresponding to the eigenvalue $\lambda_n = -(\frac{\gamma}{2(n+l+1)})^2$. Otherwise we have $v_j \ge \frac{(2-\varepsilon)^j}{j!}$ for j sufficiently large. Hence by adding a polynomial to $v(x)$ we can get a function $\tilde{v}(x)$ such that $\tilde{v}_j \ge \frac{(2-\varepsilon)^j}{j!}$ for all j. But then $\tilde{v}(x) \ge \exp((2-\varepsilon)x)$ and thus the corresponding $u(r)$ is not square integrable near $-\infty$. □

9.4. Atomic Schrödinger operators

In this section we want to have a look at the Hamiltonian corresponding to more than one interacting particle. It is given by

$$H = -\sum_{j=1}^{N} \Delta_j + \sum_{j<k}^{N} V_{j,k}(x_j - x_k). \tag{9.61}$$

We first consider the case of two particles, which will give us a feeling for how the many particle case differs from the one particle case and how the difficulties can be overcome.

We denote the coordinates corresponding to the first particle by $x_1 = (x_{1,1}, x_{1,2}, x_{1,3})$ and those corresponding to the second particle by $x_2 =$

$(x_{2,1}, x_{2,2}, x_{2,3})$. If we assume that the interaction is again of the Coulomb type, the Hamiltonian is given by

$$H = -\Delta_1 - \Delta_2 - \frac{\gamma}{|x_1 - x_2|}, \quad \mathfrak{D}(H) = H^2(\mathbb{R}^6). \tag{9.62}$$

Since Theorem 8.17 does not allow singularities for $n \ge 3$, it does *not* tell us whether H is self-adjoint or not. Let

$$(y_1, y_2) = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{I} & \mathbb{I} \\ -\mathbb{I} & \mathbb{I} \end{pmatrix} (x_1, x_2), \tag{9.63}$$

then H reads in this new coordinates

$$H = (-\Delta_1) + (-\Delta_2 - \frac{\gamma/\sqrt{2}}{|y_2|}). \tag{9.64}$$

In particular, it is the sum of a free particle plus a particle in an external Coulomb field. From a physics point of view, the first part corresponds to the center of mass motion and the second part to the relative motion.

Using that $\gamma/(\sqrt{2}|y_2|)$ has $(-\Delta_2)$-bound 0 in $L^2(\mathbb{R}^3)$ it is not hard to see that the same is true for the $(-\Delta_1 - \Delta_2)$-bound in $L^2(\mathbb{R}^6)$ (details will follow in the next section). In particular, H is self-adjoint and semi-bounded for any $\gamma \in \mathbb{R}$. Moreover, you might suspect that $\gamma/(\sqrt{2}|y_2|)$ is relatively compact with respect to $-\Delta_1 - \Delta_2$ in $L^2(\mathbb{R}^6)$ since it is with respect to $-\Delta_2$ in $L^2(\mathbb{R}^6)$. However, this is *not* true! This is due to the fact that $\gamma/(\sqrt{2}|y_2|)$ does not vanish as $|y| \to \infty$.

Let us look at this problem from the physical view point. If $\lambda \in \sigma_{ess}(H)$, this means that the movement of the whole system is somehow unbounded. There are two possibilities for this.

Firstly, both particles are far away from each other (such that we can neglect the interaction) and the energy corresponds to the sum of the kinetic energies of both particles. Since both can be arbitrarily small (but positive), we expect $[0, \infty) \subseteq \sigma_{ess}(H)$.

Secondly, both particles remain close to each other and move together. In the last coordinates this corresponds to a bound state of the second operator. Hence we expect $[\lambda_0, \infty) \subseteq \sigma_{ess}(H)$, where $\lambda_0 = -\gamma^2/8$ is the smallest eigenvalue of the second operator if the forces are attracting ($\gamma \ge 0$) and $\lambda_0 = 0$ if they are repelling ($\gamma \le 0$).

It is not hard to translate this intuitive ideas into a rigorous proof. Let $\psi_1(y_1)$ be a Weyl sequence corresponding to $\lambda \in [0, \infty)$ for $-\Delta_1$ and $\psi_2(y_2)$ be a Weyl sequence corresponding to λ_0 for $-\Delta_2 - \gamma/(\sqrt{2}|y_2|)$. Then, $\psi_1(y_1)\psi_2(y_2)$ is a Weyl sequence corresponding to $\lambda + \lambda_0$ for H and thus $[\lambda_0, \infty) \subseteq \sigma_{ess}(H)$. Conversely, we have $-\Delta_1 \ge 0$ respectively $-\Delta_2 -$

$\gamma/(\sqrt{2}|y_2|) \geq \lambda_0$ and hence $H \geq \lambda_0$. Thus we obtain

$$\sigma(H) = \sigma_{ess}(H) = [\lambda_0, \infty), \qquad \lambda_0 = \begin{cases} -\gamma^2/8, & \gamma \geq 0 \\ 0, & \gamma \leq 0 \end{cases}. \tag{9.65}$$

Clearly, the physically relevant information is the spectrum of the operator $-\Delta_2 - \gamma/(\sqrt{2}|y_2|)$ which is hidden by the spectrum of $-\Delta_1$. Hence, in order to reveal the physics, one first has to *remove* the center of mass motion.

To avoid clumsy notation, we will restrict ourselves to the case of one atom with N electrons whose nucleus is fixed at the origin. In particular, this implies that we do not have to deal with the center of mass motion encountered in our example above. The Hamiltonian is given by

$$\begin{aligned} H^{(N)} &= -\sum_{j=1}^{N} \Delta_j - \sum_{j=1}^{N} V_{ne}(x_j) + \sum_{j=1}^{N} \sum_{j<k}^{N} V_{ee}(x_j - x_k), \\ \mathfrak{D}(H^{(N)}) &= H^2(\mathbb{R}^{3N}), \end{aligned} \tag{9.66}$$

where V_{ne} describes the interaction of one electron with the nucleus and V_{ee} describes the interaction of two electrons. Explicitly we have

$$V_j(x) = \frac{\gamma_j}{|x|}, \qquad \gamma_j > 0, \ j = ne, ee. \tag{9.67}$$

We first need to establish self-adjointness of $H^{(N)}$. This will follow from Kato's theorem.

Theorem 9.8 (Kato)**.** *Let $V_k \in L^\infty_\infty(\mathbb{R}^d) + L^2(\mathbb{R}^d)$, $d \leq 3$, be real-valued and let $V_k(y^{(k)})$ be the multiplication operator in $L^2(\mathbb{R}^n)$, $n = Nd$, obtained by letting $y^{(k)}$ be the first d coordinates of a unitary transform of $\mathbb{R}^n$. Then V_k is H_0 bounded with H_0-bound 0. In particular,*

$$H = H_0 + \sum_k V_k(y^{(k)}), \qquad \mathfrak{D}(H) = H^2(\mathbb{R}^n), \tag{9.68}$$

is self-adjoint.

Proof. It suffices to consider one k. After a unitary transform of $\mathbb{R}^n$ we can assume $y^{(1)} = (x_1, \dots, x_d)$ since such transformation leave both the scalar product of $L^2(\mathbb{R}^n)$ and H_0 invariant. Now let $\psi \in \mathcal{S}(\mathbb{R}^n)$, then

$$\|V_k \psi\|^2 \leq a^2 \int_{\mathbb{R}^n} |\Delta_1 \psi(x)|^2 d^n x + b^2 \int_{\mathbb{R}^n} |\psi(x)|^2 d^n x, \tag{9.69}$$

where $\Delta_1 = \sum_{j=1}^d \partial^2/\partial^2 x_j$, by our previous lemma. Hence we obtain

$$\begin{aligned}\|V_k\psi\|^2 &\leq a^2 \int_{\mathbb{R}^n} |\sum_{j=1}^{d} p_j^2 \hat{\psi}(p)|^2 d^n p + b^2\|\psi\|^2 \\ &\leq a^2 \int_{\mathbb{R}^n} |\sum_{j=1}^{n} p_j^2 \hat{\psi}(p)|^2 d^n p + b^2\|\psi\|^2 \\ &= a^2\|H_0\psi\|^2 + b^2\|\psi\|^2, \end{aligned} \tag{9.70}$$

which implies that V_k is relatively bounded with bound 0. □

The considerations of the beginning of this section show that it is not so easy to determine the essential spectrum of $H^{(N)}$ since the potential does not decay in all directions as $|x| \to \infty$. However, there is still something we can do. Denote the infimum of the spectrum of $H^{(N)}$ by λ^N. Then, let us split the system into $H^{(N-1)}$ plus a single electron. If the single electron is far away from the remaining system such that there is little interaction, the energy should be the sum of the kinetic energy of the single electron and the energy of the remaining system. Hence arguing as in the two electron example of the previous section we expect

Theorem 9.9 (HVZ)**.** *Let $H^{(N)}$ be the self-adjoint operator given in (9.66). Then $H^{(N)}$ is bounded from below and*

$$\sigma_{ess}(H^{(N)}) = [\lambda^{N-1}, \infty), \tag{9.71}$$

where $\lambda^N = \min \sigma(H^{(N)}) < 0$.

In particular, the ionization energy (i.e., the energy needed to remove one electron from the atom in its ground state) of an atom with N electrons is given by $\lambda^N - \lambda^{N-1}$.

Our goal for the rest of this section is to prove this result which is due to Zhislin, van Winter and Hunziker and known as HVZ theorem. In fact there is a version which holds for general N-body systems. The proof is similar but involves some additional notation.

The idea of proof is the following. To prove $[\lambda^{N-1}, \infty) \subseteq \sigma_{ess}(H^{(N)})$ we choose Weyl sequences for $H^{(N-1)}$ and $-\Delta_N$ and proceed according to our intuitive picture from above. To prove $\sigma_{ess}(H^{(N)}) \subseteq [\lambda^{N-1}, \infty)$ we will *localize* $H^{(N)}$ on sets where either one electron is far away from the others or all electrons are far away from the nucleus. Since the error turns out relatively compact, it remains to consider the infimum of the spectra of these operators. For all cases where one electron is far away it is λ^{N-1} and for the case where all electrons are far away from the nucleus it is 0 (since the electrons repel each other).

We begin with the first inclusion. Let $\psi^{N-1}(x_1,\dots,x_{N-1}) \in H^2(\mathbb{R}^{3(N-1)})$ such that $\|\psi^{N-1}\| = 1$, $\|(H^{(N-1)} - \lambda^{N-1})\psi^{N-1}\| \le \varepsilon$ and $\psi^1 \in H^2(\mathbb{R}^3)$ such that $\|\psi^1\| = 1$, $\|(-\Delta_N - \lambda)\psi^{N-1}\| \le \varepsilon$ for some $\lambda \ge 0$. Now consider $\psi_r(x_1,\dots,x_N) = \psi^{N-1}(x_1,\dots,x_{N-1})\psi_r^1(x_N)$, $\psi_r^1(x_N) = \psi^1(x_N - r)$, then

$$\begin{aligned}\|(H^{(N)} - \lambda - \lambda^{N-1})\psi_r\| \le{}& \|(H^{(N-1)} - \lambda^{N-1})\psi^{N-1}\|\|\psi_r^1\| \\ &+ \|\psi^{N-1}\|\|(-\Delta_N - \lambda)\psi_r^1\| \\ &+ \|(V_N - \sum_{j=1}^{N-1} V_{N,j})\psi_r\|,\end{aligned} \tag{9.72}$$

where $V_N = V_{ne}(x_N)$ and $V_{N,j} = V_{ee}(x_N - x_j)$. Since $(V_N - \sum_{j=1}^{N-1} V_{N,j})\psi^{N-1} \in L^2(\mathbb{R}^{3N})$ and $|\psi_r^1| \to 0$ pointwise as $|r| \to \infty$ (by Lemma 8.16), the third term can be made smaller than ε by choosing $|r|$ large (dominated convergence). In summary,

$$\|(H^{(N)} - \lambda - \lambda^{N-1})\psi_r\| \le 3\varepsilon \tag{9.73}$$

proving $[\lambda^{N-1}, \infty) \subseteq \sigma_{ess}(H^{(N)})$.

The second inclusion is more involved. We begin with a **localization formula**, which can be verified by a straightforward computation

Lemma 9.10 (IMS localization formula)**.** *Suppose $\phi_j \in C^\infty(\mathbb{R}^n)$, $0 \le j \le N$, is such that*

$$\sum_{j=0}^{N} \phi_j(x)^2 = 1, \quad x \in \mathbb{R}^n, \tag{9.74}$$

then

$$\Delta\psi = \sum_{j=0}^{N} \phi_j \Delta \phi_j \psi - |\partial\phi_j|^2 \psi, \qquad \psi \in H^2(\mathbb{R}^n). \tag{9.75}$$

Abbreviate $B = \{x \in \mathbb{R}^{3N} | |x| \ge 1\}$. Now we will choose ϕ_j, $1 \le j \le N$, in such a way that $x \in \mathrm{supp}(\phi_j) \cap B$ implies that the j-th particle is far away from all the others and from the nucleus. Similarly, we will choose ϕ_0 in such a way that $x \in \mathrm{supp}(\phi_0) \cap B$ implies that all particle are far away from the nucleus.

Lemma 9.11. *There exists functions $\phi_j \in C^\infty(\mathbb{R}^n, [0,1])$, $0 \le j \le N$, is such that (9.74) holds,*

$$\begin{aligned}\mathrm{supp}(\phi_j) \cap B &\subseteq \{x \in B | \, |x_j - x_\ell| \ge C|x| \textit{ for all } \ell \ne j, \textit{ and } |x_j| \ge C|x|\}, \\ \mathrm{supp}(\phi_0) \cap B &\subseteq \{x \in B | \, |x_\ell| \ge C|x| \textit{ for all } \ell\}\end{aligned} \tag{9.76}$$

for some $C \in [0,1]$, and $|\partial\phi_j(x)| \to 0$ as $|x| \to \infty$.

Proof. Consider the sets

$$\begin{aligned} U_j^n &= \{x \in S^{3N-1} \,|\, |x_j - x_\ell| > n^{-1} \text{ for all } \ell \neq j, \text{ and } |x_j| > n^{-1}\}, \\ U_0^N &= \{x \in S^{3N-1} \,|\, |x_\ell| > n^{-1} \text{ for all } \ell\}. \end{aligned} \tag{9.77}$$

We claim that

$$\bigcup_{n=1}^{\infty} \bigcup_{j=0}^{N} U_j^n = S^{3N-1}. \tag{9.78}$$

Indeed, suppose there is an $x \in S^{3N-1}$ which is not an element of this union. Then $x \notin U_0^n$ for all n implies $0 = |x_j|$ for some j, say $j = 1$. Next, since $x \notin U_1^n$ for all n implies $0 = |x_j - x_1| = |x_j|$ for some $j > 1$, say $j = 2$. Proceeding like this we end up with $x = 0$, a contradiction. By compactness of S^{3N-1} we even have

$$\bigcup_{j=0}^{N} U_j^n = S^{3N-1} \tag{9.79}$$

for n sufficiently large. It is well-known that there is a partition of unity $\tilde{\phi}_j(x)$ subordinate to this cover. Extend $\tilde{\phi}_j(x)$ to a smooth function from $\mathbb{R}^{3N}\backslash\{0\}$ to $[0,1]$ by

$$\tilde{\phi}_j(\lambda x) = \tilde{\phi}_j(x), \qquad x \in S^{3N-1},\ \lambda > 0, \tag{9.80}$$

and pick a function $\tilde{\phi} \in C^\infty(\mathbb{R}^{3N}, [0,1])$ with support inside the unit ball which is 1 in a neighborhood of the origin. Then

$$\phi_j = \frac{\tilde{\phi} + (1-\tilde{\phi})\tilde{\phi}_j}{\sqrt{\sum_{\ell=0}^N \tilde{\phi} + (1-\tilde{\phi})\tilde{\phi}_\ell}} \tag{9.81}$$

are the desired functions. The gradient tends to zero since $\phi_j(\lambda x) = \phi_j(x)$ for $\lambda \geq 1$ and $|x| \geq 1$ which implies $(\partial\phi_j)(\lambda x) = \lambda^{-1}(\partial\phi_j)(x)$. □

By our localization formula we have

$$H^{(N)} = \sum_{j=0}^{N} \phi_j H^{(N,j)} \phi_j + K, \qquad K = \sum_{j=0}^{N} \phi_j^2 V^{(N,j)} + |\partial\phi_j|^2, \tag{9.82}$$

where

$$\begin{aligned} H^{(N,j)} &= -\sum_{\ell=1}^{N} \Delta_\ell - \sum_{\ell\neq j}^{N} V_\ell + \sum_{k<\ell,\, k,\ell\neq j}^{N} V_{k,\ell}, \quad H^{(N,0)} = -\sum_{\ell=1}^{N} \Delta_\ell + \sum_{k<\ell}^{N} V_{k,\ell} \\ V^{(N,j)} &= V_j + \sum_{\ell\neq j}^{N} V_{j,\ell}, \quad V^{(N,0)} = \sum_{\ell=1}^{N} V_\ell \end{aligned} \tag{9.83}$$

To show that our choice of the functions ϕ_j implies that K is relatively compact with respect to H we need the following

Lemma 9.12. *Let V be H_0 bounded with H_0-bound 0 and suppose that $\|\chi_{\{x||x|\geq R\}} V R_{H_0}(z)\| \to 0$ as $R \to \infty$. Then V is relatively compact with respect to H_0.*

Proof. Let ψ_n converge to 0 weakly. Note that $\|\psi_n\| \leq M$ for some $M > 0$. It suffices to show that $\|V R_{H_0}(z)\psi_n\|$ converges to 0. Choose $\phi \in C_0^\infty(\mathbb{R}^n, [0,1])$ such that it is one for $|x| \leq R$. Then

$$\begin{aligned} \|V R_{H_0}(z)\psi_n\| &\leq \|(1-\phi)V R_{H_0}(z)\psi_n\| + \|V\phi R_{H_0}(z)\psi_n\| \\ &\leq \|(1-\phi)V R_{H_0}(z)\|_\infty \|\psi_n\| + \\ &\quad a\|H_0 \phi R_{H_0}(z)\psi_n\| + b\|\phi R_{H_0}(z)\psi_n\|. \end{aligned} \tag{9.84}$$

By assumption, the first term can be made smaller than ε by choosing R large. Next, the same is true for the second term choosing a small. Finally, the last term can also be made smaller than ε by choosing n large since ϕ is H_0 compact. □

The terms $|\partial\phi_j|^2$ are bounded and vanish at ∞, hence they are H_0 compact by Lemma 6.9. The terms $\phi_j V^{(N,j)}$ are relatively compact by the lemma and hence K is relatively compact with respect to H_0. By Lemma 8.15, K is also relatively compact with respect to $H^{(N)}$ since $V^{(N)}$ is relatively bounded with respect to H_0.

In particular $H^{(N)} - K$ is self-adjoint on $H^2(\mathbb{R}^{3N})$ and $\sigma_{ess}(H^{(N)}) = \sigma_{ess}(H^{(N)} - K)$. Since the operators $H^{(N,j)}$, $1 \leq j \leq N$, are all of the form $H^{(N-1)}$ plus one particle which does not interact with the others and the nucleus, we have $H^{(N,j)} - \lambda^{N-1} \geq 0$, $1 \leq j \leq N$. Moreover, we have $H^{(0)} \geq 0$ since $V_{j,k} \geq 0$ and hence

$$\langle \psi, (H^{(N)} - K - \lambda^{N-1})\psi\rangle = \sum_{j=0}^{N} \langle \phi_j\psi, (H^{(N,j)} - \lambda^{N-1})\phi_j\psi\rangle \geq 0. \tag{9.85}$$

Thus we obtain the remaining inclusion

$$\sigma_{ess}(H^{(N)}) = \sigma_{ess}(H^{(N)} - K) \subseteq \sigma(H^{(N)} - K) \subseteq [\lambda^{N-1}, \infty) \tag{9.86}$$

which finishes the proof of the HVZ theorem.

Note that the same proof works if we add additional nuclei at fixed locations. That is, we can also treat molecules if we assume that the nuclei are fixed in space.

Finally, let us consider the example of Helium like atoms ($N = 2$). By the HVZ theorem and the considerations of the previous section we have

$$\sigma_{ess}(H^{(2)}) = [-\frac{\gamma_{ne}^2}{4}, \infty). \tag{9.87}$$

Moreover, if $\gamma_{ee} = 0$ (no electron interaction), we can take products of one particle eigenfunctions to show that

$$-\gamma_{ne}^2 \left(\frac{1}{4n^2} + \frac{1}{4m^2} \right) \in \sigma_p(H^{(2)}(\gamma_{ee} = 0)), \qquad n, m \in \mathbb{N}. \tag{9.88}$$

In particular, there are eigenvalues embedded in the essential spectrum in this case. Moreover, since the electron interaction term is positive, we see

$$H^{(2)} \geq -\frac{\gamma_{ne}^2}{2}. \tag{9.89}$$

Note that there can be no positive eigenvalues by the virial theorem. This even holds for arbitrary N,

$$\sigma_p(H^{(N)}) \subset (-\infty, 0). \tag{9.90}$$

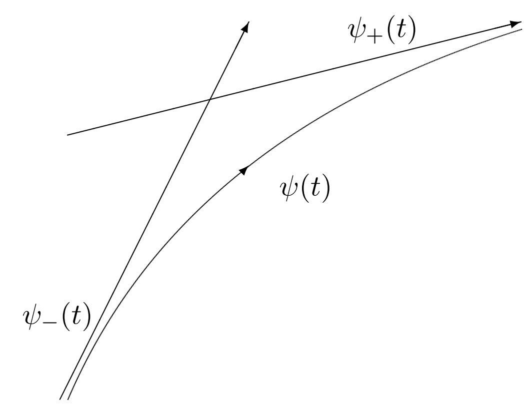
Chapter 10

Scattering theory

10.1. Abstract theory

In physical measurements one often has the following situation. A particle is shot into a region where it interacts with some forces and then leaves the region again. Outside this region the forces are negligible and hence the time evolution should be asymptotically free. Hence one expects asymptotic states $\psi_\pm(t) = \exp(-\mathrm{i}tH_0)\psi_\pm(0)$ to exist such that

$$\|\psi(t) - \psi_\pm(t)\| \to 0 \quad \text{as} \quad t \to \pm\infty. \tag{10.1}$$



Rewriting this condition we see

$$0 = \lim_{t\to\pm\infty} \|\mathrm{e}^{-\mathrm{i}tH}\psi(0) - \mathrm{e}^{-\mathrm{i}tH_0}\psi_\pm(0)\| = \lim_{t\to\pm\infty} \|\psi(0) - \mathrm{e}^{\mathrm{i}tH}\mathrm{e}^{-\mathrm{i}tH_0}\psi_\pm(0)\| \tag{10.2}$$

and motivated by this we define the **wave operators** by

$$\begin{array}{rcl} \mathfrak{D}(\Omega_\pm) &=& \{\psi \in \mathfrak{H} | \exists \lim_{t\to\pm\infty} \mathrm{e}^{\mathrm{i}tH}\mathrm{e}^{-\mathrm{i}tH_0}\psi\} \\ \Omega_\pm\psi &=& \lim_{t\to\pm\infty} \mathrm{e}^{\mathrm{i}tH}\mathrm{e}^{-\mathrm{i}tH_0}\psi \end{array}. \tag{10.3}$$

The set $\mathfrak{D}(\Omega_\pm)$ is the set of all incoming/outgoing asymptotic states $\psi_\pm$ and $\mathrm{Ran}(\Omega_\pm)$ is the set of all states which have an incoming/outgoing asymptotic state. If a state ψ has both, that is, $\psi \in \mathrm{Ran}(\Omega_+) \cap \mathrm{Ran}(\Omega_-)$, it is called a **scattering state**.

By construction we have

$$\|\Omega_\pm\psi\| = \lim_{t\to\pm\infty} \|\mathrm{e}^{\mathrm{i}tH}\mathrm{e}^{-\mathrm{i}tH_0}\psi\| = \lim_{t\to\pm\infty} \|\psi\| = \|\psi\| \tag{10.4}$$

and it is not hard to see that $\mathfrak{D}(\Omega_\pm)$ is closed. Moreover, interchanging the roles of H_0 and H amounts to replacing $\Omega_\pm$ by $\Omega_\pm^{-1}$ and hence $\mathrm{Ran}(\Omega_\pm)$ is also closed. In summary,

Lemma 10.1. *The sets* $\mathfrak{D}(\Omega_\pm)$ *and* $\mathrm{Ran}(\Omega_\pm)$ *are closed and* $\Omega_\pm : \mathfrak{D}(\Omega_\pm) \to \mathrm{Ran}(\Omega_\pm)$ *is unitary.*

Next, observe that

$$\lim_{t\to\pm\infty} \mathrm{e}^{\mathrm{i}tH}\mathrm{e}^{-\mathrm{i}tH_0}(\mathrm{e}^{-\mathrm{i}sH_0}\psi) = \lim_{t\to\pm\infty} \mathrm{e}^{-\mathrm{i}sH}(\mathrm{e}^{\mathrm{i}(t+s)H}\mathrm{e}^{-\mathrm{i}(t+s)H_0}\psi) \tag{10.5}$$

and hence

$$\Omega_\pm \mathrm{e}^{-\mathrm{i}tH_0}\psi = \mathrm{e}^{-\mathrm{i}tH}\Omega_\pm\psi, \quad \psi \in \mathfrak{D}(\Omega_\pm). \tag{10.6}$$

In addition, $\mathfrak{D}(\Omega_\pm)$ is invariant under $\exp(-\mathrm{i}tH_0)$ and $\mathrm{Ran}(\Omega_\pm)$ is invariant under $\exp(-\mathrm{i}tH)$. Moreover, if $\psi \in \mathfrak{D}(\Omega_\pm)^\perp$ then

$$\langle\varphi, \exp(-\mathrm{i}tH_0)\psi\rangle = \langle\exp(\mathrm{i}tH_0)\varphi, \psi\rangle = 0, \quad \varphi \in \mathfrak{D}(\Omega_\pm). \tag{10.7}$$

Hence $\mathfrak{D}(\Omega_\pm)^\perp$ is invariant under $\exp(-\mathrm{i}tH_0)$ and $\mathrm{Ran}(\Omega_\pm)^\perp$ is invariant under $\exp(-\mathrm{i}tH)$. Consequently, $\mathfrak{D}(\Omega_\pm)$ reduces $\exp(-\mathrm{i}tH_0)$ and $\mathrm{Ran}(\Omega_\pm)$ reduces $\exp(-\mathrm{i}tH)$. Moreover, differentiating (10.6) with respect to t we obtain from Theorem 5.1 the **intertwining property** of the wave operators.

Theorem 10.2. *The subspaces* $\mathfrak{D}(\Omega_\pm)$ *respectively* $\mathrm{Ran}(\Omega_\pm)$ *reduce* H_0 *respectively* H *and the operators restricted to these subspaces are unitarily equivalent*

$$\Omega_\pm H_0\psi = H\Omega_\pm\psi, \qquad \psi \in \mathfrak{D}(\Omega_\pm) \cap \mathfrak{D}(H_0). \tag{10.8}$$

It is interesting to know the correspondence between incoming and outgoing states. Hence we define the **scattering operator**

$$S = \Omega_+^{-1}\Omega_-, \quad \mathfrak{D}(S) = \{\psi \in \mathfrak{D}(\Omega_-) | \Omega_-\psi \in \mathrm{Ran}(\Omega_+)\}. \tag{10.9}$$

Note that we have $\mathfrak{D}(S) = \mathfrak{D}(\Omega_-)$ if and only if $\mathrm{Ran}(\Omega_+) \subseteq \mathrm{Ran}(\Omega_-)$ and $\mathrm{Ran}(S) = \mathfrak{D}(\Omega_+)$ if and only if $\mathrm{Ran}(\Omega_-) \subseteq \mathrm{Ran}(\Omega_+)$. Moreover, S is unitary from $\mathfrak{D}(S)$ onto $\mathrm{Ran}(S)$ and we have

$$H_0 S\psi = SH_0\psi, \quad \mathfrak{D}(H_0) \cap \mathfrak{D}(S). \tag{10.10}$$

However, note that this whole theory is meaningless until we can show that $\mathfrak{D}(\Omega_\pm)$ are nontrivial. We first show a criterion due to Cook.

Lemma 10.3 (Cook). *Suppose $\mathfrak{D}(H) \subseteq \mathfrak{D}(H_0)$. If*

$$\int_0^\infty \|(H - H_0)\exp(\mp \mathrm{i}tH_0)\psi\| dt < \infty, \quad \psi \in \mathfrak{D}(H_0), \tag{10.11}$$

then $\psi \in \mathfrak{D}(\Omega_\pm)$, respectively. Moreover, we even have

$$\|(\Omega_\pm - \mathbb{I})\psi\| \le \int_0^\infty \|(H - H_0)\exp(\mp \mathrm{i}tH_0)\psi\| dt \tag{10.12}$$

in this case.

Proof. The result follows from

$$\mathrm{e}^{\mathrm{i}tH}\mathrm{e}^{-\mathrm{i}tH_0}\psi = \psi + \mathrm{i}\int_0^t \exp(\mathrm{i}sH)(H - H_0)\exp(-\mathrm{i}sH_0)\psi ds \tag{10.13}$$

which holds for $\psi \in \mathfrak{D}(H_0)$. □

As a simple consequence we obtain the following result for Schrödinger operators in $\mathbb{R}^3$

Theorem 10.4. *Suppose H_0 is the free Schrödinger operator and $H = H_0 + V$ with $V \in L^2(\mathbb{R}^3)$, then the wave operators exist and $\mathfrak{D}(\Omega_\pm) = \mathfrak{H}$.*

Proof. Since we want to use Cook's lemma, we need to estimate

$$\|V\psi(s)\|^2 = \int_{\mathbb{R}^3} |V(x)\psi(s,x)|^2 dx, \quad \psi(s) = \exp(\mathrm{i}sH_0)\psi, \tag{10.14}$$

for given $\psi \in \mathfrak{D}(H_0)$. Invoking (6.37) we get

$$\|V\psi(s)\| \le \|\psi(s)\|_\infty \|V\| \le \frac{1}{(4\pi s)^{3/2}} \|\psi\|_1 \|V\|, \quad s > 0, \tag{10.15}$$

at least for $\psi \in L^1(\mathbb{R}^3)$. Moreover, this implies

$$\int_1^\infty \|V\psi(s)\| ds \le \frac{1}{4\pi^{3/2}} \|\psi\|_1 \|V\| \tag{10.16}$$

and thus any such ψ is in $\mathfrak{D}(\Omega_+)$. Since such functions are dense, we obtain $\mathfrak{D}(\Omega_+) = \mathfrak{H}$. Similarly for Ω_-. □

By the intertwining property ψ is an eigenfunction of H_0 if and only if it is an eigenfunction of H. Hence for $\psi \in \mathfrak{H}_{pp}(H_0)$ it is easy to check whether it is in $\mathfrak{D}(\Omega_\pm)$ or not and only the continuous subspace is of interest. We will say that the **wave operators exist** if all elements of $\mathfrak{H}_{ac}(H_0)$ are asymptotic states, that is,

$$\mathfrak{H}_{ac}(H_0) \subseteq \mathfrak{D}(\Omega_\pm) \tag{10.17}$$

and that they are **complete** if, in addition, all elements of $\mathfrak{H}_{ac}(H)$ are scattering states, that is,

$$\mathfrak{H}_{ac}(H) \subseteq \mathrm{Ran}(\Omega_\pm). \tag{10.18}$$

If we even have

$$\mathfrak{H}_c(H) \subseteq \mathrm{Ran}(\Omega_\pm), \tag{10.19}$$

they are called **asymptotically complete**. We will be mainly interested in the case where H_0 is the free Schrödinger operator and hence $\mathfrak{H}_{ac}(H_0) = \mathfrak{H}$. In this later case the wave operators exist if $\mathfrak{D}(\Omega_\pm) = \mathfrak{H}$, they are complete if $\mathfrak{H}_{ac}(H) = \mathrm{Ran}(\Omega_\pm)$, and asymptotically complete if $\mathfrak{H}_c(H) = \mathrm{Ran}(\Omega_\pm)$. In particular asymptotic completeness implies $\mathfrak{H}_{sc}(H) = \emptyset$ since H restricted to $\mathrm{Ran}(\Omega_\pm)$ is unitarily equivalent to H_0.

10.2. Incoming and outgoing states

In the remaining sections we want to apply this theory to Schrödinger operators. Our first goal is to give a precise meaning to some terms in the intuitive picture of scattering theory introduced in the previous section.

This physical picture suggests that we should be able to decompose $\psi \in \mathfrak{H}$ into an incoming and an outgoing part. But how should incoming respectively outgoing be defined for $\psi \in \mathfrak{H}$? Well incoming (outgoing) means that the expectation of x^2 should decrease (increase). Set $x(t)^2 = \exp(\mathrm{i}H_0 t) x^2 \exp(-\mathrm{i}H_0 t)$, then, abbreviating $\psi(t) = \mathrm{e}^{-\mathrm{i}tH_0}\psi$,

$$\frac{d}{dt}\mathbb{E}_\psi(x(t)^2) = \langle \psi(t), \mathrm{i}[H_0, x^2]\psi(t)\rangle = 4\langle \psi(t), D\psi(t)\rangle, \quad \psi \in \mathcal{S}(\mathbb{R}^n), \tag{10.20}$$

where D is the dilation operator introduced in (9.5). Hence it is natural to consider $\psi \in \mathrm{Ran}(P_\pm)$,

$$P_\pm = P_D((0, \pm\infty)), \tag{10.21}$$

as outgoing respectively incoming states. If we project a state in $\mathrm{Ran}(P_\pm)$ to energies in the interval (a^2, b^2), we expect that it cannot be found in a ball of radius proportional to $a|t|$ as $t \to \pm\infty$ (a is the minimal velocity of the particle, since we have assumed the mass to be two). In fact, we will show below that the tail decays faster then any inverse power of $|t|$.

We first collect some properties of D which will be needed later on. Note

$$\mathcal{F} D = -D \mathcal{F} \tag{10.22}$$

and hence $\mathcal{F} f(D) = f(-D)\mathcal{F}$. To say more we will look for a transformation which maps D to a multiplication operator.

Since the dilation group acts on $|x|$ only, it seems reasonable to switch to polar coordinates $x = r\omega$, $(t, \omega) \in \mathbb{R}^+ \times S^{n-1}$. Since $U(s)$ essentially transforms r into $r\exp(s)$ we will replace r by $\rho = \ln(r)$. In these coordinates we have

$$U(s)\psi(\mathrm{e}^\rho \omega) = \mathrm{e}^{-ns/2}\psi(\mathrm{e}^{(\rho - s)}\omega) \tag{10.23}$$

and hence $U(s)$ corresponds to a shift of ρ (the constant in front is absorbed by the volume element). Thus D corresponds to differentiation with respect to this coordinate and all we have to do to make it a multiplication operator is to take the Fourier transform with respect to ρ.

This leads us to the **Mellin transform**

$$\begin{array}{rccl} \mathcal{M}: & L^2(\mathbb{R}^n) & \to & L^2(\mathbb{R}\times S^{n-1}) \\ & \psi(r\omega) & \to & (\mathcal{M}\psi)(\lambda,\omega) = \frac{1}{\sqrt{2\pi}}\int_0^\infty r^{-\mathrm{i}\lambda}\psi(r\omega) r^{\frac{n}{2}-1}dr \end{array}. \tag{10.24}$$

By construction, $\mathcal{M}$ is unitary, that is,

$$\int_{\mathbb{R}}\int_{S^{n-1}} |(\mathcal{M}\psi)(\lambda,\omega)|^2 d\lambda d^{n-1}\omega = \int_{\mathbb{R}^+}\int_{S^{n-1}} |\psi(r\omega)|^2 r^{n-1} dr d^{n-1}\omega, \tag{10.25}$$

where $d^{n-1}\omega$ is the normalized surface measure on S^{n-1}. Moreover,

$$\mathcal{M}^{-1}U(s)\mathcal{M} = \mathrm{e}^{-\mathrm{i}s\lambda} \tag{10.26}$$

and hence

$$\mathcal{M}^{-1}D\mathcal{M} = \lambda. \tag{10.27}$$

From this it is straightforward to show that

$$\sigma(D) = \sigma_{ac}(D) = \mathbb{R}, \qquad \sigma_{sc}(D) = \sigma_{pp}(D) = \emptyset \tag{10.28}$$

and that $\mathcal{S}(\mathbb{R}^n)$ is a core for D. In particular we have $P_+ + P_- = \mathbb{I}$.

Using the Mellin transform we can now prove Perry's estimate [8].

Lemma 10.5. *Suppose $f \in C_0^\infty(\mathbb{R})$ with $\mathrm{supp}(f) \subset (a^2, b^2)$ for some $a, b > 0$. For any $R \in \mathbb{R}$, $N \in \mathbb{N}$ there is a constant C such that*

$$\|\chi_{\{x|\,|x|<2a|t|\}} \mathrm{e}^{-\mathrm{i}tH_0} f(H_0) P_D((\pm R, \pm\infty))\| \le \frac{C}{(1+|t|)^N}, \quad \pm t \ge 0, \tag{10.29}$$

respectively.

Proof. We prove only the + case, the remaining one being similar. Consider $\psi \in \mathcal{S}(\mathbb{R}^n)$. Introducing

$$\begin{aligned} \psi(t,x) &= \mathrm{e}^{-\mathrm{i}tH_0} f(H_0) P_D((R,\infty))\psi(x) = \langle K_{t,x}, \mathcal{F} P_D((R,\infty))\psi\rangle \\ &= \langle K_{t,x}, P_D((-\infty,-R))\hat{\psi}\rangle, \end{aligned} \tag{10.30}$$

where

$$K_{t,x}(p) = \frac{1}{(2\pi)^{n/2}} \mathrm{e}^{\mathrm{i}(\frac{p^2}{t} + px)} f(p^2)^*, \tag{10.31}$$

we see that it suffices to show

$$\|P_D((-\infty,-R))K_{t,x}\|^2 \le \frac{const}{(1+|t|)^{2N}}, \qquad \text{for } |x| < 2a|t|,\ t > 0. \tag{10.32}$$

Now we invoke the Mellin transform to estimate this norm

$$\|P_D((-\infty,-R))K_{t,x}\|^2 = \int_{-\infty}^{R}\int_{S^{n-1}} |(\mathcal{M}K_{t,x})(\lambda,\omega)|^2 d\lambda d^{n-1}\omega. \quad (10.33)$$

Since

$$(\mathcal{M}K_{t,x})(\lambda,\omega) = \frac{1}{(2\pi)^{(n+1)/2}}\int_0^\infty \tilde{f}(r)\mathrm{e}^{\mathrm{i}\alpha(r)}dr \quad (10.34)$$

with $\tilde{f}(r) = f(r^2)^* r^{n/2-1} \in C_0^\infty((a^2,b^2))$, $\alpha(r) = tr^2 + r\omega x - \lambda\ln(r)$. Estimating the derivative of α we see

$$\alpha'(r) = 2tr + \omega x - \lambda/r > 0, \quad r \in (a,b), \quad (10.35)$$

for $\lambda \le -R$ and $t > R(2\varepsilon a)^{-1}$, where ε is the distance of a to the support of $\tilde{f}$. Hence we can find a constant such that

$$\frac{1}{|\alpha'(r)|} \le \frac{const}{1+|\lambda|+|t|}, \quad r \in (a,b), \quad (10.36)$$

and λ, t as above. Using this we can estimate the integral in (10.34)

$$\left|\int_0^\infty \tilde{f}(r)\frac{1}{\alpha'(r)}\frac{d}{dr}\mathrm{e}^{\mathrm{i}\alpha(r)}dr\right| \le \frac{const}{1+|\lambda|+|t|}\left|\int_0^\infty \tilde{f}'(r)\mathrm{e}^{\mathrm{i}\alpha(r)}dr\right|, \quad (10.37)$$

(the last step uses integration by parts) for λ, t as above. By increasing the constant we can even assume that it holds for $t \ge 0$ and $\lambda \le -R$. Moreover, by iterating the last estimate we see

$$|(\mathcal{M}K_{t,x})(\lambda,\omega)| \le \frac{const}{(1+|\lambda|+|t|)^N} \quad (10.38)$$

for any $N \in \mathbb{N}$ and $t \ge 0$ and $\lambda \le -R$. This finishes the proof. □

Corollary 10.6. *Suppose that $f \in C_0^\infty((0,\infty))$ and $R \in \mathbb{R}$. Then the operator $P_D((\pm R,\pm\infty))f(H_0)\exp(-\mathrm{i}tH_0)$ converges strongly to 0 as $t \to \mp\infty$.*

Proof. Abbreviating $P_D = P_D((\pm R,\pm\infty))$ and $\chi = \chi_{\{x|\,|x|<2a|t|\}}$ we have

$$\|P_D f(H_0)\mathrm{e}^{-\mathrm{i}tH_0}\psi\| \le \|\chi\mathrm{e}^{\mathrm{i}tH_0}f(H_0)^*P_D\|\,\|\psi\| + \|f(H_0)\|\|(\mathbb{I}-\chi)\psi\|. \quad (10.39)$$

since $\|A\| = \|A^*\|$. Taking $t \to \mp\infty$ the first term goes to zero by our lemma and the second goes to zero since $\chi\psi \to \psi$. □

10.3. Schrödinger operators with short range potentials

By the RAGE theorem we know that for $\psi \in \mathfrak{H}_c$, $\psi(t)$ will eventually leave every compact ball (at least on the average). Hence we expect that the time evolution will asymptotically look like the free one for $\psi \in \mathfrak{H}_c$ if the potential decays sufficiently fast. In other words, we expect such potentials to be asymptotically complete.

Suppose V is relatively bounded with bound less than one. Introduce

$$h_1(r) = \|VR_{H_0}(z)\chi_r\|, \quad h_2(r) = \|\chi_r VR_{H_0}(z)\|, \quad r \ge 0, \tag{10.40}$$

where

$$\chi_r = \chi_{\{x|\, |x| \ge r\}}. \tag{10.41}$$

The potential V will be called **short range** if these quantities are integrable. We first note that it suffices to check this for h_1 or h_2 and for one $z \in \rho(H_0)$.

Lemma 10.7. *The function h_1 is integrable if and only if h_2 is. Moreover, h_j integrable for one $z_0 \in \rho(H_0)$ implies h_j integrable for all $z \in \rho(H_0)$.*

Proof. Pick $\phi \in C_0^\infty(\mathbb{R}^n, [0,1])$ such that $\phi(x) = 0$ for $0 \le |x| \le 1/2$ and $\phi(x) = 0$ for $1 \le |x|$. Then it is not hard to see that h_j is integrable if and only if $\tilde{h}_j$ is integrable, where

$$\tilde{h}_1(r) = \|VR_{H_0}(z)\phi_r\|, \quad \tilde{h}_2(r) = \|\phi_r VR_{H_0}(z)\|, \quad r \ge 1, \tag{10.42}$$

and $\phi_r(x) = \phi(x/r)$. Using

$$\begin{aligned} [R_{H_0}(z), \phi_r] &= -R_{H_0}(z)[H_0(z), \phi_r]R_{H_0}(z) \\ &= R_{H_0}(z)(\Delta\phi_r + (\partial\phi_r)\partial)R_{H_0}(z) \end{aligned} \tag{10.43}$$

and $\Delta\phi_r = \phi_{r/2}\Delta\phi_r$, $\|\Delta\phi_r\|_\infty \le \|\Delta\phi\|_\infty/r^2$ respectively $(\partial\phi_r) = \phi_{r/2}(\partial\phi_r)$, $\|\partial\phi_r\|_\infty \le \|\partial\phi\|_\infty/r^2$ we see

$$|\tilde{h}_1(r) - \tilde{h}_2(r)| \le \frac{c}{r}\tilde{h}_1(r/2), \quad r \ge 1. \tag{10.44}$$

Hence $\tilde{h}_2$ is integrable if $\tilde{h}_1$ is. Conversely,

$$\tilde{h}_1(r) \le \tilde{h}_2(r) + \frac{c}{r}\tilde{h}_1(r/2) \le \tilde{h}_2(r) + \frac{c}{r}\tilde{h}_2(r/2) + \frac{2c}{r^2}\tilde{h}_1(r/4) \tag{10.45}$$

shows that $\tilde{h}_2$ is integrable if $\tilde{h}_1$ is.

Invoking the first resolvent formula

$$\|\phi_r VR_{H_0}(z)\| \le \|\phi_r VR_{H_0}(z_0)\|\|\mathbb{I} - (z - z_0)R_{H_0}(z)\| \tag{10.46}$$

finishes the proof. □

As a first consequence note

Lemma 10.8. *If V is short range, then $R_H(z) - R_{H_0}(z)$ is compact.*

Proof. The operator $R_H(z)V(\mathbb{I}-\chi_r)R_{H_0}(z)$ is compact since $(\mathbb{I}-\chi_r)R_{H_0}(z)$ is by Lemma 6.9 and $R_H(z)V$ is bounded by Lemma 8.15. Moreover, by our short range condition it converges in norm to

$$R_H(z)VR_{H_0}(z) = R_H(z) - R_{H_0}(z) \tag{10.47}$$

as $r \to \infty$ (at least for some subsequence). □

In particular, by Weyl's theorem we have $\sigma_{ess}(H) = [0, \infty)$. Moreover, V short range implies that H and H_0 look alike far outside.

Lemma 10.9. *Suppose $R_H(z) - R_{H_0}(z)$ is compact, then so is $f(H) - f(H_0)$ for any $f \in C_\infty(\mathbb{R})$ and*

$$\lim_{r\to\infty} \|(f(H) - f(H_0))\chi_r\| = 0. \tag{10.48}$$

Proof. The first part is Lemma 8.13 and the second part follows from part (2) of Lemma 8.8 since χ_r converges strongly to 0. □

However, this is clearly not enough to prove asymptotic completeness and we need a more careful analysis. The main ideas are due to Enß [**4**].

We begin by showing that the wave operators exist. By Cook's criterion (Lemma 10.3) we need to show that

$$\begin{aligned}\|V \exp(\mp \mathrm{i} t H_0)\psi\| \le\ & \|V R_{H_0}(-1)\| \|(\mathbb{I} - \chi_{2a|t|}) \exp(\mp \mathrm{i} t H_0)(H_0 + \mathbb{I})\psi\| \\ & + \|V R_{H_0}(-1)\chi_{2a|t|}\| \|(H_0 + \mathbb{I})\psi\| \end{aligned} \tag{10.49}$$

is integrable for a dense set of vectors ψ. The second term is integrable by our short range assumption. The same is true by Perry's estimate (Lemma 10.5) for the first term if we choose $\psi = f(H_0)P_D((\pm R, \pm\infty))\varphi$. Since vectors of this form are dense, we see that the wave operators exist,

$$\mathfrak{D}(\Omega_\pm) = \mathfrak{H}. \tag{10.50}$$

Since H restricted to $\mathrm{Ran}(\Omega_\pm^*)$ is unitarily equivalent to H_0, we obtain $[0, \infty) = \sigma_{ac}(H_0) \subseteq \sigma_{ac}(H)$. And by $\sigma_{ac}(H) \subseteq \sigma_{ess}(H) = [0, \infty)$ we even have $\sigma_{ac}(H) = [0, \infty)$.

To prove asymptotic completeness of the wave operators we will need that $(\Omega_\pm - \mathbb{I})f(H_0)P_\pm$ are compact.

Lemma 10.10. *Let $f \in C_0^\infty((0, \infty))$ and suppose ψ_n converges weakly to 0. Then*

$$\lim_{n\to\infty} \|(\Omega_\pm - \mathbb{I})f(H_0)P_\pm\psi_n\| = 0, \tag{10.51}$$

that is, $(\Omega_\pm - \mathbb{I})f(H_0)P_\pm$ is compact.

Proof. By (10.13) we see

$$\|R_H(z)(\Omega_\pm - \mathbb{I})f(H_0)P_\pm\psi_n\| \le \int_0^\infty \|R_H(z)V \exp(-\mathrm{i} s H_0)f(H_0)P_\pm\psi_n\| dt. \tag{10.52}$$

Since $R_H(z)VR_{H_0}$ is compact we see that the integrand

$$\begin{aligned} R_H(z)V \exp(-\mathrm{i} s H_0)f(H_0)P_\pm\psi_n = \\ R_H(z)VR_{H_0} \exp(-\mathrm{i} s H_0)(H_0 + 1)f(H_0)P_\pm\psi_n \end{aligned} \tag{10.53}$$

converges pointwise to 0. Moreover, arguing as in (10.49) the integrand is bounded by an L^1 function depending only on $\|\psi_n\|$. Thus $R_H(z)(\Omega_\pm - \mathbb{I})f(H_0)P_\pm$ is compact by the dominated convergence theorem. Furthermore, using the intertwining property we see that

$$(\Omega_\pm - \mathbb{I})\tilde{f}(H_0)P_\pm = R_H(z)(\Omega_\pm - \mathbb{I})f(H_0)P_\pm - (R_H(z) - R_{H_0}(z))f(H_0)P_\pm \qquad (10.54)$$

is compact by Lemma 8.13, where $\tilde{f}(\lambda) = (\lambda + 1)f(\lambda)$. □

Now we have gathered enough information to tackle the problem of asymptotic completeness.

We first show that the singular continuous spectrum is absent. This is not really necessary, but avoids the use of Cesàro means in our main argument.

Abbreviate $P = P_H^{sc} P_H((a, b))$, $0 < a < b$. Since H restricted to $\mathrm{Ran}(\Omega_\pm)$ is unitarily equivalent to H_0 (which has purely absolutely continuous spectrum), the singular part must live on $\mathrm{Ran}(\Omega_\pm)^\perp$, that is, $P_H^{sc}\Omega_\pm = 0$. Thus $Pf(H_0) = P(\mathbb{I}-\Omega_+)f(H_0)P_+ + P(\mathbb{I}-\Omega_-)f(H_0)P_-$ is compact. Since $f(H) - f(H_0)$ is compact, it follows that $Pf(H)$ is also compact. Choosing f such that $f(\lambda) = 1$ for $\lambda \in [a, b]$ we see that $P = Pf(H)$ is compact and hence finite dimensional. In particular $\sigma_{sc}(H) \cap (a, b)$ is a finite set. But a continuous measure cannot be supported on a finite set, showing $\sigma_{sc}(H) \cap (a, b) = \emptyset$. Since $0 < a < b$ are arbitrary we even have $\sigma_{sc}(H) \cap (0, \infty) = \emptyset$ and by $\sigma_{sc}(H) \subseteq \sigma_{ess}(H) = [0, \infty)$ we obtain $\sigma_{sc}(H) = \emptyset$.

Observe that replacing P_H^{sc} by P_H^{pp} the same argument shows that all nonzero eigenvalues are finite dimensional and cannot accumulate in $(0, \infty)$.

In summary we have shown

Theorem 10.11. *Suppose V is short range. Then*

$$\sigma_{ac}(H) = \sigma_{ess}(H) = [0, \infty), \qquad \sigma_{sc}(H) = \emptyset. \qquad (10.55)$$

All nonzero eigenvalues have finite multiplicity and cannot accumulate in $(0, \infty)$.

Now we come to the anticipated asymptotic completeness result of Enß. Choose

$$\psi \in \mathfrak{H}_c(H) = \mathfrak{H}_{ac}(H) \quad \text{such that} \quad \psi = f(H)\psi \qquad (10.56)$$

for some $f \in C_0^\infty((0, \infty))$. By the RAGE theorem the sequence $\psi(t)$ converges weakly to zero as $t \to \pm\infty$. Abbreviate $\psi(t) = \exp(-\mathrm{i}tH)\psi$. Introduce

$$\varphi_\pm(t) = f(H_0)P_\pm\psi(t). \qquad (10.57)$$

which satisfy

$$\lim_{t\to\pm\infty} \|\psi(t) - \varphi_+(t) - \varphi_-(t)\| = 0. \tag{10.58}$$

Indeed this follows from

$$\psi(t) = \varphi_+(t) + \varphi_-(t) + (f(H) - f(H_0))\psi(t) \tag{10.59}$$

and Lemma 8.13. Moreover, we even have

$$\lim_{t\to\pm\infty} \|(\Omega_\pm - \mathbb{I})\varphi_\pm(t)\| = 0 \tag{10.60}$$

by Lemma 10.10. Now suppose $\psi \in \mathrm{Ran}(\Omega_\pm)^\perp$, then

$$\begin{aligned} \|\psi\|^2 &= \lim_{t\to\pm\infty} \langle \psi(t), \psi(t)\rangle \\ &= \lim_{t\to\pm\infty} \langle \psi(t), \varphi_+(t) + \varphi_-(t)\rangle \\ &= \lim_{t\to\pm\infty} \langle \psi(t), \Omega_+\varphi_+(t) + \Omega_-\varphi_-(t)\rangle. \end{aligned} \tag{10.61}$$

By Theorem 10.2, $\mathrm{Ran}(\Omega_\pm)^\perp$ is invariant under H and hence $\psi(t) \in \mathrm{Ran}(\Omega_\pm)^\perp$ implying

$$\begin{aligned} \|\psi\|^2 &= \lim_{t\to\pm\infty} \langle \psi(t), \Omega_\mp\varphi_\mp(t)\rangle \qquad (10.62)\\ &= \lim_{t\to\pm\infty} \langle P_\mp f(H_0)^* \Omega_\mp^* \psi(t), \psi(t)\rangle. \end{aligned}$$

Invoking the intertwining property we see

$$\|\psi\|^2 = \lim_{t\to\pm\infty} \langle P_\mp f(H_0)^* \mathrm{e}^{-\mathrm{i}tH_0} \Omega_\mp^* \psi, \psi(t)\rangle = 0 \tag{10.63}$$

by Corollary 10.6. Hence $\mathrm{Ran}(\Omega_\pm) = \mathfrak{H}_{ac}(H) = \mathfrak{H}_c(H)$ and we thus have shown

Theorem 10.12 (Enß)**.** *Suppose V is short range, then the wave operators are asymptotically complete.*

For further results and references see for example [**3**].

Bibliography

[1] W. O. Amrein, J. M. Jauch, and K. B. Sinha, *Scattering Theory in Quantum Mechanics*, W. A. Benajmin Inc., New York, 1977

[2] H. Bauer, *Measure and Integration Theory*, de Gruyter, Berlin, 2001.

[3] H. L. Cycon, R. G. Froese, W. Kirsch and B. Simon, *Schrödinger Operators*, Springer, Berlin, 1987.

[4] V. Enss, *Asymptotic completeness for Quantum mechanical potential scattering*, Comm. Math. Phys. **61**, 285-291 (1978).

[5] V. Enß, *Schrödinger Operators*, lecture notes (private communication).

[6] P. D. Hislop and I. M. Sigal, *Introduction to Spectral Theory* Springer, New York, 1996.

[7] Yoram Last, *Quantum dynamics and decompositions of singular continuous spectra*, J. Funct. Anal. **142**, 406-445 (1996).

[8] P. Perry, *Mellin transforms and scattering Theory*, Duke Math. J. **47**, 187-193 (1987).

[9] E. Prugovečki, *Quantum Mechanics in Hilbert Space*, 2nd edition, Academic Press, New York, 1981.

[10] M. Reed and B. Simon, *Methods of Modern Mathematical Physics I. Functional Analysis*, rev. and enl. edition, Academic Press, San Diego, 1980.

[11] M. Reed and B. Simon, *Methods of Modern Mathematical Physics II. Fourier Analysis, Self-Adjointness*, Academic Press, San Diego, 1975.

[12] M. Reed and B. Simon, *Methods of Modern Mathematical Physics III. Scattering Theory*, Academic Press, San Diego, 1979.

[13] M. Reed and B. Simon, *Methods of Modern Mathematical Physics IV. Analysis of Operators*, Academic Press, San Diego, 1978.

[14] W. Rudin, *Real and Complex Analysis*, 3rd edition, McGraw-Hill, New York, 1987.

[15] M. Schechter, *Operator Methods in Quantum Mechanics*, North Holland, New York, 1981.

[16] M. Schechter, *Spectra of Partial Differential Operators*, 2nd ed., North Holland, Amsterdam, 1986.

[17] W. Thirring, *Quantum Mechanics of Atoms and Molecules*, Springer, New York, 1981.

[18] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, 2nd ed., Cambridge University Press, Cambridge, 1962.

[19] J. Weidmann, *Linear Operators in Hilbert Spaces*, Springer, New York, 1980.

Glossary of notations

$AC(I)$	... absolutely continuous functions, 29
$\mathfrak{B}$	$= \mathfrak{B}^1$
$\mathfrak{B}^n$	... Borel σ-field of $\mathbb{R}^n$, 3.
$\mathfrak{C}(\mathfrak{H})$	... set of compact operators, 79.
$C(U)$	... set of continuous functions from U to $\mathbb{C}$.
$C_\infty(U)$	... set of functions in $C(U)$ which vanish at ∞.
$C(U,V)$	... set of continuous functions from U to V.
$C_0^\infty(U,V)$	... set of compactly supported smooth functions
$\chi_\Omega(.)$	... characteristic function of the set Ω
dim	... dimension of a linear space
$\mathfrak{D}(.)$	... domain of an operator
e	... exponential function, $\mathrm{e}^z = \exp(z)$
$\mathbb{E}(A)$	... expectation of an operator A, 23
$\mathcal{F}$	... Fourier transform, 85
H	... Schrödinger operator, 109
H_0	... free Schrödinger operator, 88
$H^j(\mathbb{R}^n)$	... Sobolev space, 87
hull(.)	... convex hull
$\mathfrak{H}$	... a separable Hilbert space
i	... complex unity, $\mathrm{i}^2 = -1$
Im(.)	... imaginary part of a complex number
inf	... infimum
Ker(A)	... kernel of an operator A
$\mathfrak{L}(X,Y)$	... set of all bounded linear operators from X to Y, 10
$\mathfrak{L}(X)$	$= \mathfrak{L}(X,X)$
$L^p(M,d\mu)$	... Lebesgue space of p integrable functions, 10

$L^\infty(M, d\mu)$	...Lebesgue space of bounded functions, 11
$L^\infty_\infty(\mathbb{R}^n)$	...Lebesgue space of bounded functions vanishing at ∞
λ	...a real number
max	...maximum
$\mathcal{M}$	...Mellin transform, 139
μ_ψ	...spectral measure, 51
$\mathbb{N}$	...the set of positive integers
$\mathbb{N}_0$	$= \mathbb{N} \cup \{0\}$
Ω	...a Borel set
$\Omega_\pm$	...wave operators, 135
$P_A(.)$	...family of spectral projections of an operator A
$P_\pm$	...projector onto outgoing/incoming states, 138
$\mathfrak{Q}(.)$	...form domain of an operator, 53
$R(I, X)$	...set of regulated functions, 64
$R_A(z)$	...resolvent of A, 36
$\mathrm{Ran}(A)$	...range of an operator A
$\mathrm{rank}(A)$	$= \dim \mathrm{Ran}(A)$, rank of an operator A, 79
$\mathrm{Re}(.)$	...real part of a complex number
$\rho(A)$	...resolvent set of A, 36
$\mathbb{R}$	...the set of real numbers
$S(I, X)$	...set of simple functions, 64
$\mathcal{S}(\mathbb{R}^n)$	...set of smooth functions with rapid decay, 85
sup	...supremum
supp	...support of a function
$\sigma(A)$	...spectrum of an operator A, 36
$\sigma_{ac}(A)$	...absolutely continuous spectrum of A, 57
$\sigma_{sc}(A)$	...singular continuous spectrum of A, 57
$\sigma_{pp}(A)$	...pure point spectrum of A, 57
$\sigma_p(A)$	...point spectrum (set of eigenvalues) of A, 57
$\sigma_d(A)$	...discrete spectrum of A, 103
$\sigma_{ess}(A)$	...essential spectrum of A, 103
$\mathrm{span}(M)$	...set of finite linear combinations from M, 9
$\mathbb{Z}$	...the set of integers
z	...a complex number

$\mathbb{I}$	... identity operator
$\sqrt{z}$	... square root of z with branch cut along $(-\infty, 0)$
z^*	... complex conjugation
A^*	... adjoint of A, 27
$\overline{A}$	... closure of A, 31
$\hat{f}$	$= \mathcal{F}f$, Fourier transform of f
$\check{f}$	$= \mathcal{F}^{-1}f$, inverse Fourier transform of f
$\|.\|$	... norm in the Hilbert space $\mathfrak{H}$
$\|.\|_p$	... norm in the Banach space L^p
$\langle ., .. \rangle$	... scalar product in $\mathfrak{H}$
$\oplus$	... orthogonal sum of linear spaces or operators, 20
Δ	... Laplace operator, 88
∂	... gradient, 85
∂_α	... derivative, 85
$M^\perp$	... orthogonal complement, 19
(λ_1, λ_2)	$= \{\lambda \in \mathbb{R} \,\vert\, \lambda_1 < \lambda < \lambda_2\}$, open interval
$[\lambda_1, \lambda_2]$	$= \{\lambda \in \mathbb{R} \,\vert\, \lambda_1 \le \lambda \le \lambda_2\}$, closed interval

Index