Spectral theory of differential operators

Lecture notes of a one-semester introductory course

Carl von Ossietzky University of Oldenburg
April – July 2022

Konstantin Pankrashkin
http://uol.de/pankrashkin

Contents

Notation

Here we list some conventions used throughout the text.

The symbol $\mathbb{N}$ denotes the sets of the natural numbers starting from 1, while $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

In what follows the word combination "Hilbert space" should be understood as "separable complex Hilbert space". If the symbol "$\mathcal{H}$" appears without explanations, it denotes a certain Hilbert space (in the above sense).

If $\mathcal{H}$ is a Hilbert space and $x, y \in \mathcal{H}$, then by $\langle x, y\rangle$ we denote the scalar product of x and y. If there is more than one Hilbert space in play, we use the more detailed notation $\langle x, y\rangle_{\mathcal{H}}$. We always assume that the scalar product is linear with respect to the *second* argument and conjugate linear with respect to the first one, i.e. that for all $\alpha \in \mathbb{C}$ we have $\langle x, \alpha y\rangle = \langle \overline{\alpha} x, y\rangle = \alpha\langle x, y\rangle$. This means, for example, that the scalar product in the standard space $L^2(\Omega)$ is defined by

$$\langle f, g\rangle = \int_\Omega \overline{f(x)} g(x)\,\mathrm{d}x.$$

If A is a finite or countable set, we denote by $\ell^2(A)$ the vector space of the functions $\xi : A \to \mathbb{C}$ with

$$\sum_{a\in A} \big|\xi(a)\big|^2 < \infty,$$

which is a Hilbert space with the scalar product

$$\langle \xi, \xi'\rangle = \sum_{a\in A} \overline{\xi(a)}\xi'(a).$$

We will make an exception for the scalar product notaion in Euclidean spaces. If $z = (z_1, \dots, z_n), w = (w_1, \dots, w_n) \in \mathbb{C}^n$ we denote

$$z \cdot w := z_1 w_1 + \cdots + z_n w_n.$$

In particular, the standard scalar product of z and w in $\mathbb{C}^n$ is then given by $\overline{z} \cdot w$.

If $\mathcal{H}$ and $\mathcal{G}$ are Hilbert spaces, then by $\mathcal{B}(\mathcal{H}, \mathcal{G})$ and $\mathcal{K}(\mathcal{H}, \mathcal{G})$ we denote the spaces of the continuous linear operators and the one of the compact operators from $\mathcal{H}$ and $\mathcal{G}$, respectively. Furtheremore, $\mathcal{B}(\mathcal{H}) := \mathcal{B}(\mathcal{H}, \mathcal{H})$ and $\mathcal{K}(\mathcal{H}) := \mathcal{K}(\mathcal{H}, \mathcal{H})$.

Recommended books

- The very first version of the lecture notes was based on a preliminary version of the book

  B. Helffer: *Spectral theory and its applications.* Cambridge University Press, 2012.

- The following recent textbook has rapidly become very popular:

  D. Borthwick: *Spectral theory. Basic concepts and applications.* Springer, 2020.

Additional references on particular topics will be given during the course.

At many points we will be obliged to use some facts on distributions and Sobolev spaces. I tried to include some elementary facts in these notes with partial proofs) and I hope that it will be sufficient. Nevertheless, if one wants to study these questions in details, I recommend to study the textbook

- G. Grubb: *Distributions and operators.* Springer, 2011.

and/or to follow a dedicated course on partial differential equations and distributions.

1 Unbounded operators

1.1 Closed and closable operators

A *linear operator* T in $\mathcal{H}$ is a linear map from a subspace (the domain of T) $D(T) \subset \mathcal{H}$ to $\mathcal{H}$. The range of T is the set $\operatorname{ran} T := \{Tx : x \in D(T)\}$. We say that a linear operator T is *bounded* if the quantity

$$\mu(T) := \sup_{\substack{x \in D(T) \\ x \neq 0}} \frac{\|Tx\|}{\|x\|}$$

is finite. In what follows, the word combination "an unbounded operator" should be understood as "an operator which is not assumed to be bounded". If $D(T) = \mathcal{H}$ and T is bounded, we arrive at the notion of a continuous linear operator in $\mathcal{H}$; the space of such operators is denoted by $\mathcal{B}(\mathcal{H})$. This is a Banach space equipped with the norm $\|T\| := \mu(T)$.

During the whole course, by considering a linear operator we always assume that its domain is dense (if the contrary is not stated explicitly).

If T is a bounded operator in $\mathcal{H}$, it can be uniquely extended to a continuous linear operator. Let us discuss a similar idea for unbounded operators.

The *graph* of a linear operator T in $\mathcal{H}$ is the set

$$\operatorname{gr} T := \big\{(x, Tx) : x \in D(T)\big\} \subset \mathcal{H} \times \mathcal{H}.$$

For two linear operators T_1 and T_2 in $\mathcal{H}$ we write $T_1 \subset T_2$ if $\operatorname{gr} T_1 \subset \operatorname{gr} T_2$. In other words, $T_1 \subset T_2$ means that $D(T_1) \subset D(T_2)$ and that $T_2 x = T_1 x$ for all $x \in D(T_1)$; the operator T_2 is then called an *extension* of T_1 and T_1 is called a *restriction* of T_2, and one writes $T_1 = T_2|_{D(T_1)}$.

Definition 1.1 (Closed operator, closable operator).

- A linear operator T in $\mathcal{H}$ is called *closed* if its graph is a closed subspace in $\mathcal{H} \times \mathcal{H}$.
- A linear operator T in $\mathcal{H}$ is called *closable*, if the closure $\overline{\operatorname{gr} T}$ of the graph of T in $\mathcal{H} \times \mathcal{H}$ is still the graph of a certain operator linear $\overline{T}$. This $\overline{T}$ is then called the *closure* of T. □

The following propositions follows directly from the above definition:

Proposition 1.2. *A linear operator T in $\mathcal{H}$ is closed if and only if the three conditions*

- $x_n \in D(T)$,
- x_n *converge to* x *in* $\mathcal{H}$,
- Tx_n *converge to* y *in* $\mathcal{H}$

imply the inclusion $x \in D(T)$ and the equality $y = Tx$.

Proposition 1.3. *A linear operator T in $\mathcal{H}$ is closable if and only if for any two sequences $(x_n) \subset D(T)$ and $(x'_n) \subset D(T)$ such that:*

- $\lim x_n = \lim x'_n =: x$,
- *there exist the limits $y := \lim Tx_n$ and $y' := \lim Tx'_n$,*

there holds $y = y'$.

Definition 1.4 (Graph norm). Let T be a linear operator in $\mathcal{H}$. Define on $D(T)$ a new scalar product by

$$\langle x, y\rangle_T = \langle x, y\rangle + \langle Tx, Ty\rangle.$$

The associated norm $\|x\|_T := \sqrt{\langle x, x\rangle_T} = \sqrt{\|x\|^2 + \|Tx\|^2}$ is called the *graph norm* for T. $\square$

The following assertion is also evident.

Proposition 1.5. *Let T be a linear operator in $\mathcal{H}$.*

- *T is closed iff $D(T)$ is complete in the graph norm (or, equivalently, if $D(T)$ equipped with the scalar product $\langle\cdot,\cdot\rangle_T$ is a Hilbert space).*
- *If T is closable, then $D(\overline{T})$ is exactly the completion of $D(T)$ with respect to the graph norm.*

Informally, one can say that $D(\overline{T})$ consists of those x for which there is a unique candidate for $\overline{T}x$ if one tries to extend T by density.

Let us consider some simple examples. More sophisticated examples involving differential operators will be discussed later in Section 1.4.

Example 1.6 (Bounded linear operators are closed). By the closed graph theorem, a linear operator T in $\mathcal{H}$ with $D(T) = \mathcal{H}$ is closed if and only if it is bounded. In this course we consider mostly unbounded closed operators.

Example 1.7 (Multiplication operator). Let $\Omega \subset \mathbb{R}^d$ be an open set and $\mathcal{H} := L^2(\Omega)$ and pick $f \in C^0(\Omega)$. Introduce a linear operator M_f in $\mathcal{H}$ as follows:

$$D(M_f) = \{u \in L^2(\Omega) : fu \in L^2(\Omega)\} \text{ and } \quad M_f u = fu \text{ for } u \in D(M_f).$$

It can be easily seen that $D(M_f)$ equipped with the graph norm coincides with the weighted space $L^2\big(\Omega, (1 + |f(x)|^2)dx\big)$, which is complete. This shows that M_f is closed.

On the other hand, denote by T the restriction of M_f to the functions with compact supports. The functions vanishing outside compact subsets of Ω are dense in $L^2\big(\Omega, (1 + |f|^2)dx\big)$, hence, the closure $\overline{T}$ of T is exactly M_f. It also follows that that T is not closed.

It is clear that the example can be generalized by taking f with lower regularity: the continuity is not really needed, but one needs f to be bounded on each compact subset if one one wants to be sure that $fu \in L^2(\mathbb{R}^d)$ for any $u \in L^2(\mathbb{R}^d)$ vanishing outside a compact subset.

Example 1.8 (Non-closable operator). Take $\mathcal{H} = L^2(\mathbb{R})$ and pick a $g \in \mathcal{H}$ with $g \neq 0$. Consider the operator L defined on $D(L) = C^0(\mathbb{R}) \cap L^2(\mathbb{R})$ by $Lf = f(0)g$.

One can find two sequences $(f_n), (g_n)$ in $D(L)$ such that both converge in the L^2 norm to f but such that $f_n(0) = 0$ and $g_n(0) = 1$ for all n. Then $Lf_n = 0$, $Lg_n = g$ for all n, and both sequences Lf_n and Lg_n converge, but to different limits. This implies that L is not closable (Proposition 1.3). □

Remark 1.9 (Graph="no vertical lines"). It is easy to see that a linear subspace $V \subset \mathcal{H} \times \mathcal{H}$ is the graph of a linear operator if and only if it does not contain any vector of the form $(0, x)$ with $x \neq 0$, i.e. if V does not contain any "vertical line" $0 \times \mathbb{C}x$ with $x \neq 0$. Then the following trivial observation will be useful several times:

Proposition 1.10. *If a linear operator T is closable and S is another linear operator with $S \subset T$, then S is closable too, with $\overline{S} \subset \overline{T}$.*

Proof. One has $\operatorname{gr} S \subset \operatorname{gr} T$, then $\overline{\operatorname{gr} S} \subset \overline{\operatorname{gr} T}$. As $\overline{\operatorname{gr} T}$ is the graph of a linear operator, it does not contain "vertical lines" (Remark 1.9), then the same holds for $\overline{\operatorname{gr} S}$, and then $\overline{\operatorname{gr} S}$ is the graph of some linear operator. □

In most areas of analysis one only works with closable operators (some reasons for that will become evident later in this course).

1.2 Adjoint, symmetric, self-adjoint operators

Recall that for $T \in \mathcal{B}(\mathcal{H})$ its adjoint T^* is defined by the relation

$$\langle x, Ty \rangle = \langle T^* x, y \rangle \text{ for all } x, y \in \mathcal{H}.$$

The proof of the existence comes from the Riesz representation theorem: for each $x \in \mathcal{H}$ the map $\mathcal{H} \ni y \mapsto \langle x, Ty \rangle \in \mathbb{C}$ is a continuous linear functional. By Riesz theorem, there exists a unique vector, denoted by T^*x, with $\langle x, Ty \rangle = \langle T^*x, y \rangle$ for all $y \in \mathcal{H}$. One then shows easily that the map $x \mapsto T^*x$ is linear, and by estimating the scalar product one shows that T^* is also continuous. Let us use the same idea for unbounded operators.

Definition 1.11 (Adjoint operator). If T be a linear operator in $\mathcal{H}$ (with dense domain!), then its *adjoint* T^* is defined as follows. The domain $D(T^*)$ consists of the vectors $u \in \mathcal{H}$ for which the map $D(T) \ni v \mapsto \langle u, Tv \rangle \in \mathbb{C}$ is bounded with respect to the $\mathcal{H}$-norm. For such u there exists, by the Riesz theorem, a unique vector denoted by T^*u such that $\langle u, Tv \rangle = \langle T^*u, v \rangle$ for all $v \in D(T)$. □

We note that the implicit assumption $\overline{D(T)} = \mathcal{H}$ is important here: if it is not satisfied, then the value T^*u is not uniquely determined, one can add to T^*u an arbitrary vector from $D(T)^\perp$. As an easy exercise one shows that $T^* : D(T^*) \to \mathcal{H}$ is a linear map.

Let us give a geometric interpretation of the adjoint operator. Recall first that $\mathcal{H}\times\mathcal{H}$ can be viewed as a Hilbert space with the scalar product

$$\big\langle (x,y),(x',y')\big\rangle_{\mathcal{H}\times\mathcal{H}} := \langle x,x'\rangle_{\mathcal{H}} + \langle y,y'\rangle_{\mathcal{H}}.$$

Consider a unitary linear operator

$$J:\mathcal{H}\times\mathcal{H}\to\mathcal{H}\times\mathcal{H},\quad J(x,y)=(y,-x).$$

It is easy to check that J commutes with the operation of orthogonal complement in $\mathcal{H}\times\mathcal{H}$, i.e. $J(V)^\perp = J(V^\perp)$ for any $V\subset\mathcal{H}\times\mathcal{H}$. Then Definition 1.11 can be reformulated as follows:

Proposition 1.12 (Geometric interpretation of the adjoint). *Let T be a linear operator in $\mathcal{H}$. The following two assertions are equivalent:*

- *$u\in D(T^*)$ and $f=T^*u$,*
- *$\big\langle (u,T^*u), J(v,Tv)\big\rangle_{\mathcal{H}\times\mathcal{H}} = 0$ for all $v\in D(T)$.*

In other words,

$$\operatorname{gr} T^* = J(\operatorname{gr} T)^\perp. \tag{1.1}$$

As a simple application we obtain

Proposition 1.13. *One has $(\overline{T}\,)^* = T^*$, and T^* is a closed operator.*

Proof. Follows from (1.1): the orthogonal complement is always closed, and $J(\operatorname{gr} T)^\perp = J(\overline{\operatorname{gr} T})^\perp$. □

Up to now we do not know if the domain of the adjoint contains non-zero vectors. This is discussed in the following proposition.

Proposition 1.14 (Domain of the adjoint). *Let T be a closable operator $\mathcal{H}$, then:*

(i) $D(T^*)$ is a dense subspace of $\mathcal{H}$,

(ii) $T^{**} := (T^*)^* = \overline{T}$.

Proof. The item (ii) easily follows from (i) and (1.1): one applies the same operations again and remark that $J^2=-1$ and that taking twice the orthogonal complement results in taking the closure.

Now let us prove the item (i). Let a vector $w\in\mathcal{H}$ be orthogonal to $D(T^*)$: $\langle u,w\rangle = 0$ for all $u\in D(T^*)$. Then one has $\langle J(u,T^*u),(0,w)\rangle_{\mathcal{H}\times\mathcal{H}} \equiv \langle u,w\rangle + \langle T^*u,0\rangle = 0$ for all $u\in D(T^*)$, which means that $(0,w)\in J(\operatorname{gr} T^*)^\perp = \overline{\operatorname{gr} T}$. As the operator T is closable, its closure $\overline{T}$ is defined and $\overline{\operatorname{gr} T} = \operatorname{gr}\overline{T}$. Then $(0,w)\in \operatorname{gr}\overline{T}$, i.e. $w=\overline{T}0=0$. □

Let us look at some examples.

Example 1.15 (Adjoint for bounded operators). The general definition of the adjoint operator is compatible with the one for continuous linear operators.

Example 1.16. As an exercise, one can show that for the multiplication operator M_f from example 1.7 one has $(M_f)^* = M_{\overline{f}}$. □

The following definition introduces further classes of linear operator that will be studied throughout the course.

Definition 1.17 (Symmetric, self-adjoint, essentially self-adjoint operators). We say that a linear operator T in $\mathcal{H}$ is *symmetric* (or *Hermitian*) if

$$\langle u, Tv\rangle = \langle Tu, v\rangle \quad \text{for all } u, v \in D(T),$$

or, equivalently, if $T \subset T^*$. Furthermore:

- T is called *self-adjoint* if $T = T^*$,
- T is called *essentially self-adjoint* if $\overline{T}$ is self-adjoint (i.e. if $\overline{T} = T^*$). □

Proposition 1.18. *Symmetric operators are closable.*

Proof. If T is symmetric, then $T \subset T^*$. As T^* is closed (in particular,, closable), one can uses Proposition 1.10. □

Example 1.19 (Bounded symmetric operators are self-adjoint). If $T \in \mathcal{B}(\mathcal{H})$, then T is symmetric if and only if T is self-adjoint. But the equivalence does not hold for unbounded operators: we will see it soon!

Example 1.20 (Self-adjoint multiplication operators). As follows from Example 1.16, the multiplication operator M_f in Example 1.7 is self-adjoint iff $f(x) \in \mathbb{R}$ for all $x \in \mathbb{R}^d$. □

A large class of self-adjoint operators comes from the following proposition.

Proposition 1.21. *Let T be an* injective *self-adjoint operator, then its inverse is also self-adjoint.*

Proof. We show first that $D(T^{-1}) := \operatorname{ran} T$ is dense in $\mathcal{H}$. Let $u \perp \operatorname{ran} T$, then $\langle u, Tv\rangle = 0$ for all $v \in D(T)$. This can be rewritten as $\langle u, Tv\rangle = \langle 0, v\rangle$ for all $v \in D(T)$, which shows that $u \in D(T^*)$ and $T^*u = 0$. As $T^* = T$, we have $u \in D(T)$ and $Tu = 0$. As T in injective, one has $u = 0$

Now consider the operator $S : \mathcal{H} \times \mathcal{H} \to \mathcal{H} \times \mathcal{H}$ given by $S(x, y) = (y, x)$. One has then $\operatorname{gr} T^{-1} = S(\operatorname{gr} T)$. We note that S commutes with J and with the operation of the orthogonal complement in $\mathcal{H} \times \mathcal{H}$:

$$\begin{aligned}\operatorname{gr}(T^{-1})^* &= J(\operatorname{gr} T^{-1})^\perp = J\big(S(\operatorname{gr} T)\big)^\perp = S\big(J(\operatorname{gr} T)^\perp\big)\\ &= S(\operatorname{gr} T^*) = S(\operatorname{gr} T) = \operatorname{gr} T^{-1},\end{aligned}$$

so $(T^{-1})^* = T^{-1}$. □

1.3 Some function spaces

Let $\Omega \subset \mathbb{R}^d$ be a non-empty open set.

If $f : \Omega \to \mathbb{C}$ is a continuous function, we denote

$$\operatorname{supp} f := \text{the closure of the set } \{x \in \Omega : f(x) \neq 0\}.$$

We further denote

$$C_c^\infty(\Omega) := \{f \in C^\infty(\Omega) : \operatorname{supp} f \text{ is a compact subset of } \Omega\}.$$

The functions $C_c^\infty(\Omega)$ are "very good" in all aspects considered in the mathematical analysis: locally (i.e. infinitely differentiable at any point) and globally (identically zero outside a compact subset: no troubles at the boundary/at infinity). The elements of $C_c^\infty(\Omega)$ are often called *test functions* on Ω. We are going to show that $C_c^\infty(\Omega)$ are dense in other functional spaces. We remark first that if $\Omega \subset \Omega'$, then any function $\varphi \in C_c^\infty(\Omega)$ can be viewed as a function in $C_c^\infty(\Omega')$ by taking the extension by zero. If particular, any test function on Ω is viewed as a test function on $\mathbb{R}^d$.

Let $d \in \mathbb{N}$, $\alpha \in \mathbb{N}_0^d$ be a multi-index, $\alpha = (\alpha_1, \dots, \alpha_d)$ with $\alpha_i \in \mathbb{N}$, $x \in \mathbb{R}^d$. We will use the writing

$$\partial^\alpha := \partial_1^{\alpha_1} \dots \partial_d^{\alpha_d}, \quad |\alpha| := \alpha_1 + \dots + \alpha_d, \quad x^\alpha := x_1^{\alpha_1} \cdot \dots \cdot x_d^{\alpha_d},$$

where $\partial_i^{\alpha_i}$ means the partial derivative with respect to the ith variable applied α_i times (in particular, the operation does nothing if is $\alpha_i = 0$). Remark that if one applies ∂^α to a C^∞ function, then the result is independent of the order in which the partial derivatives are taken.

First remark that there exist non-trivial test functions (=test functions which are not identically zero). Namely, one checks routinely that

$$\rho : x \mapsto c \begin{cases} \exp\Big(-\dfrac{1}{1-|x|^2}\Big), & |x| < 1, \\ 0, & |x| \geq 1. \end{cases}$$

belongs to $C_c^\infty(\mathbb{R}^d)$, non-negative, with $\operatorname{supp}\rho \subset \overline{B}_1(0)$, and we choose $c > 0$ in such a way that $\|\rho\|_{L^1(\mathbb{R}^d)} = 1$. For $\delta > 0$ define

$$\rho_\delta : x \mapsto \frac{1}{\delta^d}\rho\big(\frac{x}{\delta}\big),$$

then

$$\rho_\delta \geq 0, \quad \operatorname{supp}\rho_\delta \subset \overline{B}_\delta(0), \quad \int_{\mathbb{R}^d} \rho(x)\,\mathrm{d}x = 1.$$

We now briefly review some properties of the convolution. The convolution $f * g$ of two measurable functions $f, g : \mathbb{R}^d \to \mathbb{C}$ is defined by

$$(f * g)(x) = \int_{\mathbb{R}^d} f(y) g(x-y)\,\mathrm{d}y,$$

if the integral on the right-hand side exists for a.e. $x \in \mathbb{R}^d$. This notion and some of its properties were introduced in Analysis III.[1] In particular, it was shown that

$$\|f * g\|_{L^1} \le \|f\|_{L^1}\|g\|_{L^1} \text{ for any } f, g \in L^1(\mathbb{R}^d),$$
$$\|f * \rho_\delta - f\|_{L^1} \xrightarrow{\delta \to 0^+} 0 \text{ for any } f \in L^1(\mathbb{R}^d).$$

In fact, the same proofs can be easily adapted (by using the Hölder inequality at some points) to show that for any $p \in [1, \infty)$ one has

$$\|f * g\|_{L^p} \le \|f\|_{L^p}\|g\|_{L^1(\mathbb{R}^d)} \text{ for any } f \in L^p(\mathbb{R}^d), \quad g \in L^1(\mathbb{R}^d) \tag{1.2}$$
$$\|f * \rho_\delta - f\|_{L^p} \xrightarrow{\delta \to 0^+} 0 \text{ for any } f \in L^p(\mathbb{R}^d). \tag{1.3}$$

These properties allow us to show the following important assertion:

Theorem 1.22. *For any $p \in [1, \infty)$ the set $C_c^\infty(\Omega)$ is dense in $L^p(\Omega)$.*

Proof. One may approximate Ω by compact subsets, i.e. there exists a sequence $(K_j)_{j\in\mathbb{N}}$ of compact $K_j \subset \Omega$, with $K_j \subset K_{j+1}$ for all j, such that $\Omega = \bigcup_{j\in\mathbb{N}} K_j$. For example, one can take

$$K_j := \overline{B}_j(0) \cap \Big\{x \in \Omega : d_\Omega(x) \le \frac{1}{j}\Big\}, \quad d_\Omega(x) := \inf_{y\in\Omega^\complement} |x - y|$$

(remark that d_Ω is a continuous function). Denote $f_j := 1_{K_j} f$, then $f_j \to f$ in $L^p(\Omega)$ by the dominated convergence:

$$\begin{aligned}\lim_j \|f_j - f\|_{L^p}^p = \lim_j \int_\Omega 1_{\Omega\setminus K_j}(x)|f(x)|^p\,\mathrm{d}x = \lim_j \int_\Omega 1_{\Omega\setminus\cup_{i\le j}K_i}(x)|f(x)|^p\,\mathrm{d}x\\ = \int_\Omega \lim_j 1_{\Omega\setminus\cup_{i\le j}K_i}(x)|f(x)|^p\,\mathrm{d}x = \int_\Omega 0\,\mathrm{d}x = 0.\end{aligned}$$

Let $\varepsilon > 0$, then one can choose some $j \in \mathbb{N}$ with $\|f_j - f\|_{L^p} < \varepsilon$. As K_j is compact, one has

$$\delta_j := \inf_{x\in K_j} d_\Omega(x) > 0.$$

We extend f_j by zero to the whole of $\mathbb{R}^d$ and consider it as a function in $L^p(\mathbb{R}^d)$. Consider $g_\delta := f_j * \rho_\delta$ with $\delta > 0$. From the definition of $*$ it is clear that g_δ is C^∞ with $\partial^\alpha = f_j * \partial^\alpha\rho_\delta$ for any $\alpha \in \mathbb{N}_0^d$. In addition, by (1.3) one can choose δ sufficiently small to have $\|g_\delta - f_j\|_{L^p(\mathbb{R}^d)} < \varepsilon$. Without loss of generality we may assume that $\delta < \delta_j$. As f_j is zero outside K_j and $\operatorname{supp}\rho_\delta \subset B_\delta(0)$ it follows from the definition of the convolution that $g_\delta(x) = 0$ for all $x \in \mathbb{R}^d$ such that $|x - y| > \delta$ for all $y \in K_j$, i.e.

$$\operatorname{supp} g_\delta \subset \{y + z : \; y \in K_j, \; z \in \overline{B}_\delta(0)\},$$

[1] My lecture notes “Analysis III” can be downloaded from `https://uol.de/pankrashkin/lehre-teaching`, the convolution is discussed in the last chapter.

and the set on the right-hand is a compact subset of Ω (as δ is strictly small than the distance between K_j and Ω^{C}). Therefore, $g_\delta \in C_c^\infty(\Omega)$, and by construction we have

$$\begin{aligned}\|f - g_\delta\|_{L^p(\Omega)} &\le \|f - f_j\|_{L^p(\Omega)} + \|f_j - g_\delta\|_{L^p(\Omega)}\\ &= \|f - f_j\|_{L^p(\Omega)} + \|f_j - g_\delta\|_{L^p(\Omega)} < \varepsilon + \varepsilon = 2\varepsilon.\end{aligned}$$

As $\varepsilon > 0$ can be arbitrary, the result follows. □

Remark 1.23 (Local versions of function spaces). The test functions are also used to construct "local" version of various functional spaces. If with every open $\Omega \subset \mathbb{R}^d$ one associates in some canonical sense a space $\mathcal{F}(\Omega)$ consisting of the functions $f : \Omega \to \mathbb{C}$ a family of properties (for example, $\mathcal{F} = L^p$, further example will be introduced soon), then one defines

$$\mathcal{F}_{\text{loc}}(\Omega) := \{f : \Omega \to \mathbb{C} : \varphi f \in \mathcal{F}(\mathbb{R}^d) \text{ for any } \varphi \in C_c^\infty(\Omega)\}.$$

For example, according to this definition,

$$L^p_{\text{loc}}(\Omega) := \{f : \Omega \to \mathbb{C} : \varphi f \in L^p(\mathbb{R}^d) \text{ for any } \varphi \in C_c^\infty(\Omega)\}.$$

But for this specific case we are going to prove another characterization.

Lemma 1.24. *Let $\Omega \subset \mathbb{R}^d$ be open and $K \subset \Omega$ compact, then there exists $\varphi \in C_c^\infty(\Omega)$ with $\varphi = 1$ on K.*

Proof. Consider again the continuous function $d_\Omega : x \mapsto \inf_{y\in\Omega^{\mathsf{C}}} |x - y|$, then $\delta_K := \inf_{x\in K} d_\Omega(x) > 0$. Choose $\delta \in \big(0, \frac{\delta_0}{2}\big)$ and set

$$\varphi : x \mapsto \int_{\operatorname{dist}(y,K)<\delta} \rho_\delta(x - y)\,\mathrm{d}y,$$

then $\varphi \in C^\infty(\mathbb{R}^d)$ with $\operatorname{supp}\varphi \in \{x : \operatorname{dist}(x, K) \le 2\delta\}$ =compact subset of Ω. If $x \in K$, then all y with $\rho_\delta(x - y) \ne 0$ are contained in $\overline{B}_\delta(x) \subset \big\{y : \operatorname{dist}(y, K) \le \delta\big\}$, so

$$\varphi(x) = \int_{\operatorname{dist}(y,K)<\delta} \rho_\delta(x - y)\,\mathrm{d}y = \int_{\mathbb{R}^d} \rho_\delta(x - y)\,\mathrm{d}y = (y = t) = \int_{\mathbb{R}^d} \rho_\delta(t)\,\mathrm{d}t = 1. \quad \square$$

Proposition 1.25. *For any $p \in [1,\infty]$ there holds*

$$L^p_{\text{loc}}(\Omega) = \Big\{f : \Omega \to \mathbb{C} \textit{ measurable} : \int_K |f|^p\,\mathrm{d}x < \infty \textit{ for any compact } K \subset \Omega\Big\}. \tag{1.4}$$

Proof. We consider $p < \infty$ only ($p = \infty$ is an easy exercise). Let L be the set on the right-hand side of (1.4).

Assume that $f \in L$ and $\varphi \in C_c^\infty(\Omega)$. Let $K := \operatorname{supp}\varphi$, then $|\varphi| \le \|\varphi\|_\infty 1_K$ and

$$\int_{\mathbb{R}^d} |\varphi f|^p\,\mathrm{d}x \le \|\varphi\|_\infty^2 \int_{\mathbb{R}^d} |1_K f|^p\,\mathrm{d}x = \|\varphi\|_\infty^2 \int_K |f|^p\,\mathrm{d}x < \infty,$$

which shows $\varphi f \in L^p(\mathbb{R}^d)$. Hence, $L \subset L^p_{\rm loc}(\Omega)$.

Let $f \in L^p_{\rm loc}(\Omega)$ and $K \subset \Omega$ compact. Be Lemma 1.24 there exists $\varphi \in C_c^\infty(\Omega)$ with $\varphi = 1$ on K. As $\varphi f \in L^p(\mathbb{R}^d)$, the function φf is measurable on any subset of $\mathbb{R}^d$, in particular, on K. On K one has $\varphi f = f$, hence, f is measurable on K. In particular, f is measurable on any ball in Ω and then on the whole of Ω. Further we have

$$\int_K |f|^p \,\mathrm{d}x = \int_K |\varphi f|^p \,\mathrm{d}x \le \int_{\mathbb{R}^d} |\varphi f|^p \,\mathrm{d}x < \infty.$$

As K is arbitrary, it shows the inclusion $L^p_{\rm loc}(\Omega) \subset L$. □

Remark 1.26. It is clear that

$$L^p(\Omega) \subset L^p_{\rm loc}(\Omega) \text{ for any } p \ge 1.$$

Furthermore, due to the Hölder inequality, for any compact K we have

$$\int_K |f| \,\mathrm{d}x = \int_K |1_K f| \,\mathrm{d}x \le \Big(\int_K |1_K|^q\Big)^{1/q}\Big(\int_K |f|^p\Big)^{1/p} \equiv |K|^{1/q}\Big(\int_K |f|^p\Big)^{1/p}.$$

According to (1.4) it shows the inclusions

$$L^p_{\rm loc}(\Omega) \subset L^1_{\rm loc}(\Omega) \text{ for any } p \ge 1.$$

Further remark that

$$C^0(\Omega) \subset L^p_{\rm loc}(\Omega) \text{ for any } p \ge 1.$$

We conclude this section by showing the following important result:

Proposition 1.27. *If $f \in L^1_{\rm loc}(\Omega)$ such that*

$$\int_\Omega f(x)\varphi(x) \,\mathrm{d}x = 0 \textit{ for all } \varphi \in C_c^\infty(\Omega),$$

then $f = 0$ a.e. in Ω.

Proof. Let B be a ball with $\overline{B} \subset \Omega$. By Lemma 1.24 there exists $\psi \in C_c^\infty(\Omega)$ with $\psi = 1$ on $\overline{B}$. Due to $f \in L^1_{\rm loc}(\Omega)$ we have $\psi f \in L^1(\mathbb{R}^d)$, and then $\|(\psi f) * \rho_\delta - \psi f\|_{L^1(\mathbb{R}^d)} \to 0$ as $\delta \to 0^+$. On other hand,

$$(\psi f) * \rho_\delta(x) = \int_{\mathbb{R}^d} \psi(y) f(y) \rho_\delta(x-y) \,\mathrm{d}y = \int_\Omega f(y)\varphi_x(y) \,\mathrm{d}y \text{ for } \varphi_x(y) := \psi(y)\rho_\delta(x-y).$$

As $\varphi_x \in C_c^\infty(\Omega)$, the term on the right-hand side is zero by assumption, hence, $(\psi f) * \rho_\delta \equiv 0$, and then $\psi f = 0$ a.e. As $\psi = 1$ in B, one has $f = 0$ a.e. in B. Therefore, $f = 0$ a.e. in any ball in Ω and then in the whole of Ω. □

1.4 Weak derivatives

Let $\Omega \subset \mathbb{R}^d$ be a non-empty open set and $\mathcal{H} := L^2(\Omega)$. Let $m \in \mathbb{N}$ and $c_\alpha : \Omega \to \mathbb{C}$ be measurable functions, $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq m$. In the theory of differential operators one deals with realizations of differential expressions

$$P := \sum_{|\alpha| \leq m} c_\alpha \partial^\alpha \tag{1.5}$$

as linear operators with "good" properties in the Hilbert space $\mathcal{H}$, i.e. one looks for suitable $D(T) \subset L^2(\Omega)$ such that the linear operator

$$T : u \mapsto Pu \text{ with domain } D(T)$$

becomes closed/symmetric/self-adjoint etc. One should immediately say that the problem is very difficult and no general solutions exists so far. In this course, we mostly deal some particular P with smooth or even constant coefficients c_α (anyway, we remark that some non-constant and non-smooth coefficients will appear when we will deal with Schrödinger operators).

Definition 1.28 (Weak derivative). Let $f \in L^1_{\text{loc}}(\Omega)$ and $\alpha \in \mathbb{N}_0^d$. One says that a function $g \in L^1_{\text{loc}}(\Omega)$ is the *weak ∂^α-derivative* of f in Ω, if for all $\varphi \in C_c^\infty(\Omega)$ one has the equality

$$\int_\Omega f \partial^\alpha \varphi = (-1)^{|\alpha|} \int_\Omega g\varphi. \tag{1.6}$$

If such g exists, then it is unique by Proposition 1.27. One says that f *admit a weak ∂^α-derivative*, and for the moment we use the writing

$$g = \widetilde{\partial^\alpha} f$$

in order to distinguish from the usual derivatives. If f admits weak ∂^α-derivatives for all $|\alpha| \leq m$, then one says that f is *m times weakly differentiable.*

Proposition 1.29. *If $f \in C^m(\Omega)$ with some $m \in \mathbb{N}$, then weak derivatives up to order m exist and coincide with the usual derivatives.*

Proof. Let $\varphi \in C_c^\infty(\Omega)$, then $\varphi f \in C^m(\mathbb{R}^d)$ and for sufficiently large $R > 0$ there holds $\operatorname{supp}(\varphi f) \subset (-R, R)^d$. Furthermore,

$$\int_{[-R,R]^d} \partial_1(\varphi f)\,\mathrm{d}x = \int_{-R}^{R} \cdots \int_{-R}^{R} \int_{-R}^{R} \partial_1\big(\varphi f\big)(x_1, \ldots, x_d)\,\mathrm{d}x_1\,\mathrm{d}x_2 \ldots \mathrm{d}x_d$$
$$\int_{-R}^{R} \cdots \int_{-R}^{R} \Big[\underbrace{(\varphi f)(R, x_2, \ldots, x_d)}_{=0} - \underbrace{(\varphi f)(-R, x_2, \ldots, x_d)}_{=0} \Big]\,\mathrm{d}x_2 \ldots \mathrm{d}x_d = 0.$$

On the other hand, using the Leibniz rule one obtains

$$\int_{[-R,R]^d} \partial_1(\varphi f)\,\mathrm{d}x = \int_\Omega \partial_1(\varphi f)\,\mathrm{d}x = \int_\Omega (f\,\partial_1\varphi + \varphi\,\partial_1 f)\,\mathrm{d}x,$$

which gives

$$\int_\Omega f\partial_1\varphi\,\mathrm{d}x = -\int_\Omega \varphi\partial_1 f\,\mathrm{d}x \text{ for all } \varphi\in C_c^\infty(\Omega),$$

hence $\partial_1 f = \widetilde{\partial}_1 f$, and one extends this argument to general α by re-enumeration and iteration. □

Example 1.30. There exist weakly differentiable functions that are not classically differentiable. For example, consider $\Omega := \mathbb{R}$ and $f(x) = |x|$, which is not differentiable at 0. For any $\varphi \in C_c^\infty(\Omega)$ one has

$$\begin{aligned}
\int_\mathbb{R} f(x)\varphi'(x)\,\mathrm{d}x &= -\int_{-\infty}^0 x\varphi'(x)\,\mathrm{d}x + \int_0^\infty x\varphi'(x)\,\mathrm{d}x\\
&= -x\varphi(x)\Big|_{-\infty}^0 + \int_{-\infty}^0 \varphi(x)\,\mathrm{d}x + x\varphi(x)\Big|_0^\infty - \int_0^\infty \varphi(x)\,\mathrm{d}x\\
&= \int_{-\infty}^0 \varphi(x)\,\mathrm{d}x - \int_0^\infty \varphi(x)\,\mathrm{d}x\\
&= -\int_\mathbb{R} \operatorname{sgn}(x)\,\varphi(x), \quad \operatorname{sgn}(x) = \begin{cases} 1, & x>0,\\ 0, & x=0,\\ -1, & x<0,\end{cases}
\end{aligned}$$

(the precise value of sgn in 0 has no importance here), which shows that sgn is the weak derivative of f.

On the other hand, not every function is weakly differentiable. For example, the function sgn is not weakly differentiable. To see this, remark first that for any $\varphi \in C_c^\infty(\Omega)$ one has

$$\begin{aligned}
\int_\mathbb{R} \operatorname{sgn}(x)\varphi'(x)\,\mathrm{d}x &= -\int_{-\infty}^0 \varphi'(x)\,\mathrm{d}x + \int_0^\infty \varphi'(x)\,\mathrm{d}x\\
&= -\varphi(x)\Big|_{-\infty}^0 + \varphi(x)\Big|_0^\infty = -2\varphi(0).
\end{aligned}$$

A weak derivative g of sgn would satisfy then

$$\int_\mathbb{R} g(x)\varphi(x)\,\mathrm{d}x = 2\varphi(0) \text{ for all } \varphi\in C_c^\infty(\mathbb{R}). \tag{1.7}$$

In particular, one would get

$$\int_0^\infty g(x)\varphi(x)\,\mathrm{d}x = 0 \text{ for all } \varphi \in C_c^\infty(0,\infty),$$

and then $g = 0$ a.e. in $(0,\infty)$ by Proposition 1.27. Analogolously $g = 0$ a.e. in $(-\infty, 0)$, so finally $g = 0$ a.e. in $\mathbb{R}$. If one now takes $\varphi \in C_c^\infty(\mathbb{R})$ with $\varphi(0) = 1$, one obtains a contradiction with (1.7).

Example 1.31. The operation $\widetilde{\partial^\alpha}$ is also called *distributional derivative*. Distributions represent useful generalizations of functions (every function $f \in L^1_{\text{loc}}$ is a distribution), and any distribution is infinitely differentiable. In particular, all derivatives of any function $f \in L^1_{\text{loc}}$ exist, but the price to pay is that they are not functions anymore: for example, $(\text{sgn})' = 2\delta$, where δ is the so-called Dirac distribution. Nevertheless, if some derivatives happen to be functions, then they are exactly the weak derivatives defined above.

Remark 1.32. The construction can be generalized to more general differential expressions. Let P be as in (1.5) with *smooth* coefficients $c_\alpha \in C^\infty(\Omega)$ and $f, g \in L^1_{\text{loc}}(\Omega)$, then one says that $g = Pf$ *weakly* (or writes $g = \widetilde{P}u$) if

$$\int_\Omega g\varphi = \sum_{|\alpha|\le m} (-1)^{|\alpha|} \int_\Omega f \partial^\alpha (c_\alpha \varphi) \text{ for any } \varphi \in C_c^\infty(\Omega).$$

Important: this does not imply the existence all weak derivatives appearing in the expression of Pf, as one can have a kind of "compensation" of various "bad" terms. If $f \in C^m(\Omega)$, then one shows as in Remark 1.29 that that the usual and weak version Pf coincide.

We emphasize some important properties of weak derivatives:

Proposition 1.33 (Leibniz rule). *The weak derivatives satisfy the Leibniz rule: if $f \in L^1_{\text{loc}}(\Omega)$ has the weak derivative $\widetilde{\partial}_j f \in L^1_{\text{loc}}(\Omega)$ and $\chi \in C^\infty(\Omega)$, then also χf admits the weak ∂_j derivative, and $\widetilde{\partial}_j(\chi f) = f\, \partial_j \chi + \chi \widetilde{\partial}_j f$.*

Proof. For any $\varphi \in C_c^\infty(\Omega)$ one has $\chi\varphi \in C_c^\infty(\mathbb{R}^d)$, hence,

$$\int_\Omega \widetilde{\partial}_j f \cdot \chi\varphi = -\int_\Omega f \partial_1(\chi\varphi) = -\int_\Omega f \partial_1 \chi \cdot \varphi - \int_\Omega f\chi \cdot \partial_1 \varphi.$$

This can be rewritten as

$$\int_\Omega (\chi\, \widetilde{\partial}_1 f + f\, \partial_1 \chi)\varphi = -\int_\Omega \chi f \cdot \partial_1 \varphi, \quad \varphi \in C_c^\infty(\Omega),$$

and $\widetilde{\partial}_1(\chi f) = f\partial_1\chi + \chi\widetilde{\partial}_1 f$ by Proposition 1.27, as the function on the right-hand side is in $L^1_{\text{loc}}(\Omega)$. □

Proposition 1.34 (Weak derivatives and convolutions). *If $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ has the weak derivative $\widetilde{\partial}_j f \in L^1_{\text{loc}}(\mathbb{R}^d)$, then*

$$\partial_j (f * \varphi) = (\widetilde{\partial}_j f) * \varphi \text{ for any } \varphi \in C_c^\infty(\mathbb{R}^d).$$

Remark that $f * \varphi \in C^\infty$, therefore, one takes the usual derivative on the left-hand side.

Proof. One easily checks that one is allowed to interchange the derivative with the integral, so for any $x \in \mathbb{R}^d$ there holds

$$\begin{aligned}\partial_j(f * \varphi)(x) &= \partial_j \int_{\mathbb{R}^d} f(y)\varphi(x-y)\,\mathrm{d}x = \int_{\mathbb{R}^d} f(y)(\partial_j\varphi)(x-y)\,\mathrm{d}x\\ &= -\int_{\mathbb{R}^d} f(y)\partial_j\varphi_x(y)\,\mathrm{d}x \text{ for } \varphi_x(y) = \varphi(x-y).\end{aligned}$$

As $\varphi_x \in C_c^\infty(\mathbb{R}^d)$, one continues as

$$\begin{aligned}&= \int_{\mathbb{R}^d} \widetilde{\partial}_j f(y)\cdot\varphi_x(y)\,\mathrm{d}x = \int_{\mathbb{R}^d} \widetilde{\partial}_j f(y)\cdot\varphi(x-y)\,\mathrm{d}x\\ &= \big((\widetilde{\partial}_j f) * \varphi\big)(x).\end{aligned}$$

□

From now on we denote $\widetilde{\partial^\alpha}$ and ∂^α by the same symbol ∂^α. (If it will be important, we will say explicitly which derivative type is used.)

Now let us return back to the differential expression P with *smooth* coefficients c_α as in (1.5) and the following densely defined linear operator in $\mathcal{H}$:

$$Tu = Pu, \quad D(T) = C_c^\infty(\Omega).$$

Using the definition of the adjoint operator one sees that

$$T^*u = P'u \text{ (weakly)}, \quad D(T^*) = \big\{u \in L^2(\Omega) : P'u \in L^2(\Omega)\big\},$$

where P' is the so-called *formal adjoint* of P, i.e.

$$P' : u \mapsto \sum_{|\alpha|\le m} (-1)^{|\alpha|}\partial^\alpha(\overline{c_\alpha}u).$$

The formal adjoint has the property that $\langle\varphi, P\psi\rangle = \langle P'\varphi, \psi\rangle$ for any $\varphi, \psi \in C_c^\infty(\Omega)$, and the differential expression P will be called *formally self-adjoint* if $P = P'$, i.e. $c_\alpha = (-1)^{|\alpha|}\overline{c_\alpha}$ for all α. **For the rest of the section we assume that P is formally self-adjoint and that all coefficients c_α are constant.** Important examples are

$$P = -i\partial_j, \quad P = -\Delta \equiv -\sum_{j=1}^d \partial_j^2 \text{ (Laplacian)} .$$

Then one easily sees that $T \subset T^*$, i.e. that T is symmetric and, hence, closable. The closure of T is usually called the *minimal operator* generated by the differential expression P and is denoted $P_{\min}$. The operator T^* is called the *maximal operator* generated by the differential expression P and is denoted by $P_{\max}$.

It is natural to ask if one has $P_{\min} = P_{\max}$: if the equality holds, then $\overline{T} = T^*$, hence, T is essentially self-adjoint, while $T^* = P_{\max}$ is self-adjoint. If the equality fails, then $\overline{T}$ is just symmetric, but is not self-adjoint. Checking $P_{\min} = P_{\max}$ is a difficult question as, in general, it depends on the geometry of Ω or, more precisely of the regularity properties of its boundary. It is not our objective to study the most general case (in fact, this is one of the hardest parts of analysis), but we are going to look at some important examples.

1.5 Sobolev spaces

We continue to work with a non-empty open set $\Omega \subset \mathbb{R}^d$.

Definition 1.35 (Sobolev space). For $k \in \mathbb{N}$ the kth *Sobolev space* $H^k(\Omega)$ on Ω is defined as

$$H^k(\Omega) = \left\{ \begin{array}{r} f \in L^2(\Omega) : f \text{ is } k \text{ times weakly differentiable,} \qquad\qquad\qquad \\ \text{with } \partial^\alpha f \in L^2(\Omega) \text{ for all } \alpha \in \mathbb{N}_0^d \text{ with } |\alpha| \le k \end{array} \right\}.$$

We introduce a scalar product in $H^k(\Omega)$ by

$$\langle f, g\rangle_{H^k(\Omega)} = \sum_{|\alpha|\le k} \langle \partial^\alpha f, \partial^\alpha g\rangle_{L^2(\Omega)},$$

and in order to have a uniform notation we set $H^0(\Omega) := L^2(\Omega)$.

Theorem 1.36. *The Sobolev space $H^k(\Omega)$ with the above scalar product is a Hilbert space.*

Proof. Let $(v_j) \in H^k(\Omega)$ be a Cauchy sequence with respect to $\|\cdot\|_{H^k}$. For each α with $|\alpha| \le k$ one has $\|\partial^\alpha u\|_{L^2} \le \|u\|_{H^k}$, and it follows that $(\partial^\alpha v_j)$ is a Cauchy sequence in $L^2(\Omega)$. As $L^2(\Omega)$ is complete, there exist $v^\alpha := L^2 - \lim \partial^\alpha v_j$. Denote $v := v^0$, then for any $\varphi \in C_c^\infty(\Omega)$ one has

$$\begin{aligned} \langle v^\alpha, \varphi\rangle_{L^2} &= \lim \langle \partial^\alpha v_j, \varphi\rangle_{L^2} = \lim \int_\Omega \overline{\partial^\alpha v_j} \cdot \varphi \, \mathrm{d}x \\ &= \lim (-1)^{|\alpha|} \int_\Omega \overline{v_j} \partial^\alpha \varphi \, \mathrm{d}x = (-1)^{|\alpha|} \lim \langle v_j, \partial^\alpha \varphi\rangle_{L^2} = (-1)^{|\alpha|} \langle v, \partial^\alpha \varphi\rangle, \end{aligned}$$

which means that v^α is the weak ∂^α-derivative of v. This shows that $v \in H^k(\Omega)$. □

It is a remarkable fact that in the absence of boundaries (i.e. for $\Omega = \mathbb{R}^d$) there is an alternative description of the Sobolev spaces. Namely, the Sobolev spaces $H^k(\mathbb{R}^d)$ can be characterized using the Fourier transform, which we will briefly address now.

Recall[2] that the Fourier transform $\widehat{f}$ of a function $f \in L^1(\mathbb{R}^d)$ is given by

$$\widehat{f}(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(x) e^{-i\xi\cdot x} \, dx, \quad x \in \mathbb{R}^d, \tag{1.8}$$

and it is a continuous function satisfying

$$\|\widehat{f}\|_\infty \le \frac{1}{(2\pi)^{d/2}} \|f\|_{L^1}. \tag{1.9}$$

If $f \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d) \cap C^0(\mathbb{R}^d)$ such that $\widehat{f} \in L^1$, then the *Fourier inversion formula* holds

$$f(x) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \widehat{f}(\xi) e^{i\xi\cdot x} \, d\xi, \quad x \in \mathbb{R}^d. \tag{1.10}$$

[2] The Fourier transform was defined in Analysis III without the coefficient in front of the integral, but for our purposes is will be useful.

Remark that for any $\varphi \in C_c^\infty(\Omega)$ and any $\alpha \in \mathbb{N}_0^d$ obtains (with the help of the partial integration)

$$\begin{aligned}\widehat{\partial^\alpha \varphi}(\xi) &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \partial^\alpha \varphi(x) \cdot e^{-i\xi\cdot x}\, dx = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \varphi(x)(-\partial_x)^\alpha e^{-i\xi\cdot x}\, dx \\ &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \varphi(x)(i\xi)^\alpha e^{-i\xi\cdot x}\, dx = (i\xi)^\alpha \widehat{\varphi}(\xi).\end{aligned}$$

and by (1.9) all $\xi^\alpha \varphi$ are bounded. It follows that $\widehat{\varphi}$ decays rapidly at infinity (and converges to zero faster that any rational function), in particular, $\widehat{\varphi} \in L^1 \cap L^\infty$. This means that the above inversion formula holds for all test functions. We further remark that for any test function φ one has $\widehat{\varphi} \in L^1 \cap L^\infty \subset L^2$.

Theorem 1.37 (Fourier transform as a unitary map). *The linear map* $C_c^\infty(\mathbb{R}^d) \ni f \mapsto \widehat{f} \in L^2(\mathbb{R}^d)$ *extends uniquely to a unitary operator*

$$\mathcal{F} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d), \text{ which satisfies } \mathcal{F}\mathcal{F}f = f(-\cdot). \tag{1.11}$$

Proof. For $\varphi, \psi \in C_c^\infty(\mathbb{R}^d)$ one has, using the inversion formula,

$$\langle \varphi, \psi \rangle_{L^2} = \int_{\mathbb{R}^d} \overline{\varphi(x)} \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{i\xi\cdot x} \widehat{\psi}(\xi)\, \mathrm{d}\xi\, \mathrm{d}x$$

one can interchange the integrals, as $(x,\xi) \mapsto \overline{\varphi(x)}\widehat{\psi}(\xi)$ in in $L^1(\mathbb{R}^d \times \mathbb{R}^d)$

$$\begin{aligned}&= \int_{\mathbb{R}^d} \overline{\frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-i\xi\cdot x} \varphi(x)\, \mathrm{d}x}\, \widehat{\psi}(\xi)\, \mathrm{d}\xi \\ &= \int_{\mathbb{R}^d} \overline{\widehat{\varphi}(\xi)} \widehat{\psi}(\xi)\, \mathrm{d}\xi = \langle \widehat{\varphi}, \widehat{\psi} \rangle_{L^2}.\end{aligned}$$

For $\varphi = \psi$ one obtains $\|\varphi\|_{L^2} = \|\widehat{\varphi}\|_{L^2}$. As $C_c^\infty(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d)$, the map $\varphi \mapsto \widehat{\varphi}$ uniquely extends to a linear operator $\mathcal{F}$ in $L^2(\mathbb{R}^d)$ with $\|\mathcal{F}f\|_{L^2} = \|f\|_{L^2}$ for all $f \in L^2(\mathbb{R}^d)$, and then $\operatorname{ran} \mathcal{F}$ is automatically closed. It remains to check that $\operatorname{ran} \mathcal{F} = L^2(\mathbb{R}^d)$. Assume that $f \in L^2(\mathbb{R}^d)$ with $f \perp \operatorname{ran} \mathcal{F}$. Let $f_j \in C_c^\infty(\mathbb{R}^d)$ with $\lim \|f_j - f\|_{L^2} = 0$, then $\mathcal{F}f = L^2 - \lim \widehat{f_j}$. For any $\varphi \in C_c^\infty(\mathbb{R}^d)$ we have

$$\begin{aligned}0 &= \langle f, \widehat{\varphi} \rangle_{L^2} = \lim \langle f_j, \widehat{\varphi} \rangle_{L^2} \\ &= \lim_j \int_{\mathbb{R}^d} \overline{f_j(\xi)} \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-i\xi\cdot x}\, \varphi(x)\, \mathrm{d}x\, \mathrm{d}\xi \\ &= \lim_j \int_{\mathbb{R}^d} \overline{\frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f_j(\xi) e^{i\xi\cdot x}\, \mathrm{d}\xi}\, \varphi(x)\, \mathrm{d}x \\ &= \lim_j \int_{\mathbb{R}^d} \overline{\widehat{f_j}(-x)} \varphi(x)\, \mathrm{d}x = \lim_j \langle \widehat{f_j}(-\cdot), \varphi \rangle_{L^2} = \langle \mathcal{F}f(-\cdot), \varphi \rangle_{L^2},\end{aligned}$$

which shows that $\mathcal{F}f(-\cdot) = 0$, then $\mathcal{F}f = 0$ and $f = 0$.

For the identity in (1.11) one simply extends the inversion formula (1.8) be density. $\square$

Remark 1.38. It is common to keep the notation $\widehat{f}$ instead of $\mathcal{F}f$ for all $f \in L^2(\mathbb{R}^d)$.

Recall that we deal with the differential expression

$$P := \sum_{|\alpha|\le m} c_\alpha \partial^\alpha, \quad c_\alpha \in \mathbb{C},$$

which is assumed formally self-adjoint.

Proposition 1.39 (Weak derivatives and the Fourier transform). *For any $f \in L^2(\mathbb{R}^d)$, one can the equivalence*

$$Pf \in L^2(\mathbb{R}^d) \text{ (weakly)} \quad \Leftrightarrow \quad p\widehat{f} \in L^2(\mathbb{R}^d),$$

where

$$p(\xi) = \sum_{|\alpha|\le m} c_\alpha (i\xi)^\alpha.$$

Moreover, in this case one has $Pf = g$, where g is the unique L^2-function with $\widehat{g} = p(\xi)\widehat{f}$.

Proof. We remark that $p(\xi)$ is real-valued (as P is assumed formally self-adjoint).

The proof becomes quite technical at some points. Consider the space[3]

$$\mathcal{S}(\mathbb{R}^d) := \big\{ f \in C^\infty(\mathbb{R}^d) : x^\alpha \partial^\beta f \in L^\infty(\mathbb{R}^d) \text{ for all } \alpha, \beta \in \mathbb{N}_0^d \big\}.$$

Clearly, $C_c^\infty(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$, and using the inversion formula and Theorem 1.37 one checks that $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ is bijective.

$\Rightarrow$: Let $f \in L^2(\mathbb{R}^d)$ with $g := Pu \in L^2(\mathbb{R}^d)$, then for any $\varphi \in C_c^\infty(\mathbb{R}^d)$ one has

$$\int_{\mathbb{R}^d} \overline{P\varphi} f \,\mathrm{d}x = \int_{\mathbb{R}^d} \overline{\varphi} g \,\mathrm{d}x. \tag{1.12}$$

We claim that the identity also holds for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Namely, let $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Take $\chi \in \mathbb{R}^d$ such that $\chi(x) = 1$ for $|x| \le 1$ and $\chi(x) = 0$ for $|x| \ge 2$ and denote

$$\chi_N : x \mapsto \chi\big(\frac{x}{N}\big), \quad \varphi_N := \chi_N \varphi.$$

One has clearly $\varphi_N \in C_c^\infty(\mathbb{R}^d)$, hence,

$$\int_{\mathbb{R}^d} \overline{P\varphi_N} f \,\mathrm{d}x = \int_{\mathbb{R}^d} \overline{\varphi_N} g \,\mathrm{d}x, \tag{1.13}$$

and $\varphi_N \to \varphi$ in $L^2(\mathbb{R}^d)$ as $N \to \infty$. On the other hand,

$$\partial_j \varphi(x) = \frac{1}{N} (\partial_j \chi)\big(\frac{x}{N}\big)\varphi(x) + \chi_N(x) \partial_j \varphi(x),$$

[3] In the theory of distributions it is called Schwartz space.

and one easily shows that the first and the second summand converge respectively to 0 and $\chi\partial_j\varphi$ in $L^2(\mathbb{R}^d)$. Therefore, $\partial_j\varphi_N \xrightarrow{L^2} \partial_j\varphi$. Similarly, $\partial^\alpha\varphi_N \xrightarrow{L^2} \partial^\alpha\varphi$ for every α, and then by linearity one has $P\varphi_N \xrightarrow{L^2} P\varphi$. Hence, one passes to the limit $N \to \infty$ in (1.13) and obtains (1.12) for any $\varphi \in \mathcal{S}(\mathbb{R}^d)$.

Then (1.12) can be rewritten as $\langle P\varphi, f\rangle_{L^2} = \langle \varphi, g\rangle_{L^2}$ for any $\varphi \in \mathcal{S}(\mathbb{R}^d)$, and by Theorem 1.37 one has $\langle \widehat{P\varphi}, \widehat{f}\rangle = \langle \widehat{\varphi}, \widehat{g}\rangle$ and then

$$\int_{\mathbb{R}^d} p(\xi)\overline{\widehat{\varphi}(\xi)}\widehat{f}(\xi)d\xi = \int_{\mathbb{R}^d} \overline{\widehat{\varphi}(\xi)}\widehat{g}(\xi)d\xi.$$

As $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ is bijective, this can be rewritten as

$$\int_{\mathbb{R}^d} p(\xi)\widehat{f}(\xi)\psi(\xi)d\xi = \int_{\mathbb{R}^d} \widehat{g}(\xi)\psi(\xi)d\xi$$

for all $\psi \in \mathcal{S}(\mathbb{R}^d)$, in particular, for all $\psi \in C_c^\infty(\mathbb{R}^d)$. Due to $\widehat{f}, \widehat{g} \in L^2$ and $p \in C^0$ one has $pf, g \in L^1_{\text{loc}}(\mathbb{R}^d)$, and by Proposition 1.27 one has $pf = g$ a.e. and then $pf \in L^2(\mathbb{R}^d)$.

$\Leftarrow$: Now assume that $f \in L^2(\mathbb{R}^d)$ with $p\widehat{f} \in L^2(\mathbb{R}^d)$, then there exists a unique $g \in L^2(\mathbb{R}^d)$ with $\widehat{g} = p\widehat{f}$. Then for any $\varphi \in C_c^\infty(\mathbb{R}^d)$ one has

$$\int_{\mathbb{R}^d} \overline{\widehat{\varphi}(\xi)}p(\xi)\widehat{f}(\xi)d\xi = \int_{\mathbb{R}^d} \overline{\widehat{\varphi}(\xi)}\widehat{g}(\xi)d\xi,$$

which can be regrouped into $\langle \widehat{P\varphi}, \widehat{f}\rangle = \langle \widehat{\varphi}, \widehat{g}\rangle$ and then $\langle P\varphi, f\rangle = \langle \varphi, g\rangle$, which is exactly

$$\int_{\mathbb{R}^d} \overline{P\varphi}\, f \,\mathrm{d}x = \int_{\mathbb{R}^d} \overline{\varphi} g \text{ for any } \varphi \in C_c^\infty(\mathbb{R}^d)$$

and means that $g = Pf$. □

Corollary 1.40. *If P is formally self-adjoint, then the maximal operator $P_{\max}$ is self-adjoint.*

Proof. By Proposition 1.39 the operator $P_{\max}$ is unitarily equivalent to the self-adjoint multiplication operator M_p with a continuous real-valued p. (The unitary equivalence was defined in the exercises.) □

Corollary 1.41 (Characterizing Sobolev spaces using the Fourier transform). *For any $k \in \mathbb{N}$ one has*

$$H^k(\mathbb{R}^d) = \big\{ f \in L^2(\mathbb{R}^d) : \langle\xi\rangle^k \widehat{f} \in L^2(\mathbb{R}^d)\big\}, \quad \langle\xi\rangle := \sqrt{1 + |\xi|^2}. \tag{1.14}$$

Proof. Let L be the set on the right-hand side of (1.14). Let $f \in H^k(\mathbb{R}^d)$, then by Proposition 1.39 one has $\xi^\alpha \widehat{f} \in L^2(\mathbb{R}^d)$ for $|\alpha| \le k$. Using $\langle\xi\rangle \le 1 + |\xi_1| + \dots + |\xi_d|$ we estimate

$$\big|\,\langle\xi\rangle^k \widehat{f}\,\big| \le \big(1 + |\xi_1| + \dots + |\xi_d|\big)^k |\widehat{f}| \le \sum_{|\alpha|\le k} b_\alpha |\xi^\alpha \widehat{f}|,$$

where b_α are suitable constants. By Proposition 1.39 each summand on the right-hand side is an L^2-function, which shows that $\langle\xi\rangle^k\widehat{f} \in L^2(\mathbb{R}^d)$. This gives the inclusion $H^k(\mathbb{R}^d) \subset L$.

On the other hand, let $f \in L$. For $|\alpha| \le k$ one has $|\xi^\alpha| \le \langle\xi\rangle^{|\alpha|} \le \langle\xi\rangle^k$, therefore, $|\xi^\alpha\widehat{f}| \le \langle\xi\rangle^k|\widehat{f}| \in L^2(\mathbb{R}^d)$, implying $\xi^\alpha\widehat{f} \in L^2(\mathbb{R}^d)$. By Proposition 1.39 this means that $\partial^\alpha f \in L^2(\mathbb{R}^d)$. As this holds for arbitrary α with $|\alpha| \le k$, one arrives at the inclusion $L \subset H^k(\mathbb{R}^d)$. $\square$

Corollary 1.42 (Global elliptic regularity for the Laplacian). *Let $u \in L^2(\mathbb{R}^d)$ with $\Delta u \in H^k(\mathbb{R}^d)$, then $u \in H^{k+2}(\mathbb{R}^d)$.*

Proof. By assumption $\widehat{u} \in L^2$ and $\langle\xi\rangle^k|\xi^2|\widehat{u} \in L^2$. For $|\xi| \le 1$ one has $|\langle\xi\rangle^{k+1}\widehat{u}| \le |\widehat{u}|$. For $|\xi| \ge 1$ we have $\langle\xi\rangle^2 = 1 + |\xi^2| \le 2|\xi|^2$, therefore, $|\langle\xi\rangle^{k+2}\widehat{u}| = 2|\xi|^2|\langle\xi\rangle^k\widehat{u}|$. The above estimates can be summarized as $|\langle\xi\rangle^{k+2}\widehat{u}| \le 1_{|\xi|\le 1}|\widehat{u}| + 2\,1_{|\xi|\ge 1}|\xi|^2|\langle\xi\rangle^k\widehat{u}|$, and the two functions on the right-hand side are in $L^2(\mathbb{R}^d)$ due to the initial assumptions. This implies $\langle\xi\rangle^{k+2}\widehat{u} \in L^2(\mathbb{R}^d)$. $\square$

While the spaces H^k may look unusual at the beginning, they can be compared with the classical spaces C^m (as they represent an alternative way to measure the smoothness). He we will only provide the final results, and the proofs will be discussed in the exercises:

Proposition 1.43 (Density of test functions). *The set $C_c^\infty(\mathbb{R}^d)$ is dense in $H^k(\mathbb{R}^d)$ for any $k \in \mathbb{N}$.*

Theorem 1.44 (Sobolev embedding theorem). *Let $k \in \mathbb{N}$ and $m \in \mathbb{N}_0$ with $k > m + \frac{d}{2}$. Equip the vector space*

$$C^m_{L^\infty}(\mathbb{R}^d) := \big\{u \in C^\infty(\mathbb{R}^d) : \partial^\alpha u \in L^\infty(\mathbb{R}^d) \text{ for all } \alpha \in \mathbb{N}_0^d \text{ with } |\alpha| \le m\}$$

with the norm $\|u\|_{m,\infty} := \sum_{|\alpha|\le m} \|\partial^\alpha u\|_\infty$, then it becomes a Banach space, one has

$$H^k(\mathbb{R}^d) \subset C^m_{L^\infty}(\mathbb{R}^d),$$

and the embedding is continuous.

With the preceding notions and constructions, let us now discuss a very important example of Laplacians.

Example 1.45 (Laplacians in $\mathbb{R}^d$). Take $\mathcal{H} = L^2(\mathbb{R}^d)$ and consider several operators in $\mathcal{H}$ associated with the differential expression

$$P = -\Delta = -\sum_{j=1}^d \partial_j^2$$

called *the d-dimensional Laplacian.* Namely, define

$$T_0 = -\Delta u, \quad D(T_0) = C_c^\infty(\mathbb{R}^d),$$
$$T_1 = P_{\min}, \quad T_2 = P_{\max}.$$

Recall that by the preceding definitions and constructions the following holds: $T_1 = \overline{T_0}$ and $T_2 = T_0^*$, both T_1 and T_2 act as $u \mapsto -\Delta u$ (in the weak sense).

By Corollary 1.42 we have $D(T_2) = H^2(\mathbb{R}^d)$, and T_2 is self-adjoint by Corollary 1.40. For $u \in D(T_2)$ its graph norm is given by

$$\|u\|_{T_2}^2 = \|u\|_{L^2(\mathbb{R}^d)}^2 + \|\Delta u\|_{L^2(\mathbb{R}^d)}^2 \equiv \|\widehat{u}\|_{L^2(\mathbb{R}^d)}^2 + \||\xi|^2\widehat{u}\|_{L^2(\mathbb{R}^d)}^2 = \int_{\mathbb{R}^d} \big(1+|\xi|^4\big)\big|\widehat{u}(\xi)\big|^2 \,\mathrm{d}\xi,$$

while its H^2-norm is given by

$$\|u\|_{H^2}^2 = \sum_{|\alpha|\le 2} \|\partial^\alpha u\|_{L^2(\mathbb{R}^d)}^2 \equiv \sum_{|\alpha|\le 2} \|\xi^\alpha \widehat{u}\|_{L^2(\mathbb{R}^d)}^2 = \int_{\mathbb{R}^d} \big(\sum_{|\alpha|\le 2} |\xi^\alpha|^2\big)\big|\widehat{u}(\xi)\big|^2 \,\mathrm{d}\xi.$$

For any $\xi \in \mathbb{R}^d$ we clearly have

$$1 + |\xi|^4 \le \sum_{|\alpha|\le 2} |\xi^\alpha|^2,$$

which shows $\|u\|_{T_2} \le \|u\|_{H^2}$. At the same time, denoting $c := \#\{\alpha \in \mathbb{N}_0^d : |\alpha| \le 2\}$, we obtain

$$\begin{aligned}\sum_{|\alpha|\le 2} |\xi^\alpha|^2 \le \sum_{|\alpha|\le 2} \langle\xi\rangle^{2|\alpha|} \le \sum_{|\alpha|\le 2} \langle\xi\rangle^4 \\ = c\langle\xi\rangle^4 = c(1+|\xi|^2)^2 \le 2c(1+|\xi|^4),\end{aligned}$$

which gives $\|u\|_{H^2}^2 \le 2c\|u\|_{T_2}^2$. Therefore, the both norms are equivalent. It follows by Proposition 1.43 then that $C_c^\infty(\mathbb{R}^d)$ is dense in $D(T_2)$ in the graph norm, i.e. that $\overline{T_0} = T_2$ (i.e. T_0 is essentially self-adjoint). □

Remark 1.46. One observes that most of the above constructions are based on the fact that for $P = -\Delta$ one has $p(\xi) = |\xi|^2 > 0$ for $\xi \ne 0$. Most assertions can be extended to other differential expressions P with constant coefficients such that $p(\xi)$ do not vanish at least for sufficiently large $|\xi|$: such P are usually called *elliptic*.

Definition 1.47 (Free Laplacian in $\mathbb{R}^d$). The operator T in $L^2(\mathbb{R}^d)$ defined by

$$D(T) = H^2(\mathbb{R}^d), \quad Tu = -\Delta u,$$

is called the *free Laplacian in* $\mathbb{R}^d$. As discussed in Example 1.45, it is a self-adjoint operator, and it is essentially self-adjoint on $C_c^\infty(\mathbb{R}^d)$ (i.e. its restriction on the test functions is an essentially self-adjoint operator).

The free Laplacian T will be of importance for the rest of the course. In fact, many operators we are going to study will be of the form $T + V$ with some perturbation V. □

Therefore, for $\Omega = \mathbb{R}^d$ and $P = -\Delta$ we have shown that $P_{\min} = P_{\max}$ and $D(P_{\max}) = H^2(\Omega)$. Nevertheless, it should be noted that these equalities do not hold for general open sets Ω (even if P remains the same). The theory of Sobolev spaces on general open sets is much more involved, as no direct equivalent of the Fourier transform is available. We discuss now some key points that will be important later.

Definition 1.48. By $C^\infty(\overline{\Omega})$ we will denote the set of functions defined on Ω that can be extended to a function in $C_c^\infty(\mathbb{R}^d)$.

Definition 1.49 (Open sets with "good" boundaries). One says that an open set $\Omega \subset \mathbb{R}^d$ has C^m (respectively, *Lipschitz*) *boundary*, if for any $p \in \partial\Omega$ there exist $\varepsilon > 0$, a Cartesian coordinate system $(y_1, \dots, y_d)$ centered at p and a C^m (respectively, Lipschitz) function h defined on a neighborhood of zero in $\mathbb{R}^{d-1}$, with $h(0) = 0$, such that

$$\Omega \cap B_\varepsilon(p) = \big\{ y \in B_\varepsilon(0) : \, y_d < h(y_1, \dots, y_{d-1}) \big\}.$$

Note that open sets with C^1 boundaries are usually considered in Analysis III when one discusses Gauss integral formulas and partial integration in higher dimensions.

Remark 1.50. In the above definition, it is important that Ω lies "on one side" of its boundary. A simple example of an open set Ω which is not covered by the above definition is $\Omega := \mathbb{R}^2 \setminus I$, where I is any closed interval ("plane with a cut").

The following important (but very technical) result can be viewed as a kind of replacement for the above Proposition 1.43 (we give it without proof):

Theorem 1.51 (Density of smooth functions). *Let $\Omega \subset \mathbb{R}^d$ be an open set with C^0 boundary, then $C^\infty(\overline{\Omega})$ is dense in $H^k(\Omega)$ for any $k \in \mathbb{N}$.*

Therefore, if the open set Ω has a "good" boundary, then $H^k(\Omega)$ can be alternatively defined as the completion of $C^\infty(\overline{\Omega})$ in the H^k-norm.

Remark 1.52 (Example of $P_{\min} \neq P_{\max}$). Let $P = -\Delta$ and $\Omega \subset \mathbb{R}^d$ be a bounded open set with smooth boundary (so that one can apply the Gauss integral formula). It is clear that $C^\infty(\overline{\Omega}) \subset D(P_{\max})$. On the other hand, for $u, v \in C^\infty(\overline{\Omega})$ one has (Green formula!)

$$\langle u, P_{\max} v\rangle - \langle P_{\max} u, v\rangle = \int_\Omega \overline{\Delta u}\, v \,\mathrm{d}x - \int_\Omega \overline{u}\, \Delta v \,\mathrm{d}x = \int_{\partial\Omega} \big(\overline{\partial_n u}\, v - \overline{u}\, \partial_n v\big)\,\mathrm{d}s,$$

where $\partial_n u := n \cdot \nabla u$ is the outer normal derivative (n is the smooth unit normal vector field on $\partial\Omega$ pointing to the exterior of Ω) and $\mathrm{d}s$ means the integration with respect to the hypersurface measure. It is clear that u and v can be chosen in such a way that the result is non-zero, and it follows that $P_{\max}$ is not symmetric (so it cannot be self-adjoint). On the other hand, $P_{\min}$ is always symmetric, so $P_{\min} \neq P_{\max}$. In fact, one needs to take a restriction of $P_{\max}$ in order to obtain a self-adjoint operator, and usually such a restriction is formulated in terms of a boundary condition that guarantees that the integral over the boundary in the above identity becomes identically zero.

According to the general rule (Remark 1.23), define

$$H^k_{\mathrm{loc}}(\Omega) := \{ f : \Omega \to \mathbb{C} : \varphi f \in H^k(\mathbb{R}^d) \text{ for any } \varphi \in C_c^\infty(\Omega)\}.$$

If one proceeds as in Proposition 1.25, one shows the equality

$$H^k_{\mathrm{loc}}(\Omega) = \{f : \Omega \to \mathbb{C} : f \in H^1(B) \text{ for any ball } B \text{ with } \overline{B} \subset \Omega\},$$

but both characterizations are important. By applying Theorem 1.44 (Sobolev embedding) to the products φf one arrives at

Corollary 1.53 (Local Sobolev embedding). *For any $k \in \mathbb{N}$ and $m \in \mathbb{N}_0$ with $k > m + \frac{d}{2}$ one has $H^k_{\mathrm{loc}}(\Omega) \subset C^m(\Omega)$.*

Finally let us mention the following important result, which is a local version of Corollary 1.42:

Theorem 1.54 (Interior elliptic regularity). *Let $u \in L^2_{\mathrm{loc}}(\Omega)$ with $\Delta u \in H^k_{\mathrm{loc}}(\Omega)$, then $u \in H^{k+2}_{\mathrm{loc}}(\Omega)$.*

Proof idea. Let $\varphi \in C^\infty_c(\Omega)$, then $\varphi u \in L^2(\mathbb{R}^d)$ and

$$\Delta(\varphi u) = u\Delta\varphi + 2\nabla\varphi \cdot \nabla u + \varphi\Delta u.$$

One can give a complete proof for $k = 0$ and under the additional assumption $u \in H^1_{\mathrm{loc}}(\Omega)$: in that case all components of ∇u are in L^2_{loc}, and the function on the right-hand side is in $L^2(\mathbb{R}^d)$. Now one applies Corollary 1.42 and obtains $\varphi u \in H^2(\mathbb{R}^d)$.

In the general case, the argument is in the same spirit, but the components of ∇u should be considered as functions in $H^s(\mathbb{R}^d)$ with some $s < 0$. (Related results are discussed in dedicated PDE courses.) $\square$

If one iteratively applies Theorem 1.54 and uses Corollary 1.53, one obtains:

Proposition 1.55. *Let $u \in L^2(\Omega)$ with $\Delta u = Vu + f$ in Ω for some $V, f \in C^\infty(\Omega)$, then $u \in C^\infty(\Omega)$.*

Proof. It is clear that the product of a C^∞ function with a H^k_{loc} function belongs to H^k_{loc}. If $u \in L^2$, then $Vu + f \in L^2_{\mathrm{loc}} = H^0_{\mathrm{loc}}$, then $u \in H^2_{\mathrm{loc}}$ by Theorem 1.54. Then $Vu + f \in H^2_{\mathrm{loc}}$, and by Theorem 1.54 one obtains $u \in H^4_{\mathrm{loc}}$ and so on. Therefore, $u \in H^k_{\mathrm{loc}}$ for any $k \in \mathbb{N}$, and by Corollary 1.53 one obtains $u \in C^m$ with any $m \in \mathbb{N}$. $\square$

1.6 Operators defined by forms

Definition 1.56. Let $\mathcal{H}$ be a Hilbert space and $D(t) \subset \mathcal{H}$ a dense subspace. A *sesquilinear form* t in a Hilbert space $\mathcal{H}$ with domain $D(t)$ is a map

$$t : \mathcal{H} \times \mathcal{H} \supset D(t) \times D(t) \to \mathbb{C}$$

which is linear with respect to the second argument and conjugate linear with respect to the first one.[4] It is called:

[4] In the literature, one uses sometimes the terms *bilinear form* and *quadratic form* for the same objects.

- *symmetric* (or *Hermitian*) if $t(u,v) = \overline{t(v,u)}$ for all $u, v \in D(t)$,
- *semibounded from below* if t is symmetric and for some $c \in \mathbb{R}$ one has
$$t(u,u) \geq -c\|u\|^2 \text{ for all } u \in D(t);$$
in this case we write $t \geq -c$,
- *closed* if $t \geq -c$ and the domain $D(t)$ with the scalar product
$$\langle u, v\rangle_t := t(u,v) + (c+1)\langle u, v\rangle_{\mathcal{H}}$$
is a Hilbert space. It is an easy exercise to show that this property does not depend on the particular choice of c.

For a symmetric form t we often use the shorthand notation:

$$t(u) := t(u,u).$$

Remark 1.57. (a) Any scalar product is a symmetric sesquilinear non-negative form.

(b) From the linear algebra and the functional analysis is it known that a symmetric sesqulinear form t is uniquely determined by its "diagonal values" $t(u)$ with $u \in D(t)$: for any $u, v \in D(t)$ one has the polarization identity

$$t(u,v) = \sum_{k=1}^{4} i^k t(u - i^k v).$$

Remark 1.58. It is known from the functional analysis that if t is a symmetric sesquilinear form with $D(t) = \mathcal{H}$ such that for some $c > 0$ there holds

$$|t(u,v)| \leq c\|u\|_{\mathcal{H}}\,\|v\|_{\mathcal{H}} \text{ for all } u, v \in \mathcal{H},$$

then there exists a uniquely defined operator $T = T^* \in \mathcal{L}(\mathcal{H})$ such that

$$t(u,v) = \langle u, Tv\rangle_{\mathcal{H}} \text{ for all } u, v \in \mathcal{H}.$$

We are going to find an analogous result for more general sesqulinear forms.

Definition 1.59 (Operator generated by a closed form). Let t be a closed sesquilinear form in $\mathcal{H}$. The linear operator T in $\mathcal{H}$ *generated by* (or *associated with*) the form t is defined by

$$\big(v \in D(T) \text{ and } f = Tv\big) \text{ iff } v \in D(t) \text{ with } t(u,v) = \langle u, f\rangle_{\mathcal{H}} \text{ for all } u \in D(t).$$

The following proposition can be considered is an easy exercise (it shows that Definition 1.59 is compatible with the respective construction involving bounded operators as in Remark 1.58).

Proposition 1.60. *Let t be a closed sesquilinear form in $\mathcal{H}$ and T be the operator generated by t. Furthermore, let $B = B^* \in \mathcal{L}(\mathcal{H})$. Then*

$$t_B : (u, v) \mapsto t(u, v) + \langle u, Bv\rangle_{\mathcal{H}}, \quad D(t_B) = D(t),$$

is a closed sesqulinear form, and the operator T_B generated by t_B is $T_B : u \mapsto Tu+Bu$ with $D(T_B) = D(T)$.

The following result is of crucial importance for many subsequent examples and computations. In fact, many operators we are going to study will be defined through their sesquilinear forms.

Theorem 1.61. *The operator T in Definition 1.59 is self-adjoint in $\mathcal{H}$, and $D(T)$ is dense in $D(t)$ with respect to $\langle\cdot,\cdot\rangle_t$.*

Proof. We consider the case $t \geq 1$, then $\langle u, v\rangle_t = t(u, v)$ and $t(u, u) = \|u\|_t^2 \geq \|u\|^2_{\mathcal{H}}$. (The general case easily follows by Proposition 1.60: exercise). Remark first that for $v \in D(T)$ we have $\|v\|^2_{\mathcal{H}} \leq t(v, v) = \langle v, Tv\rangle_{\mathcal{H}} \leq \|v\|_{\mathcal{H}}\,\|Tv\|_{\mathcal{H}}$ and then $\|Tv\|_{\mathcal{H}} \geq \|v\|_{\mathcal{H}}$, which shows that T is injective.

Now let us show that $T : D(T) \to \mathcal{H}$ is surjective. Let $f \in \mathcal{H}$. For any $u \in D(t)$ one has

$$\big|\langle u, f\rangle_{\mathcal{H}}\big| \leq \|u\|_{\mathcal{H}} \cdot \|f\|_{\mathcal{H}} \leq \|f\|_{\mathcal{H}}\|u\|_t.$$

Hence, $D(t) \ni u \mapsto \langle u, f\rangle_{\mathcal{H}} \in \mathbb{C}$ is a continuous antilinear map, and by the Riesz theorem there is $v \in D(t)$ with $\langle u, f\rangle_{\mathcal{H}} = \langle u, v\rangle_t \equiv t(u, v)$ for all $u \in D(t)$. By definition this means that $v \in D(T)$ with $f = Tv$. This shows the surjectivity.

We further remark that for any $u, v \in D(T)$ we have, using the symmetry of t,

$$\langle u, Tv\rangle_{\mathcal{H}} = t(u, v) = \overline{t(v, u)} = \overline{\langle v, Tu\rangle_{\mathcal{H}}} = \langle Tu, v\rangle_{\mathcal{H}}.$$

Therefore, T is symmetric, and then T^{-1} is symmetric as well (using the same argument as in Proposition 1.21). Hence, the operator T^{-1} is symmetric and defined everywhere, hence, it is self-adjoint. Then $T = (T^{-1})^{-1}$ is self-adjoint by Proposition 1.21.

To prove the remaining statement (density) let $h \in D(t)$ with $\langle v, h\rangle_t = 0$ for all $v \in D(T)$, then we need to show that $h = 0$. Remark that by assumption we have

$$0 = \langle v, h\rangle_t = t(v, h) = \overline{t(h, v)} = \overline{\langle h, Tv\rangle_{\mathcal{H}}} = \langle Tv, h\rangle_{\mathcal{H}}.$$

As the vectors Tv cover the whole of $\mathcal{H}$ as v runs through $D(T)$, one has $h = 0$. □

Now let us give some "canonical" examples of operators defined by forms. We will see them very often.

Example 1.62 (Free Laplacian revisited). Consider $\mathcal{H} = L^2(\mathbb{R}^d)$ and the form

$$t(u, v) = \int_{\mathbb{R}^d} \overline{\nabla u} \cdot \nabla v \,\mathrm{d}x, \quad D(t) = H^1(\mathbb{R}^d),$$

which is clearly closed: in fact, $t \geq 0$ and $\langle \cdot, \cdot \rangle_t$ is exactly the H^1-scalar product, and $H^1(\mathbb{R}^d)$ is complete. Let us find the associated operator T, which is already known to be self-adjoint due to Theorem 1.61.

Let $v \in D(T)$ and $f := Tv$, then for any $u \in H^1(\mathbb{R}^d)$ we have

$$\int_{\mathbb{R}^d} \overline{\nabla u} \cdot \nabla v \, dx = \int_{\mathbb{R}^d} \overline{u} f \, \mathrm{d}x.$$

In particular, this equality holds for $u \in C_c^\infty(\mathbb{R}^d) \subset H^1(\mathbb{R}^d)$, which gives

$$\int_{\mathbb{R}^d} \overline{u} f \, \mathrm{d}x = \int_{\mathbb{R}^d} \overline{\nabla u} \cdot \nabla v \, \mathrm{d}x \quad \text{for all } u \in C_c^\infty(\mathbb{R}^d).$$

We now use the definition of weak derivatives:

$$\begin{aligned}\int_{\mathbb{R}^d} \overline{\nabla u} \cdot \nabla v \, \mathrm{d}x &= \sum_{j=1}^d \int_{\mathbb{R}^d} \overline{\partial_j u} \, \partial_j \nabla v \, \mathrm{d}x = -\sum_{j=1}^d \int_{\mathbb{R}^d} \overline{\partial_j^2 u} \, v \, \mathrm{d}x \\ &= \int_{\mathbb{R}^d} \overline{\Big(-\sum_{j=1}^d \partial_j^2 u \Big)} \, v \, \mathrm{d}x = \int_{\mathbb{R}^d} \overline{-\Delta u} \, v \, \mathrm{d}x,\end{aligned}$$

which means that $f = -\Delta v \in L^2(\mathbb{R}^d)$ (weakly). Due to the global elliptic regularity (Corollary 1.42) one obtains $v \in H^2(\mathbb{R}^d)$, which means that T is a restriction of the free Laplacian in $\mathbb{R}^d$ (see Definition 1.47). The maximality property of self-adjoint operators implies that T is exactly the free Laplacian in $\mathbb{R}^d$.

Example 1.63 (Neumann Laplacian). Let $\Omega \subset \mathbb{R}^d$ and $\mathcal{H} = L^2(\Omega)$. Consider the sesquilinear form

$$t_N(u, v) = \int_\Omega \overline{\nabla u} \cdot \nabla v \, \mathrm{d}x, \quad D(t) = H^1(\Omega).$$

The form is closed due to the completeness of $H^1(\Omega)$, and the associated self-adjoint operator T_N is called the *Neumann Laplacian* on Ω.

In order to understand T_N in a better way, remark first that $C_c^\infty(\Omega) \subset D(t_N)$. Therefore, if $v \in D(T_N)$ and $f = T_N v$, then

$$\int_\Omega \overline{\nabla u} \cdot \nabla v \, dx = \int_\Omega \overline{u} f \, \mathrm{d}x \text{ for all } u \in C_c^\infty(\Omega),$$

and similarly to the preceding example one obtains $f = -\Delta v$, i.e. T acts as the weak Laplacian.

Now assume that Ω has a "good" boundary such that the Gauss integral formula is valid (for example, bounded, with C^1 boundary). Let $v \in C^\infty(\overline{\Omega})$. As $C^\infty(\overline{\Omega})$ is dense in $H^1(\Omega)$, the inclusion $v \in D(T_N)$ is equivalent to the

$$\int_\Omega \overline{\nabla u} \cdot \nabla v \, \mathrm{d}x = \int_\Omega \overline{u}(-\Delta v) \, \mathrm{d}x \text{ for all } u \in C^\infty(\overline{\Omega}). \tag{1.15}$$

The Gauss integral formula gives

$$\int_\Omega \overline{u}(-\Delta v)\,\mathrm{d}x = -\int_{\partial\Omega} \overline{u}\,\partial_n v\,\mathrm{d}s + \int_\Omega \overline{\nabla u}\cdot\nabla v\,\mathrm{d}x,$$

and (1.15) is satisfied if and only if

$$\int_{\partial\Omega} \overline{u}\,\partial_n v\,\mathrm{d}s = 0 \text{ for all } u \in C_c^\infty(\overline{\Omega}).$$

As this point we admit that the restrictions of functions from $C_c^\infty(\overline{\Omega})$ form a dense subset of $L^2(\partial\Omega)$ (which is an easy consequence of the Stone-Weierstrass theorem), which then shows that for a function $v \in C^\infty(\overline{\Omega})$ one has the equivalence

$$v \in D(T_N) \quad \Leftrightarrow \quad \partial_n v = 0 \text{ on } \partial\Omega.$$

The condition $\partial_n v = 0$ on $\partial\Omega$ is called the *Neumann boundary condition.*

For the next example we will introduce an additional class of Sobolev spaces:

Definition 1.64 (Spaces H_0^k). For a non-empty open set $\Omega \subset \mathbb{R}^d$ and $k \in \mathbb{N}$ define

$$H_0^k(\Omega) := \text{the closure of } C_c^\infty(\Omega) \text{ in } H^k(\Omega).$$

Example 1.65 (Dirichlet Laplacian: first attempt). Let $\Omega \subset \mathbb{R}^d$ and $\mathcal{H} = L^2(\Omega)$. Consider the sesquilinear form

$$t_D(u,v) = \int_\Omega \overline{\nabla u}\cdot\nabla v\,\mathrm{d}x, \quad D(t) = H_0^1(\Omega).$$

The form is closed due to the preceding definition of $H_0^1(\Omega)$, and the associated self-adjoint operator T_D is called the *Dirichlet Laplacian* on Ω. One shows again that T_D acts as $v \mapsto -\Delta v$.

As $C_c^\infty(\Omega)$ is dense in $H_0^1(\Omega)$, a function $v \in H_0^1(\Omega)$ belongs to $D(T_D)$ if and only if one has

$$\int_\Omega \overline{\nabla u}\cdot\nabla v\,\mathrm{d}x = \int_\Omega \overline{u}(-\Delta v)\,\mathrm{d}x = 0 \text{ for all } u \in C_c^\infty(\Omega).$$

On the other hand, this condition is always satisfied due to the definition of the weak derivatives. Therefore,

$$D(T_D) = \{v \in H_0^1(\Omega) :\ \Delta v \in L^2(\Omega)\}, \quad T_D v = -\Delta v.$$

The problem with this equality is that it is not very informative: the structure of the set on the right-hand side remains unclear. For example: which functions $v \in C^\infty(\overline{\Omega})$ belong to $H_0^1(\Omega)$? This will be answered below.

Remark 1.66. For some Ω one has $H_0^1(\Omega) = H^1(\Omega)$, for example for $\Omega = \mathbb{R}^d$ (Proposition 1.43), in that case the Dirichlet and Neumann Laplacians coincide. Now we are going to discuss cases with $H_0^1(\Omega) \neq H^1(\Omega)$.

Proposition 1.67. *Let $\Omega \subset \mathbb{R}^d$ be an open set. For a function $u : \Omega \to \mathbb{C}$ we denote by $\widetilde{u}$ its extension by zero to the whole of $\mathbb{R}^d$. Then*

$$H_0^k(\Omega) \subset \big\{u \in H^k(\Omega) : \widetilde{u} \in H^k(\mathbb{R}^d)\big\}. \tag{1.16}$$

Proof. Let L be the set on the right-hand side of (1.16). Remark that if $u \in L$, then $\|u\|_{H^k(\Omega)} = \|\widetilde{u}\|_{H^k(\mathbb{R}^d)}$ as $\widetilde{u}$ and all its derivatives are zero outside Ω. One has the obvious inclusion $C_c^\infty(\Omega) \subset L$. Let $u \in H_0^k(\Omega)$, then there exist $(\varphi_n) \subset C_c^\infty(\Omega)$ with $\|u - \varphi_n\|_{H^k(\Omega)} \overset{n\to\infty}{\to} 0$. In particular, φ_n converge to u in $L^2(\Omega)$. On the other hand, for any $\varphi \in C_c^\infty(\Omega)$ one has $\widetilde{\varphi} \in C_c^\infty(\mathbb{R}^d)$ with $\|\widetilde{\varphi}\|_{H^k(\mathbb{R}^d)} = \|\varphi\|_{H^k(\Omega)}$. It follows that $(\widetilde{\varphi}_n)$ is a Cauchy sequence in $H^k(\mathbb{R}^d)$ and $\widetilde{\varphi}_n$ converges in $H^k(\mathbb{R}^d)$ to some $g \in H^k(\mathbb{R}^d)$. One has

$$\int_\Omega |\varphi_n - g|^2 \,\mathrm{d}x = \int_\Omega |\widetilde{\varphi}_n - g|^2 \,\mathrm{d}x \le \int_{\mathbb{R}^d} |\widetilde{\varphi}_n - g|^2 \,\mathrm{d}x \le \|\widetilde{\varphi}_n - g\|^2_{H^k(\mathbb{R}^d)} \to 0,$$

and due to the uniqueness of the limit one has $g = f$ a.e. in Ω. Similarly,

$$\int_{\Omega^\complement} |g|^2 \,\mathrm{d}x = \int_{\Omega^\complement} |\widetilde{\varphi}_n - g|^2 \,\mathrm{d}x \le \int_{\mathbb{R}}^d |\widetilde{\varphi}_n - g|^2 \,\mathrm{d}x \le \|\widetilde{\varphi}_n - g\|^2_{H^k(\mathbb{R}^d)} \to 0,$$

which shows that $g = 0$ a.e. in $\Omega^\complement$. Therefore, $g = \widetilde{u}$. As $g \in H^k(\mathbb{R}^d)$, one has the inclusion $H_0^k(\Omega) \subset L$. □

Remark that the above proof holds for any open Ω. Unter additional assumption one can prove the reverse inclusion, which leads to the following assertion:

Proposition 1.68. *If $\Omega \subset \mathbb{R}^d$ is a bounded open set with Lipschitz boundary, then the inclusion* (1.16) *becomes an equality.*

Proof idea. Let L be the set on the right-hand side of (1.16). In view of Proposition 1.67 one needs to show $L \subset H_0^k(\Omega)$.

Assume first that Ω has a special shape: if (r, θ) are the standard polar coordinates in $\mathbb{R}^d$, then

$$\Omega = \{(r, \theta) : r < h(\theta)\}, \quad h : \mathbb{S}^{d-1} \to (0, \infty) \text{ continuous.} \tag{1.17}$$

Let $t > 0$, then one easily shows that the linear map $\Phi_t : H^k(\mathbb{R}^d) \to H^k(\mathbb{R}^d)$ given by $\Phi_t u(x) = u\big((1+t)x\big)$ satisfies $\Phi_t u \to u$ as $t \to 0^+$ for any $u \in H^k(\mathbb{R}^d)$. Now let $u \in L$. Pick $\varepsilon > 0$ and find first some $t > 0$ such that $\|\Phi_t\widetilde{u} - \widetilde{u}\|_{H^k(\mathbb{R}^d)} < \varepsilon$. Remark that $\widetilde{u} = 0$ outside Ω, hence, $\Phi_t\widetilde{u} = 0$ outside $(1+t)^{-1}\Omega$. In particular, there is a compact subset $K \subset \Omega$ such that $\Phi_t\widetilde{u} = 0$ outside K. Then we consider $v_\delta := \rho_\delta \star (\Phi_t\widetilde{u})$: for sufficiently small δ one has $v \in C_c^\infty(\Omega)$ and $\|v_\delta - \Phi_t\widetilde{u}\|_{H^k(\mathbb{R}^d)} < \varepsilon$. Then $\|u - v_\delta\|_{H^k(\Omega)} = \|\widetilde{u} - v_\delta\|_{H^k(\mathbb{R}^d)} \le \|\Phi_t\widetilde{u} - \widetilde{u}\|_{H^k(\mathbb{R}^d)} + \|v_\delta - \Phi_t\widetilde{u}\|_{H^k(\mathbb{R}^d)} < 2\varepsilon$. As $\varepsilon > 0$ is arbitrary, we obtain $u \in H_0^1(\Omega)$, which shows the inclusion $L \subset H_0^k(\Omega)$ for Ω as in (1.17).

General Ω are handled using partitions of unity. One covers $\partial\Omega$ by balls B_j, $j \in \{1, \dots, n\}$, such that $\Omega \cap B_j$ as in (1.17) (the Lipschitz conidition is important for this part of construction). Then there exists a subordinated partition of unity, i.e. functions $\chi_j \in C_c^\infty(\mathbb{R}^d)$, $j \in \{0, \dots, n\}$, such that $\operatorname{supp} \chi_0 \subset \Omega$ and $\operatorname{supp} \chi_j \subset B_j$ for $j \geq 1$, and $\sum_j \chi_j = 1$ in Ω. If $u \in L$, then using the first part of the argument one shows that $\chi_j u \in H_0^k(\Omega)$ for each j. Taking the sum over j one shows that $u \in H_0^1(\Omega)$. This shows the inclusion $L \subset H_0^k(\Omega)$ for Ω with bounded Lipschitz boundaries. $\square$

Remark 1.69 (Dirichlet Laplacian: second attempt). Let $\Omega \subset \mathbb{R}^d$ be a bounded open set with a C^1 boundary (so that the Gauss integral formula is valid). Let $u \in C^\infty(\overline{\Omega})$, then $u \in H_0^1(\Omega)$ if and only if $\partial_j \widetilde{u} \in L^2(\Omega)$ for any $j \in \{1, \dots, d\}$ (Proposition 1.67). Due to the definition of weak derivatives this is equivalent to the existence of $g_j \in L^2(\mathbb{R}^d)$ such that

$$\int_\Omega u\, \partial_j \varphi \,\mathrm{d}x \equiv \int_{\mathbb{R}^d} \widetilde{u}\, \partial_j \varphi \,\mathrm{d}x = -\int_{\mathbb{R}^d} g_j \varphi \,\mathrm{d}x.$$

Using the integration by parts on the left-hand side one obtains, with n being the outer unit normal on $\partial\Omega$,

$$-\int_\Omega \partial_j u\, \varphi \,\mathrm{d}x + \int_{\partial\Omega} n_j u \varphi \,\mathrm{d}s = -\int_\Omega g_j \varphi_j \,\mathrm{d}x - \int_{\Omega^\complement} g_j \varphi_j \,\mathrm{d}x$$

for all $\varphi \in C_c^\infty(\mathbb{R}^d)$ and all $j \in \mathbb{N}$. Taking φ supported in the interior of the exterior of Ω one sees that $g_j = \partial_j u$ in Ω and $g_j = 0$ in $\Omega^\complement$, and the above conditions holds if and only if

$$\int_{\partial\Omega} n_j u \varphi \,\mathrm{d}s = 0$$

for all $\varphi \in C_c^\infty(\mathbb{R}^d)$ and all $j \in \mathbb{N}$, i.e. if $u = 0$ on $\partial\Omega$.

Therefore, for Ω as above and $u \in C^\infty(\overline{\Omega})$ one has the equivalence

$$u \in H_0^1(\Omega) \quad \Leftrightarrow \quad u = 0 \text{ on } \partial\Omega.$$

In particular, for the Dirichlet Laplacian T_D in Ω one has

$$C^\infty(\overline{\Omega}) \cap D(T_D) = \{u \in C^\infty(\overline{\Omega}) : \ u = 0 \text{ on } \partial\Omega\}.$$

From the proof it is seen that the same conclusion holds for a larger class of domains, for example, for polyhedra (as the Gauss integral formula still holds).

Remark 1.70. In the two above examples (Dirichlet and Neumann Laplacians) we see several important features of forms and associated operators:

- Closed sesquilinear forms do *not* have the maximality property, i.e. a closed sesquilinear form have can a closed extension with a strictly larger domain,
- The fact that one closed form extends another closed form does *not* imply the same relation for the associated operators. $\square$

Remark 1.71. If $\Omega \subset \mathbb{R}^d$ is bounded with C^∞ boundary, then with the help of advanced methods (elliptic regularity at the boundary, which is almost never covered during the university studies) one can show that the domains of the Neumann Laplacian T_N and the Dirichlet Laplacian T_D on Ω are

$$\begin{aligned} D(T_N) &= \{u \in H^2(\Omega) : \partial_n u|_{\partial\Omega} = 0\}, \\ D(T_D) &= H^2(\Omega) \cap H_0^1(\Omega) = \{u \in H^2(\Omega) : u|_{\partial\Omega} = 0\}, \end{aligned}$$

while $u|_{\partial\Omega}$ and $\partial_n u|_{\partial\Omega}$ are the so-called traces defined in a special way (pointwise definition for functions smooth up to boundary, then extension by density) as functions in $L^2(\partial\Omega)$. This description is important in some cases, but it will not be used in this course. Remark that the description fails for non-smooth boundaries (there are simple examples in which the domains are not contained in H^2).

1.7 Semibounded operators

For what follows we will need an additional notion:

Definition 1.72 (Closable form). A symmetric sesquilinear form t is called *closable*, if there exists a closed sesquilinear form extending t. The closed sesquilinear form extending t and having the smallest domain is called the *closure* of t and denoted $\bar{t}$.

Proposition 1.73 (Criterion of closability). *A lower semibounded sesquilinear form t in $\mathcal{H}$ is closable if and only if*

$$\begin{aligned} &\textit{for any sequence } (w_n) \subset D(t) \textit{ with } w_n \overset{n\to\infty}{\longrightarrow} 0 \textit{ in } \mathcal{H} \textit{ and} \\ &t(w_m - w_n) \overset{m,n\to\infty}{\longrightarrow} 0 \textit{ there holds } \lim_{n\to\infty} t(w_n) = 0. \end{aligned} \tag{1.18}$$

If this condition is satisfied, then $D(\bar{t})$ is the completion of $D(t)$ with respect to $\|\cdot\|_t$ and $\bar{t}$ is the extension of t by continuity.

Proof. It follows from the definitions that:

- a lower semibounded sequilinear form t is closed if and only if the conditions

  $$u_n \in D(t), \quad u_n \overset{n\to\infty}{\longrightarrow} u \text{ in } \mathcal{H}, \quad t(u_m - u_n) \overset{m,n\to\infty}{\longrightarrow} 0,$$

  imply $u \in D(t)$ with $t(u) = \lim_n t(u_n)$,

- a symmetric lower semibounded sesquilinear form t is closable if and only if for any two sequences $(u_n), (v_n) \subset D(t)$ such that

  $$u_n - v_n \overset{n\to\infty}{\longrightarrow} 0 \text{ in } \mathcal{H}, \qquad t(u_m - u_n) \overset{m,n\to\infty}{\longrightarrow} 0, \qquad t(v_m - v_n) \overset{m,n\to\infty}{\longrightarrow} 0 \tag{1.19}$$

  there holds $\lim_{n\to\infty} t(u_n) = \lim_{n\to\infty} t(v_n)$.

Assume that t is closable. Let $(w_n) \subset D(t)$ such that $w_n \overset{n\to\infty}{\longrightarrow} 0$ in $\mathcal{H}$ and $t(w_m - w_n) \overset{m,n\to\infty}{\longrightarrow} 0$. The sequences $u_n := w_n$ and $v_n := 0$ satify (1.19), therefore,

$$\lim_{n\to\infty} t(w_n) = \lim_{n\to\infty} t(u_n) = \lim_{n\to\infty} t(v_n) = \lim_{n\to\infty} t(0) = \lim_{n\to\infty} 0 = 0.$$

This shows that (1.18) is satisfied.

Now let t satisfy (1.18). Let $(u_n), (v_n) \subset D(t)$ be two sequences obeying (1.19). We need to show that $\lim_{n\to\infty} t(u_n) = \lim_{n\to\infty} t(v_n)$, which is equivalent to $\lim_{n\to\infty} \|u_n\|_t = \lim_{n\to\infty} \|v_n\|_t$. For $w_n := u_n - v_n$ one has $w_n \to 0$ in $\mathcal{H}$ and

$$\|w_m - w_n\|_t = \|u_m - v_m - u_n + v_n\|_t \le \|u_m - u_n\|_t + \|v_m - v_n\|_t \overset{m,n\to\infty}{\longrightarrow} 0.$$

Therefore, the sequence (w_n) satisfies (1.18), and one obtains with the help of the triangle inequality $\big|\|u_n\|_t - \|v_n\|_t\big| \le \|u_n - v_n\|_t = \|w_n\|_t \to 0$. Hence, t is closable.

The last assertion is an easy exercise. $\square$

Example 1.74 (Non-closable form). Take $\mathcal{H} = L^2(\mathbb{R})$ and consider the form $t(u,v) = \overline{u(0)}v(0)$ defined on $D(t) = L^2(\mathbb{R}) \cap C^0(\mathbb{R})$. This form is densely defined, symmetric, with $t \ge 0$.

Take any $u \in D(t)$ such that $u(0) = 1$ and consider $u_n : x \mapsto u(nx)$ for $n \in \mathbb{N}$, then $(u_n) \subset D(t)$ with $u_n(0) = 1$ and we have

$$\|u_n\|_{L^2(\mathbb{R})} \to 0, \quad t(u_n - u_m) \equiv 0,$$

but $t(u_n) \equiv 1$. Proposition 1.73 shows that t is not closable. $\square$

We now arrive at a canonical construction of self-adjoint operators, which will allow us to associate self-adjoint operators with some differential expressions having non-smooth coefficients.

Definition 1.75 (Semibounded operator). A symmetric operator T in $\mathcal{H}$ is called *semibounded from below* if there exists a constant $c \in \mathbb{R}$ such that

$$\langle u, Tu\rangle \ge -c\langle u, u\rangle_{\mathcal{H}} \text{ for all } u \in D(T),$$

and in this will be written as $T \ge -c$. $\square$

Proposition 1.76. *Let T be a symmetric, densely defined, semibounded from below linear operator in $\mathcal{H}$, then the induced sesquilinear form t in $\mathcal{H}$ given by*

$$t(u,v) := \langle u, Tv\rangle, \quad D(t) := D(T), \tag{1.20}$$

is semibounded from below and closable.

Proof. The semiboundedness of t is clear. To show the closability we remark that without loss of generality one can assume $T \ge 1$ (the general case is reduced to this one: easy exercise), then $\|u\|_t^2 = t(u,u) \ge \|u\|_{\mathcal{H}}^2$.

Let $(u_n) \subset D(t)$ be $\|\cdot\|_t$-Cauchy $u_n \to 0$ in $\mathcal{H}$. By Proposition 1.3 we need to show that $\lim \|u_n\|_t = 0$. Due to $\big|\|u_n\|_t - \|u_m\|_t\big| \le \|u_n - u_m\|_t$ the real-valued sequence

$\|u_n\|_t$ is Cauchy in $[0,\infty)$, hence, there exists the limit $\lim \|u_n\|_t = \alpha \in [0,\infty)$. We suppose that $\alpha > 0$ and try to arrive at a contradiction.

We have $t(u_n,u_m) = t(u_n,u_n)+t(u_n,u_m-u_n)$. By the Cauchy-Schwarz inequality for $\langle\cdot,\cdot\rangle_t$ we also have $\big|t(u_n,u_m-u_n)\big| \le \|u_n\|_t\|u_m-u_n\|_t$. Recall that $\|u_m-u_n\|_t$ goes to zero for large m,n and that $\|u_n\|_t$ converges to α, in particular, is bounded. We conclude that for any $\varepsilon > 0$ there exists $N > 0$ such that $\big|t(u_n,u_m)-\alpha^2\big| \le \varepsilon$ for all $n,m > N$. Take $\varepsilon = \frac{1}{2}\alpha^2$ and the associated N, then for $n,m > N$ we have $\big|\langle u_n, Tu_m\rangle\big| \equiv \big|t(u_n,u_m)\big| \ge \frac{1}{2}\,\alpha^2$. On the other hand, the term on the left-hand side goes to 0 as $n \to \infty$ (as u_n converges to 0 in $\mathcal{H}$ by assumption). So we obtain a contradiction, and the assertion is proved. □

Definition 1.77 (Friedrichs extension). Let T be a densely defined lower semibounded linear operator in $\mathcal{H}$. Define a sesquilinear form t by (1.20). The self-adjoint operator T_F generated by its closure $\bar{t}$ is called the *Friedrichs extension* of T. □

Corollary 1.78. *Any densely defined lower semibounded linear operator has a self-adjoint extension.*

Remark 1.79 (Form domain). If T is a self-adjoint operator semibounded from below, then it is the Friedrichs extension of itself. The domain of the associated form $\bar{t}$ is usually called the *form domain* of T and is denoted $Q(T)$. The form domain plays an important role in the analysis of self-adjoint operators, in particular, in the variational characterization of eigenvalues using the min-max principle, which will be a central point later. □

Example 1.80 (Schrödinger operators). A basic example for the Friedrichs extension is delivered by Schrödinger operators with lower semibounded potentials.

Let $\mathcal{H} := L^2(\mathbb{R}^d)$ and let $V \in L^2_{\text{loc}}(\mathbb{R}^d)$ and $V \ge -C$ for some $C \in \mathbb{R}$ (i.e. V is real-valued and semibounded from below). In $\mathcal{H} = L^2(\mathbb{R}^d)$ consider the operator T acting as

$$T : u \mapsto -\Delta u + Vu, \quad D(T) = C_c^\infty(\mathbb{R}^d).$$

For any $u,v \in D(T)$ there holds

$$\begin{aligned} t(u,v) := \langle u, Tv\rangle &= \int_{\mathbb{R}^d} \overline{u}(-\Delta v)\,\mathrm{d}x + \int_{\mathbb{R}^d} V\overline{u}v\,\mathrm{d}x = \int_{\mathbb{R}^d} \overline{\nabla u}\cdot\nabla v\,\mathrm{d}x + \int_{\mathbb{R}^d} V\overline{u}v\,\mathrm{d}x,\\ \langle u, Tu\rangle &= \int_{\mathbb{R}^d} |\nabla u|^2\,\mathrm{d}x + \int_{\mathbb{R}^d} V|u|^2\,\mathrm{d}x \ge -C\int_{\mathbb{R}^d} |u|^2\,\mathrm{d}x = -C\|u\|^2_{\mathcal{H}}. \end{aligned}$$

The Friedrichs extension T_F of T will be called the *Schrödinger operator with potential V*. One can show (using standard truncations and convolutions with $*\rho_\delta$) that

$$D(\bar{t}) = H^1_V(\mathbb{R}^d) := \big\{u \in H^1(\mathbb{R}^d) : \sqrt{|V|}u \in L^2(\mathbb{R}^d)\big\}$$

and that $\bar{t}$ is given by the same expression as t. □

Let us extend the above example by including a class of potentials V which are *not* semibounded from below. This will be done using the following classical inequality.

Proposition 1.81 (Hardy inequality). *Let $d \geq 3$ and $u \in C_c^\infty(\mathbb{R}^d)$, then*

$$\int_{\mathbb{R}^d} \big|\nabla u(x)\big|^2 \,\mathrm{d}x \geq \frac{(d-2)^2}{4} \int_{\mathbb{R}^d} \frac{\big|u(x)\big|^2}{|x|^2} \,\mathrm{d}x.$$

Proof. For any $\gamma \in \mathbb{R}$ one has

$$\int_{\mathbb{R}^d} \Big|\nabla u(x) + \gamma \,\frac{x u(x)}{|x|^2}\Big|^2 \,\mathrm{d}x \geq 0,$$

which may be rewritten in the form

$$\begin{aligned}\int_{\mathbb{R}^d} \big|\nabla u(x)\big|^2 \,\mathrm{d}x + \gamma^2 \int_{\mathbb{R}^d} \frac{\big|u(x)\big|^2}{|x|^2} \,\mathrm{d}x \\ \geq -\gamma \int_{\mathbb{R}^d} \bigg[x \cdot \overline{\nabla u(x)} \,\frac{u(x)}{|x|^2} + x \cdot \nabla u(x) \,\frac{\overline{u(x)}}{|x|^2} \bigg] \,\mathrm{d}x. \quad (1.21)\end{aligned}$$

Using the identities

$$\nabla |u|^2 \equiv \nabla(\overline{u} u) = \overline{u}\, \nabla u + u \overline{\nabla u}, \quad \operatorname{div} \frac{x}{|x|^2} = \frac{d-2}{|x|^2},$$

and the integration by parts we obtain

$$\begin{aligned}\int_{\mathbb{R}^d} \bigg[x \cdot \overline{\nabla u(x)} \,\frac{u(x)}{|x|^2} + x \cdot \nabla u(x) \,\frac{\overline{u(x)}}{|x|^2} \bigg] \,\mathrm{d}x &= \int_{\mathbb{R}^d} \nabla \big|u(x)\big|^2 \cdot \frac{x}{|x|^2} \,\mathrm{d}x \\ &= -\int_{\mathbb{R}^d} \big|u(x)\big|^2 \operatorname{div} \frac{x}{|x|^2} \,\mathrm{d}x = -(d-2) \int_{\mathbb{R}^d} \frac{\big|u(x)\big|^2}{|x|^2} \,\mathrm{d}x.\end{aligned}$$

Inserting this equality into (1.21) gives

$$\int_{\mathbb{R}^d} \big|\nabla u(x)\big|^2 \,\mathrm{d}x \geq \gamma\big((d-2) - \gamma\big) \int_{\mathbb{R}^d} \frac{\big|u(x)\big|^2}{|x|^2} \,\mathrm{d}x.$$

We optimize the coefficient on the right-hand side by taking $\gamma = (d-2)/2$, which gives the claim. □

Note that the integral on the right-hand side of the Hardy inequality is not defined for $d \leq 2$, because the function $x \mapsto |x|^{-2}$ is not integrable anymore.

By combining the Hardy inequality with the constructions of Example 1.80 one easily shows the following result:

Corollary 1.82. *Let $d \geq 3$ and $V \in L^2_{\mathrm{loc}}(\mathbb{R}^d)$ be real-valued such that for some $C \in \mathbb{R}$ and a.e. $x \in \mathbb{R}^d$ one has*

$$V(x) \geq -C - \frac{(d-2)^2}{4|x|^2},$$

then the operator $T = -\Delta + V$ defined on $C_c^\infty(\mathbb{R}^d)$ is semibounded from below and, hence, has a self-adjoint extension (Friedrichs extension).

Example 1.83 (Coulomb potential). We would like to show that the operator

$$T = -\Delta + \frac{q}{|x|}$$

in $L^2(\mathbb{R}^3)$ is semibounded from below for any real q. The operator is of importance in quantum physics, the potential $x \mapsto q/|x|$ is referred to as the Coulomb potential of charge q placed at the origin. For $q \geq 0$ we are in the situation of Example 1.80 (the potential is ≥ 0).

If $q < 0$, we estimate with any $p > 0$

$$\frac{|q|}{|x|} = |q|\Big(p \cdot \frac{1}{p|x|}\Big) \leq |q|\,\frac{1}{2}\Big(p^2 + \frac{1}{p^2|x|^2}\Big) = \frac{|q|p^2}{2} + \frac{|q|}{2p^2|x|^2},$$

and for for $p = \sqrt{|q|/8}$ one obtains one obtains

$$\frac{q}{|x|} = -\frac{|q|}{|x|} \geq -\frac{|q|^2}{16} - \frac{1}{4|x|^2},$$

which is covered by Corollary 1.82.

Therefore, for any $q \in \mathbb{R}$ the above operator T has a self-adjoint extension (Friedrichs extension). We will see later that this self-adjoint extension is unique. □

Remark 1.84. The questions addressed in this chapter are non-trivial in the sense that *there exist symmetric operators having no self-adjoint extensions.* For example, the operator

$$T : u \mapsto iu', \quad D(T) = H_0^1(0,\infty),$$

is symmetric in $\mathcal{H} = L^2(0,\infty)$, while its adjoint

$$T^* : u \mapsto iu', \quad D(T^*) = H^1(0,\infty),$$

is not symmetric. Any self-adjoint extensions S of T must obey $T \subset S \subset T^*$, but one easily checks that T^* is the unique closed extension of T. □

2 Spectrum and resolvent

In this section we collect first definitions concerning the spectrum. Some of them are supposed to be known the functional analysis course when applied to bounded operators. Nevertheless, we reinterpret these notions from the point of view of unbounded operators and see some new aspects.

2.1 Definitions and examples

Definition 2.1 (Resolvent set, spectrum, point spectrum). Let T be a linear operator in a Hilbert space $\mathcal{H}$.

- The *resolvent set* $\operatorname{res} T$ consists of the complex numbers z for which the operator $T - z : D(T) \ni u \mapsto Tu - zu \in \mathcal{H}$ is bijective and the inverse $(T-z)^{-1}$ is bounded.
- The *spectrum* $\operatorname{spec} T$is defined by $\operatorname{spec} T := \mathbb{C} \setminus \operatorname{res} T$.
- The *point spectrum* $\operatorname{spec}_p T$ is defined as the set of the eigenvalues of T. □

Note that very often the resolvent set and the spectrum of T are denoted by $\rho(T)$ and $\sigma(T)$, respectively.

Proposition 2.2. *If* $\operatorname{res} T \neq \emptyset$*, then* T *is a closed operator.*

Proof. Let $z \in \operatorname{res} T$, then $\operatorname{gr}(T-z)^{-1}$ is closed by the closed graph theorem, but then the graph of $T-z$ is also closed, as $\operatorname{gr}(T-z)$ and $\operatorname{gr}(T-z)^{-1}$ are isometric in $\mathcal{H} \times \mathcal{H}$. □

Proposition 2.3. *For a* <u>*closed*</u> *operator* T *one has the following equivalence:*

$$z \in \operatorname{res} T \quad \textit{iff} \quad \begin{cases} \ker(T-z) = \{0\}, \\ \operatorname{ran}(T-z) = \mathcal{H}. \end{cases}$$

Proof. The $\Rightarrow$ direction follows from the definition.

Now let T be closed and $z \in \mathbb{C}$ with $\ker(T-z) = \{0\}$ and $\operatorname{ran}(T-z) = \mathcal{H}$. The inverse $(T-z)^{-1}$ is then defined everywhere and has a closed graph (as the graph of $T-z$ is closed), and is then bounded by the closed graph theorem. □

Theorem 2.4 (Properties of the resolvent). *The resolvent set* $\operatorname{res} T$ *is always open (hence, the spectrum* $\operatorname{spec} T$ *is always closed). The operator function*

$$\operatorname{res} T \ni z \mapsto R_T(z) := (T-z)^{-1} \in \mathcal{B}(\mathcal{H})$$

is called the resolvent *of* T*. It is holomorphic and satisfies the identities*

$$R_T(z_1) - R_T(z_2) = (z_1 - z_2) R_T(z_1) R_T(z_2), \tag{2.1}$$

$$R_T(z_1) R_T(z_2) = R_T(z_2) R_T(z_1), \tag{2.2}$$

$$\frac{d}{dz} R_T(z) = R_T(z)^2 \tag{2.3}$$

for all $z, z_1, z_2 \in \operatorname{res} T$.

Proof. Let $z_0 \in \operatorname{res} T$. On $D(T)$ we have the equality

$$T - z = \big(I - (z - z_0)R_T(z_0)\big)(T - z_0). \tag{2.4}$$

Let $|z - z_0| < 1/\|R_T(z_0)\|$, then the bounded operator $I - (z - z_0)R_T(z_0) : \mathcal{H} \to \mathcal{H}$ is bijective and has bounded inverse

$$\big(I - (z - z_0)R_T(z_0)\big)^{-1} = \sum_{j=0}^{\infty} (z - z_0)^j R_T(z_0)^j.$$

Then the operator on the right-hand side of (2.4) is bijective and has a bounded inverse too, which means that $z \in \operatorname{res} T$, and

$$R_T(z) = R_T(z_0)\big(I - (z - z_0)R_T(z_0)\big)^{-1} = \sum_{j=0}^{\infty} (z - z_0)^j R_T(z_0)^{j+1}, \tag{2.5}$$

which shows that R_T is holomorphic.

If $z_1, z_2 \in \operatorname{res} T$, then for any $u \in D(T)$ one has $(T - z_2)u = (T - z_1)u + (z_1 - z_2)u$. Let $v \in \mathcal{H}$ and $u := R_T(z_1)v$, then $(T - z_2)R_T(z_1)v = v + (z_1 - z_2)R_T(z_1)v$. If one applies $R_T(z_2)$ on the both sides, one arrives at

$$R_T(z_1)v = R_T(z_2)v + (z_1 - z_2)R_T(z_2)R_T(z_1)v.$$

As $v \in \mathcal{H}$ is arbitrary and the final expression is symmetric with respect to $z_1 \leftrightarrow z_2$, this shows (2.1) and (2.2). For $z_1 \neq z_2$ we rewrite the last identity as

$$\frac{R_T(z_1) - R_T(z_2)}{z_1 - z_2} = R_T(z_2)R_T(z_1),$$

and for $z_1 \to z_2$ one arrives at (2.3). $\square$

We first make a general remark concerning the computation of the spectrum.

Proposition 2.5 (Weyl sequences). *Let $\lambda \in \mathbb{C}$ and $(u_n) \subset D(T)$ with $u_n \neq 0$ such that*

$$\varepsilon_n := \frac{\|(T - \lambda)u_n\|}{\|u_n\|} \xrightarrow{n \to \infty} 0,$$

(such a sequence (u_n) is called a Weyl sequence for λ), then $\lambda \in \operatorname{spec} T$.

Proof. If $T - \lambda$ is not injective, then $\lambda \in \operatorname{spec} T$ automatically. If $T - \lambda$ in injective, then $\varepsilon_n > 0$ for all, and for $v_n := (T - \lambda)^{-1}u_n$ one has $\|(T - \lambda)^{-1}v_n\| = \varepsilon_n^{-1}\|v_n\|$, which means that $\|(T - \lambda)^{-1}\| \geq \varepsilon_n^{-1}$, and $(T - \lambda)^{-1}$ cannot be bounded. $\square$

Example 2.6. Let (X, μ) be a measure space such that for any $A \subset X$ with $0 < \mu(A) \leq \infty$ there exists $A_0 \subset A$ with $0 < \mu(A_0) < \infty$. Let $f : X \to \mathbb{C}$ be a measurable function defined almost everywhere, then the operator

$$M_f : u \mapsto fu, \quad D(M_f) := \big\{u \in L^2(X, \mu) : fu \in L^2(X, \mu)\big\}$$

is closed in $\mathcal{H} := L^2(X,\mu)$. The *essential range* of f is defined by

$$\operatorname{ess\,ran} f = \Big\{\lambda \in \mathbb{C} : \mu\big\{x : |f(x) - \lambda| < \varepsilon\big\} > 0 \text{ for all } \varepsilon > 0\Big\}.$$

Remark that $\operatorname{ess\,ran} f = \operatorname{ess\,ran} g$ if $f = g$ a.e. One can easily check that if $X \subset \mathbb{R}^d$ is an open set, μ is the Lebesgue measure and f is a continuous function, then the essential range of f coincides with the closure of the usual range of f.

Proposition 2.7 (Spectrum of the multiplication operator). *There holds*

$$\operatorname{spec} M_f = \operatorname{ess\,ran} f, \qquad \operatorname{spec}_p M_f = \big\{\lambda : \mu\{x : f(x) = \lambda\} > 0\big\}.$$

Proof. Let $\lambda \notin \operatorname{ess\,ran} f$, the for some $\varepsilon > 0$ there holds $|f(x) - \lambda| > \varepsilon$ for a.e. $x \in X$, then $1/(f - \lambda)$ is defined almost everywhere and essentially bounded (i.e. coincides a.e. with a bounded function), and then the operator $M_{1/(f-\lambda)}$ is bounded, and one easily checks that this is the inverse for $M_f - \lambda$. On the other hand, let $\lambda \in \operatorname{ess\,ran} f$. For any $m \in \mathbb{N}$ denote $\widetilde{S}_m := \big\{x : |f(x) - \lambda| < 2^{-m}\big\}$, then $\mu(\widetilde{S}_m) > 0$, and by the assumptions there is a subset $S_m \subset \widetilde{S}_m$ with $0 < \mu(S_m) < \infty$. If ϕ_m is the indicator function of S_m, then one has

$$\big\|(M_f - \lambda)\phi_m\big\|^2 = \int_{S_m} \big|f(x) - \lambda\big|^2 \big|\phi_m(x)\big|^2 dx \le 2^{-2m}\big\|\phi_m\big\|^2.$$

Hence, (ϕ_m) is a Weyl sequence for λ, and $\lambda \in \operatorname{spec} M_f$ (Prop. 2.5).

To prove the second assertion we remark that the condition $\lambda \in \operatorname{spec}_p M_f$ is equivalent to the existence of $\phi \in L^2(X,\mu)$ such that $\big(f(x) - \lambda\big)\phi(x) = 0$ for a.e. x. This means that $\phi(x) = 0$ for a.e. x with $f(x) \ne \lambda$. If $\mu\{x : f(x) = \lambda\} = 0$, then $\phi = 0$ a.e., and $\lambda \notin \operatorname{spec}_p M_f$. On the other hand, if $\mu\{x : f(x) = \lambda\} > 0$, one can choose a subset $\Sigma \subset \{x : f(x) = \lambda\}$ of a strictly positive finite measure, then the indicator function ϕ of Σ is an eigenfunction of M_f for the eigenvalue λ. $\square$

It can be shown that the spectrum is invariant under unitary transformations (exercise):

Proposition 2.8 (Spectrum and unitary equivalence). *Let two operators A an B be unitarily equivalent, then* $\operatorname{spec} A = \operatorname{spec} B$ *and* $\operatorname{spec_p} A = \operatorname{spec_p} B$. *The resolvents of A and B are also unitarily equivalent.*

Example 2.9 (Spectrum of the free Laplacian). Let T be the free Laplacian in $\mathbb{R}^d$ (see Definition 1.47). As seen above, T is unitarily equivalent to the multiplication operator $f(p) \mapsto |p|^2 f(p)$ in $L^2(\mathbb{R}^d)$, i.e. to the multiplication operator M_h in $L^2(\mathbb{R}^d)$ with $h : p \mapsto |p|^2$. One has $\operatorname{ess\,ran} h = [0,\infty)$, while the set $\{p \in \mathbb{R}^d : h(p) = \lambda\}$ has zero measure for any λ. By combining Propositions 2.7 and 2.8 we obtain

$$\operatorname{spec} T = [0,+\infty), \quad \operatorname{spec_p} T = \emptyset.$$

Example 2.10 (Spectrum of discrete multiplication operators). Take $\mathcal{H} = \ell^2(\mathbb{Z})$. Consider an aribtrary function $a : \mathbb{Z} \to \mathbb{C}$, $n \mapsto a_n$, and the associated operator T:

$$D(T) = \big\{(\xi_n) \in \ell^2(\mathbb{Z}) : \ (a_n\xi_n) \in \ell^2(\mathbb{Z})\big\}, \quad (T\xi)_n = a_n\xi_n.$$

Remark that $\ell^2(\mathbb{Z})$ can be viewed as $L^2(\mathbb{Z}, \mu)$ with the discrete measure μ given by $\mu(X) = \#X$. Then $T = M_a$ (the operator of multiplication by a), and one easily checks that

$$\operatorname{spec} T := \overline{\{a_n : \ n \in \mathbb{Z}\}} \equiv \operatorname{ess\,ran} a, \quad \operatorname{spec_p} T := \{a_n : \ n \in \mathbb{Z}\}.$$

It is clear that all constructions also hold if one replaces $\mathbb{Z}$ by $\mathbb{N}$.

Example 2.11 (Spectrum and orthonormal eigenbases). Let T be a self-adjoint operator in $\mathcal{H}$ such that there exists an orthonormal basis of eigenfunctions $(\varphi_n)_{n\in\mathbb{N}}$, with $T\varphi_n = a_n\varphi_n$.

Consider the unitary map $U : \mathcal{H} \to \ell^2(\mathbb{N})$, $(Uf)(n) = \langle\varphi_n, f\rangle$, then the unitarily equivalent operator $A := UTU^{-1}$ is the operator of multiplication by (a_n) as in Example 2.10. It follows that the spectrum is given by the same expressions, i.e.

$$\operatorname{spec} T := \overline{\{a_n : \ n \in \mathbb{N}\}} \equiv \operatorname{ess\,ran} a, \quad \operatorname{spec_p} T := \{a_n : \ n \in \mathbb{N}\}.$$

Example 2.12 (Harmonic oscillator). Let $\mathcal{H} = L^2(\mathbb{R})$. Consider the operator $T = -d^2/dx^2 + x^2$ defined on $C_c^\infty(\mathbb{R})$. We have seen (Exercise 4) that T is essentially self-adjoint and that its closure $S := \overline{T}$ has an orthonormal basis of eigenfunctions with simple eigenvalues $(2n-1)$, $n \in \mathbb{N}$. Then the constructions of Example 2.11 show that

$$\operatorname{spec} S = \operatorname{spec_p} S = \{2n - 1 : \ n \in \mathbb{N}\}.$$

Example 2.13 (Empty spectrum). Let T_0 be the linear operator in $\mathcal{H} := L^2(0,1)$ given as $T_0 f = f'$ on the domain $D(T_0) = \big\{f \in C^1\big([0,1]\big) : f(0) = 0\big\}$.

For any $z \in \mathbb{C}$ and any $g \in C^0([0,1])$ there exists a unique $f \in D(T_0)$ with $(T_0 - z)f = g$, i.e. f is the solution of the initial value problem

$$f' - zf = g, \quad f(0) = 0.$$

In fact, this f is explicitly given by

$$f(x) = \int_0^x e^{z(x-t)} g(t)\,\mathrm{d}t.$$

Now let $z \in \mathbb{C}$ and consider the linear operator $A_z : \mathcal{H} \to \mathcal{H}$,

$$A_z g(x) = \int_0^x e^{z(x-t)} g(t)\,\mathrm{d}t,$$

which is clearly continuous. It is also injective, which can be seen as follows. Let $A_z g = 0$, then the function $\widetilde{g} : t \mapsto e^{-zt} g(t)$ is orthogonal to the indicator functions of $(0,x)$ for all x and, as a consequence, to the indicator functions of all subintervals of $(0,1)$. Hence $\widetilde{g} = 0$ a.e., and then $\widetilde{g} = 0$ a.e.

The above constructions show that:

- $T_0 : D(T_0) \to C^0([0,1])$ is bijective,
- $A_z : C^0([0,1]) \to D(T_0)$ is bijective,
- $(T_0 - z)A_z g = g$ for any $g \in C^0([0,1])$,
- $A_z(T_0 - z)f = f$ for any $f \in D(T_0)$.

It follows that $T_0 - z = A_z^{-1}$ on $D(T_0)$, and by taking the closure we see that the operator $T := \overline{T}_0$ is defined on $\operatorname{ran} A_z$ and satisfies $T - z = A_z^{-1}$ for any $z \in \mathbb{C}$, i.e. $(T - z)^{-1} = A_z \in \mathcal{B}(\mathcal{H})$. Therefore, $\operatorname{res} T = \mathbb{C}$ and $\operatorname{spec} T = \emptyset$.

Example 2.14 (Empty resolvent set). Let us modify the previous example. Take $\mathcal{H} = L^2(0,1)$ and consider the operator T acting as $Tf = f'$ on the domain $D(T) = H^1(0,1)$. Now for any $z \in \mathbb{C}$ we see that the function $\phi_z(x) = e^{zx}$ belongs to $D(T)$ and satisfies $(T - z)\phi_z = 0$. Therefore, $\operatorname{spec}_p T = \operatorname{spec} T = \mathbb{C}$.

As we can see in the two last examples, for general operators one cannot say much on the location of the spectrum. In what follows we will study mostly self-adjoint operators, whose spectral theory is understood much better.

2.2 Spectra of self-adjoint operators

The following proposition is certainly already known, but we include the proof for completeness:

Proposition 2.15 (Spectrum of a continuous operator). *Let $T \in \mathcal{B}(\mathcal{H})$, then* $\operatorname{spec} T$ *is a* <u>*non-empty*</u> *subset of* $\big\{z \in \mathbb{C} : |z| \le \|T\|\big\}$.

Proof. Let $z \in \mathbb{C}$ with $|z| > \|T\|$. Represent $T - z = -z(1 - T/z)$. As $\|T/z\| < 1$, the inverse to $T - z$ is defined by the series,

$$(T-z)^{-1} = -z \sum_{n=0}^{\infty} \Big(\frac{T}{z}\Big)^n \equiv -\sum_{n=0}^{\infty} T^n z^{1-n}. \tag{2.6}$$

and $z \in \operatorname{res} T$. This implies the sought inclusion.

Let us show that the spectrum is non-empty. Assume that it is not the case. Then for any $f, g \in \mathcal{H}$ the function $\mathbb{C} \ni z \mapsto F(z) := \langle f, R_T(z) g\rangle \in \mathbb{C}$ is holomorphic in $\mathbb{C}$ by Theorem 2.4. On the other hand, it follows from the series representation (2.6) that for large z the norm of $R_T(z)$ tends to zero. It follows that $F(z) \to 0$ as $|z| \to \infty$ and that F is bounded. By Liouville's theorem, F is constant, and, moreover, $F(z) = \lim_{|z|\to+\infty} F(z) = 0$. Therefore, $\langle f, R_T(z) g\rangle = 0$ for all $z \in \mathbb{C}$ and $f, g \in \mathcal{H}$, which means that $R_T(z) = 0$. This contradicts the definition of the resolvent and shows that the spectrum of T must be non-empty. □

We will also need the following relations betwen operators and their adjoints:

Proposition 2.16. *Let T be a densely defined linear operator and $z \in \mathbb{C}$, then*

$$\ker(T^* - \overline{z}) = \operatorname{ran}(T - z)^{\perp}, \tag{2.7}$$

$$\overline{\operatorname{ran}(T - z)} = \ker(T^* - \overline{z})^{\perp}. \tag{2.8}$$

Proof. Note that the second equality can be obtained from the first one by taking the orthogonal complement in the both parts. Let us prove the first equality. As $D(T)$ is dense, the condition $f \in \ker(T^* - \overline{z})$ is equivalent to $\langle (T^* - \overline{z}) f, g\rangle = 0$ for all $g \in D(T)$, which can be also rewritten as $\langle T^* f, g\rangle = z\langle f, g\rangle$ for all $g \in D(T)$. By the definition of T^*, one has $\langle T^* f, g\rangle = \langle f, Tg\rangle$ and

$$\langle f, Tg\rangle - z\langle f, g\rangle \equiv \langle f, (T-z)g\rangle = 0 \text{ for all } g \in D(T),$$

which is equivalent to $f \perp \operatorname{ran}(T - z)$. □

Now we pass to the discussion of the spectra of self-adjoint operators, and we are going to show the following fundamental fact:

Theorem 2.17 (Spectrum of a self-adjoint operator). *The spectrum of a self-adjoint operator in a Hilbert space is a non-empty closed subset of the real line.*

Proof. We have already shown that the spectrum is closed (Theorem 2.4). In Lemma 2.18 below we prove that the spectrum is real, and in Lemma 2.22 we prove that it is non-empty. □

The proof will be decomposed in several steps:

Lemma 2.18 (Spectrum of a self-adjoint operator is real). *Let T be a self-adjoint operator in a Hilbert space $\mathcal{H}$, then* $\operatorname{spec} T \subset \mathbb{R}$, *and for any $z \in \mathbb{C} \setminus \mathbb{R}$ there holds*

$$\big\|(T-z)^{-1}\big\| \le \frac{1}{|\Im z|}. \tag{2.9}$$

Proof. Let $z \in \mathbb{C} \setminus \mathbb{R}$ and $u \in D(T)$, then

$$\langle u, (T-z)u\rangle = \langle u, Tu\rangle - \Re z\langle u, u\rangle - i\Im z\langle u, u\rangle.$$

As T is self-adjoint, the number $\langle u, Tu\rangle$ is real. Therefore,

$$|\Im z|\,\|u\|^2 = |\Im\langle u, (T-z)u\rangle| \le \big|\langle u, (T-z)u\rangle\big| \le \big\|(T-z)u\big\| \cdot \|u\|,$$

which shows that

$$\big\|(T-z)u\big\| \ge |\Im z| \cdot \|u\|. \tag{2.10}$$

It follows from here that $\operatorname{ran}(T-z)$ is closed, that $\ker(T-z) = \{0\}$. Proposition 2.16 implies $\operatorname{ran}(T-z) = \mathcal{H}$. Therefore, $(T-z)^{-1} \in \mathcal{B}(\mathcal{H})$, and the estimate (2.9) follows from (2.10). □

Lemma 2.19 (Spectral edges for continuous self-adjoint operators). *Let T be a continuous self-adjoint operator in a Hilbert $\mathcal{H}$. Denote*

$$m = m(T) = \inf_{u \ne 0} \frac{\langle u, Tu\rangle}{\langle u, u\rangle}, \quad M = M(T) = \sup_{u \ne 0} \frac{\langle u, Tu\rangle}{\langle u, u\rangle},$$

then $\operatorname{spec} T \subset [m, M]$ *and* $\{m, M\} \subset \operatorname{spec} T$*.*

Proof. We proved already that $\operatorname{spec} T \subset \mathbb{R}$. For $\lambda \in (M, +\infty)$ we have

$$\|u\| \cdot \|(T-\lambda)u\| \ge \big|\langle u, (\lambda - T)u\rangle\big| \ge (\lambda - M)\|u\|^2,$$

i.e. $\|(T-\lambda)u\| \ge (\lambda - M)^{-1}\|u\|$. It follows that $\ker(T-\lambda) = \{0\}$, that $\operatorname{ran}(T-\lambda)$ is closed, and due to $\operatorname{ran}(T-\lambda)^\perp = \ker(T-\lambda)$, is dense. Hence, $(T-\lambda)^{-1} \in \mathcal{B}(\mathcal{H})$. In the same way one shows that $\operatorname{spec} T \cap (-\infty, m) = \emptyset$.

Let us show that $M \in \operatorname{spec} T$ (for m the proof is similar). Using the Cauchy-Schwarz inequality for the semi-scalar product $(u, v) \mapsto \langle u, (M-T)v\rangle$ we obtain

$$\big|\langle u, (M-T)v\rangle\big|^2 \le \langle u, (M-T)u\rangle \cdot \langle v, (M-T)v\rangle.$$

Taking the supremum over all $u \in \mathcal{H}$ with $\|u\| \le 1$ we arrive at

$$\big\|(M-T)v\big\|^2 \le \|M-T\| \cdot \big\langle v, (M-T)v\big\rangle.$$

By assumption, for some (u_n) with $\|u_n\| = 1$ one has $\langle u_n, Tu_n\rangle \to M = M\langle u, u\rangle$ as $n \to \infty$. By the above inequality we have then $(M-T)u_n \to 0$, i.e. (u_n) is a Weyl sequence for M, and $M \in \operatorname{spec} T$ (Prop. 2.5). □

Lemma 2.20. *If T is a continuous self-adjoint operator with* $\operatorname{spec} T = \{0\}$, *then* $T = 0$.

Proof. By Lemma 2.19 we have $m(T) = M(T) = 0$. This means that $\langle x, Tx\rangle = 0$ for all $x \in \mathcal{H}$, and the polar identity shows that $\langle x, Ty\rangle = 0$ for all $x, y \in \mathcal{H}$. □

Remark 2.21. If T is not self-adjoint, then $\operatorname{spec} T = \{0\}$ *does not imply* that $T = 0$. A simple example is given by

$$\mathcal{H} := \mathbb{C}^2, \quad T = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Lemma 2.22. *The spectrum of a self-adjoint operator in a Hilbert space is non-empty.*

Proof. Let T be a self-adjoint operator in a Hilbert space $\mathcal{H}$. By contradiction, assume that $\operatorname{spec} T = \emptyset$. Then $T^{-1} \in \mathcal{B}(\mathcal{H})$. Let $\lambda \in \mathbb{C} \setminus \{0\}$, then $1/\lambda \in \operatorname{res} T$, the operator

$$L_\lambda := -\frac{T}{\lambda}\Big(T - \frac{1}{\lambda}\Big)^{-1} \equiv -\frac{1}{\lambda} - \frac{1}{\lambda^2}\Big(T - \frac{1}{\lambda}\Big)^{-1}$$

is continuous. Furthermore,

$$(T^{-1} - \lambda) = -\lambda\Big(T - \frac{1}{\lambda}\Big)T^{-1} = -\lambda T^{-1}\Big(T - \frac{1}{\lambda}\Big),$$

which shows that $(T^{-1} - \lambda)L_\lambda = \mathrm{Id}_{\mathcal{H}} = L_\lambda(T^{-1} - \lambda)$. Therefore, $\lambda \in \operatorname{res}(T^{-1})$.

It follows that all $\lambda \neq 0$ belong to $\operatorname{res}(T^{-1})$, hence, $\operatorname{spec}(T^{-1}) \subset \{0\}$. As T^{-1} is bounded, its spectrum is non-empty, hence, $\operatorname{spec} T^{-1} = \{0\}$. On the other hand, T^{-1} is self-adjoint by Proposition 1.21, and $T^{-1} = 0$ by Corollary 2.20, which contradicts the definition of the inverse operator. □

For further references we also mention the following result (the proof is a minor modification of the proof of Theorem 1.61 and will be discussed as an exercise):

Theorem 2.23 (Spectrum of a semibounded self-adjoint operator). *Let T be self-adjoint lower semibounded linear operator. Denote*

$$m := \inf_{x \in D(T),\, x \neq 0} \frac{\langle x, Tx\rangle}{\langle x, x\rangle}.$$

Then $m = \inf \operatorname{spec} T$.

Proof. By construction $T - m \geq 0$. Let $\lambda \in (-\infty, m)$, then $T - \lambda = (T - m) + (m - \lambda) \geq (m - \lambda)$, and for any $x \in D(T)$ one has $\big\langle x, (T - \lambda)x\big\rangle \geq (m - \lambda)\|v\|^2$, i.e. $\|(T - \lambda)x\| \geq (m - \lambda)\|x\|$. This shows that $T - \lambda$ is injective and that $\operatorname{ran}(T - \lambda)$ is closed, then $\operatorname{ran}(T - \lambda) = \mathcal{H}$ by Prop. 2.16, and then $\lambda \in \operatorname{res} T$ by Prop. 2.3.

Assume that $m \notin \operatorname{spec} T$, then $(T - m)^{-1}$ is a well-defined continuous operator. From the definition of m it follows that there exist $(x_n) \subset D(T)$ with $\|x_n\| = 1$ and

$\langle x_n, (T-m)x_n\rangle \to 0$. As $(u,v) \mapsto \langle u, (T-m)v\rangle$ is a semi-scalar product on $D(T)$, one has the Cauchy-Schwarz inequality

$$\big|\langle u, (T-m)v\rangle\big|^2 \le \big\langle u, (T-m)u\big\rangle\big\langle v, (T-m)v\big\rangle.$$

Now use this inequality for $u := x_n$ and $v := (T-m)^{-1}x_n$:

$$\begin{aligned}
1 = \|x_n\|^2 = \langle x_n, x_n\rangle &= \big\langle x_n, (T-m)(T-m)^{-1}x_n\big\rangle \\
&\le \langle x_n, (T-m)x_n\rangle\langle (T_m)^{-1}x_n, (T-m)(T-m)^{-1}x_n\rangle \\
&= \langle x_n, (T-m)x_n\rangle\langle (T-m)^{-1}x_n, x_n\rangle \\
&\le \langle x_n, (T-m)x_n\rangle\|(T-m)^{-1}\| \overset{n\to+\infty}{\longrightarrow} 0,
\end{aligned}$$

which is a contradiction. □

2.3 Compact operators

The present section contains a lot of repetititons from earlier courses, but they are important for what follows.

Let $\mathcal{H}$ and $\mathcal{H}'$ be Hilbert spaces. A linear operator $T : \mathcal{H} \to \mathcal{H}'$ is called *compact*, if the image of the unit ball in $\mathcal{H}$ is relatively compact in $\mathcal{H}'$ (remark that the relative compactness is equivalent to the existence of a finite ε-net for any $\varepsilon > 0$, as Hilbert spaces are complete metric spaces). We denote by $\mathcal{K}(\mathcal{H}, \mathcal{H}')$ the set of all such operators. The definition can also be reformulated as follows: an operator $T : \mathcal{H} \to \mathcal{H}'$ is compact if and only if any bounded sequence $(x_n) \subset \mathcal{H}$ has a subsequence (x_{n_k}) such that Tx_{n_k} converges in $\mathcal{H}'$. We recall that:

- any compact operator is continuous,
- any continuous operator having a finite-dimensional range is compact,
- the norm limit of a sequence of compact operators is again a compact operator,
- the adjoint of a compact operator is compact,
- the composition of a continuous operator with a compact one (in any order) is again a compact operator.

Recall the following fundamental result, which is based on Fredholm's alternative and is proved in the functional analysis course:

Theorem 2.24 (Spectrum of a compact operator). *Let T be a compact linear operator in an infinite-dimensional Hilbert space, then*

(a) $0 \in \operatorname{spec} T$,

(b) $\operatorname{spec} T \setminus \{0\} = \operatorname{spec}_p T \setminus \{0\}$,

(c) we are in one and only one of the following situations:

- $\operatorname{spec} T \setminus \{0\} = \emptyset$,
- $\operatorname{spec} T \setminus \{0\}$ *is a finite set,*
- $\operatorname{spec} T \setminus \{0\}$ *is a sequence convergent to* 0.

(d) Each $\lambda \in \operatorname{spec} T \setminus \{0\}$ *is isolated (i.e. has a neighborhood containing no other values of the spectrum), and* $\dim \ker(T - \lambda) < \infty$.

The result has the following important corollary:

Theorem 2.25 (Spectrum of compact self-adjoint operators). *Let T be a compact self-adjoint operator in a Hilbert space $\mathcal{H}$, then there exist an orthonormal basis consisting of eigenvectors of T, and the respective eigenvalues form a real sequence convergent to* 0.

Proof. Let $(\lambda_n)_{n\geq 1}$ be the distinct non-zero eigenvalues of T. As T is self-adjoint, these eigenvalues are real. Set $\lambda_0 = 0$, and for $n \geq 0$ denote $E_n := \ker(T - \lambda_n)$. If $n \neq m$ and $u \in E_n$, $v \in E_m$, then

$$\lambda_n \langle u, v\rangle = \langle Tu, v\rangle = \langle u, Tv\rangle = \lambda_m \langle u, v\rangle \;\Rightarrow\; (\lambda_n - \lambda_m)\langle u, v\rangle = 0 \;\Rightarrow\; u \perp v = 0,$$

which shows that $E_n \perp E_m$ for $n \neq m$. Denote by F the linear hull of $\cup_{n\geq 0} E_n$. We are going to show that F is dense in $\mathcal{H}$.

Clearly, we have $T(F) \subset F$. Due to the self-adjointness of T we also have $T(F^\perp) \subset F^\perp$. Denote by $\widetilde{T}$ the restriction of T to $F^\perp$, then $\widetilde{T}$ is self-adjoint, and its spectrum equals $\{0\}$, so $\widetilde{T} = 0$. But this means that $F^\perp \subset \ker T = E_0 \subset F$ and shows that $F^\perp = \{0\}$. Therefore F is dense in $\mathcal{H}$.

We now choose an orthonormal basis in each subspace E_n and obtain an orthonormal basis in the whole space $\mathcal{H}$. □

We now consider an important subclass of compact operators:

Definition 2.26 (Hilbert-Schmidt operator). A continuous linear operator T in a Hilbert space $\mathcal{H}$ is *Hilbert-Schmidt* if for some orthonormal basis (e_n) of $\mathcal{H}$ the sum

$$\|T\|_2^2 := \sum_n \|Te_n\|^2 \tag{2.11}$$

is finite.

Theorem 2.27 (Hilbert-Schmidt norm and compactness). *Let T be a Hilbert-Schmidt operator in a Hilbert space $\mathcal{H}$, then:*

(a) the quantity $\|T\|_2$ (called the Hilbert-Schmidt norm of T) does not depend on the choice of the orthonormal basis,

(b) T^* is also Hilbert-Schmidt with $\|T^*\|_2 = \|T\|_2$,

(c) $\|T\| \leq \|T\|_2$,

(d) T and T^* are compact operators.

Proof. Let (e_n) and (f_n) be two orthonormal bases. Using the Parseval identity we have

$$\sum_n \|Te_n\|^2 = \sum_n \sum_m \big|\langle f_m, Te_n\rangle\big|^2 = \sum_m \sum_n \big|\langle T^* f_m, e_n\rangle\big|^2 = \sum_m \|T^* f_m\|^2.$$

This shows that the expression (2.11) is independent of the choice of (e_n) and that $\|T^*\|_2 = \|T\|_2$, which proves (a) and (b).

To show (c), let $x \in \mathcal{H}$ with $x_n := \langle e_n, x\rangle$, then using the Cauchy-Schwarz inequality in ℓ^2 we estimate

$$\|Tx\|^2 = \Big\| \sum_n x_n Te_n \Big\|^2 \le \Big(\sum_n |x_n| \|Te_n\| \Big)^2 \le \sum_n |x_n|^2 \sum_n \|Te_n\|^2 = \|T\|_2^2 \|x\|^2.$$

In order to prove (d) remark first that for any $x \in \mathcal{H}$ one has

$$x = \sum_{n=1}^{\infty} \langle e_n, x\rangle e_n, \quad Tx = \sum_{n=1}^{\infty} \langle e_n, x\rangle Te_n.$$

For $N \in \mathbb{N}$ introduce the operators T_N by

$$T_N : x \mapsto \sum_{n=1}^{N} \langle e_n, x\rangle Te_n.$$

One has

$$\|T - T_N\|^2 \le \|T - T_N\|_2^2 = \sum_{n \ge N+1} \|Te_n\|^2 \xrightarrow{N\to\infty} 0,$$

and then T is compact being the norm-limit of the finite-dimensional operators T_N. The compactness of T^* follows from (b). $\square$

If we use (2.11) for an orthonormal basis consisting of eigenvectors of T, we obtain:

Corollary 2.28. *Let T be a self-adjoint Hilbert-Schmidt operator with eigenvalues (λ_n), then*

$$\|T\|_2^2 = \sum_n \lambda_n^2.$$

The suprising thing is that a Hilbert-Schmidt operator in an L^2-space is always an integral operator. We only consider $L^2(\Omega)$ with $\Omega \subset \mathbb{R}^d$ being an open set: the extension to general measure spaces is straightforward. We will need the following fact:

Lemma 2.29. *If (e_n) is an orthonormal basis in $L^2(\Omega)$ and (e'_m) is an orthonormal basis in $L^2(\Omega')$, then the functions $f_{n,m} : (x, x') \mapsto e_n(x)e'_m(x')$ form an orthonormal basis in $L^2(\Omega \times \Omega')$.*

Proof. It follows by Fubini's theorem that $f_{n,m}$ form an orthonormal family in $L^2(\Omega \times \Omega')$, and we need to show that they span the whole Hilbert space. Let $g \perp f_{n,m}$ for all (n, m),

$$\int_\Omega \overline{e_n(x)} \int_{\Omega'} \overline{e'_m(x')} g(x, x') \,\mathrm{d}x' \,\mathrm{d}x = 0 \text{ for all } (n, m). \tag{2.12}$$

One easily shows (using the Cauchy-Schwarz inequality) that for any $m \in \mathbb{N}$ the function

$$h_m : x \mapsto \int_{\Omega'} \overline{e'_m(x')} g(x, x') \,\mathrm{d}x'$$

belongs to $L^2(\Omega)$, and (2.12) implies $h_m = 0$ a.e. For a.e. $x \in \Omega$ the function $H_x : x \mapsto g(x, x')$ belongs to $L^2(\Omega)$, and the condition $h_m = 0$ a.e. for all m implies that $H_x(x') = 0$ for a.e. (x, x'), i.e. $g = 0$. $\square$

Theorem 2.30 (Integral Hilbert-Schmidt operators). *Let $\mathcal{H} = L^2(\Omega)$. A linear operator T in $\mathcal{H} = L^2(\Omega)$ is Hilbert-Schmidt if and only if T is an integral operator,*

$$Tu(x) = \int_\Omega K(x, y)u(y) \,\mathrm{d}y, \tag{2.13}$$

with integral kernel $K \in L^2(\Omega \times \Omega)$, and in that case $\|T_K\|_2 = \|K\|_{L^2(\Omega\times\Omega)}$.

Proof. Let $K \in L^2(\Omega \times \Omega)$ and T be defined as is (2.13). Remark that T is continuous: using the Cauchy-Schwarz inequality one obtains

$$\begin{aligned}
\|Tu\|^2 &= \int_\Omega \Big| \int_\Omega K(x, y)u(y) \,\mathrm{d}y \Big|^2 \,\mathrm{d}x \\
&\leq \int_\Omega \bigg[\int_\Omega |K(x, y)|^2 \,\mathrm{d}y \int_\Omega \big|u(y)\big|^2 \,\mathrm{d}y \bigg] \,\mathrm{d}x = \|K\|^2_{L^2(\Omega\times\Omega)} \|u\|^2.
\end{aligned}$$

Let (e_n) be an orthonormal basis in $\mathcal{H}$, then the functions $e_{m,n}(x, y) = e_m(x)\overline{e_n(y)}$

form an orthonormal basis in $L^2(\Omega \times \Omega)$, see Lemma 2.29. There holds

$$\begin{aligned}
\|T\|_2^2 &= \sum_n \|Te_n\|^2 = \sum_{m,n} \big|\langle e_m, Te_n\rangle\big|^2 \\
&= \sum_{m,n} \Big| \int_\Omega \overline{e_m(x)} \Big(\int_\Omega K(x,y) e_n(y)\,\mathrm{d}y \Big)\,\mathrm{d}x \Big|^2 \\
&= \sum_{m,n} \Big| \int_\Omega \int_\Omega \overline{e_m(x)} e_n(y) K(x,y)\,\mathrm{d}x\,\mathrm{d}y \Big|^2 = \\
&= \sum_{m,n} \big|\langle e_{m,n}, K\rangle\big|^2 = \|K\|^2_{L^2(\Omega\times\Omega)},
\end{aligned}$$

which shows that T is Hilbert-Schmidt.

Now let T be a Hilbert-Schmidt operator in $\mathcal{H}$ and $u, v \in \mathcal{H}$, then

$$\begin{aligned}
\langle u, Tv\rangle &= \Big\langle \sum_m \langle e_m, u\rangle e_m, T\Big(\sum_n \langle e_n, v\rangle e_n \Big\rangle = \sum_{m,n} \langle e_n, v\rangle \langle e_m, Te_n\rangle \langle u, e_m\rangle \\
&= \sum_{m,n} \int_\Omega \int_\Omega \overline{e_n(y)} \langle e_m, Te_n\rangle e_m(x) \overline{u(x)}\, v(y)\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(x) \\
&= \sum_{m,n} \int_{\Omega\times\Omega} \langle e_m, Te_n\rangle e_{m,n}(x,y) \overline{u(x)}\, v(y)\,\mathrm{d}x\,\mathrm{d}y. \qquad (2.14)
\end{aligned}$$

Take

$$K(x,y) = \sum_{m,n} \overline{e_n(y)} \langle e_m, Te_n\rangle e_m(x) = \sum_{m,n} \langle e_m, Te_n\rangle e_{m,n}(x,y).$$

One has

$$\sum_{m,n} |\langle e_m, Te_n\rangle|^2 = \sum_n \sum_m |\langle e_m, Te_n\rangle|^2 = \sum_n \|Te_n\|^2 = \|T\|_2^2 < \infty,$$

which shows that $K \in L^2(\Omega \times \Omega)$. Then one can interchange the sum and the integral in (2.14) and one arrives at

$$\langle u, Tv\rangle = \int_{\Omega\times\Omega} K(x,y) \overline{u(x)}\, v(y)\,\mathrm{d}x\,\mathrm{d}y \text{ for all } u, v \in L^2(\Omega),$$

which proves the representation (2.13). □

Using Corollary 2.28 one immediately shows:

Corollary 2.31 (Trace formula for integral Hilbert-Schmidt operators). *Let K be the integral kernel of a self-adjoint Hilbert-Schmidt operator T with eigenvalues (λ_n) in $L^2(\Omega)$, then*

$$\int_\Omega \int_\Omega |K(x,y)|^2\,\mathrm{d}x\,\mathrm{d}y = \sum_n \lambda_n^2.$$

2.4 Operators with compact resolvents

The above can be used for a discussion of a class of unbounded operators. Namely, one says that an operator T in $\mathcal{H}$ *has compact resolvent* if $\operatorname{res} T \neq \emptyset$ and for some (and then for all) $z \in \operatorname{res} T$ the resolvent $(T-z)^{-1}$ is a compact operator.

Theorem 2.32 (Spectra of semibounded operators with compact resolvents). *Let T be a semibounded from below self-adjoint operator with compact resolvent in an infinite-dimensional Hilbert space $\mathcal{H}$, then:*

- $\operatorname{spec} T = \operatorname{spec}_p T$,
- *for each $\lambda \in \operatorname{spec} T$ there holds $\dim\ker(T-\lambda) < \infty$.*
- *the eigenvalues of T form a sequence converging to $+\infty$.*

Proof. Let $T \geq -c$, then $-(c+1) \in \operatorname{res} T$ (Theorem 2.23), and $(T+c+1)^{-1}$ is a bounded self-adjoint operator which is compact by assumption. Moreover, this operator is non-negative: for any $u \in \mathcal{H}$ denote $v := (T+c+1)^{-1} \in D(T)$, then

$$\langle u, (T+c+1)^{-1}u\rangle = \langle (T+c+1)v, v\rangle \geq \|v\|^2 \geq 0.$$

By Theorem 2.25, there exists an orthonormal basis $(e_n)_{n\in\mathbb{N}}$ of $\mathcal{H}$ such that each e_n is an eigenvector of $(T+c+1)^{-1}$: $(T+c+1)^{-1}e_n = \lambda_n e_n$, where $\lambda_n > 0$ form a sequence converging to 0. We then have $(T+c+1)e_n = \lambda_n^{-1}e_n$, i.e. each e_n is an eigenvector of T with eigenvalue $\mu_n := \lambda_n^{-1} - c - 1$, and the multiplicity of this eigenvalue is the same as that of λ_n as an eigenvalue of $(T+c+1)^{-1}$, e.g. is finite. The operator T is then essentially self-adjoint on finite linear combinations of e_n (Exercise 3), and $\operatorname{spec} T = \overline{\{\mu_n : n \in \mathbb{N}\}}$. In our case $\mu_n \to +\infty$ as $n \to +\infty$ (due to $\lambda_n \to 0$), so the closure can be omitted. □

Now we would like to obtain a class of operators with compact resolvents.

Theorem 2.33 (Compact embeddings and compact resolvents). *Let T be a self-adjoint operator generated by a closed sesquilinear form t in $\mathcal{H}$. Assume that the Hilbert space $D(t)$, equiped with $\langle\cdot,\cdot\rangle_t$, is compactly embedded in $\mathcal{H}$, then T has compact resolvent.*

Proof. Without logg of generality assume $t \geq 1$, then $T \geq 1$ as well. Recall that $\langle u, v\rangle = t(u,v)$ and $\|u\|_t^2 = t(u,u) + (c+1)\|u\|_{\mathcal{H}}^2 \geq \|u\|_{\mathcal{H}}^2$. For any $u \in D(T)$ we have $|u\|_{\mathcal{H}}\|u\|_t \leq \|u\|_t^2 = t(u,u) = \big\langle u, Tu\big\rangle_{\mathcal{H}} \leq \|u\|_{\mathcal{H}}\|Tu\|_{\mathcal{H}}$, i.e. $\|u\|_t \leq \|Tu\|_{\mathcal{H}}$. It follows that $\|T^{-1}v\|_t \leq \|v\|_{\mathcal{H}}$ for all $v \in \mathcal{H}$, which means $T^{-1} \in \mathcal{B}\big(\mathcal{H}, D(t)\big)$.

Now let $j : D(t) \hookrightarrow \mathcal{H}$ be the embedding, which is compact by assumption, then $T^{-1} = jL$, where $L : \mathcal{H} \ni v \mapsto T^{-1}v \in D(t)$. Hence, $T^{-1} : \mathcal{H} \to \mathcal{H}$ is compact as the composition of a bounded operator and a compact one. □

In order to look at concrete examples we need the following classical criterion of compactness in $L^2(\mathbb{R}^d)$:

Theorem 2.34 (Riesz-Kolmogorov criterion). *A subset $\mathcal{F} \subset L^2(\mathbb{R}^d)$ is relatively compact in $L^2(\mathbb{R}^d)$ if and only if the following conditions are satisfied:*

(a) $\mathcal{F}$ is bounded,

(b) for any $\varepsilon > 0$ one can find $R > 0$ such that

$$\int_{|x|>R} |u(x)|^2 \,\mathrm{d}x < \varepsilon^2 \textit{ for all } u \in \mathcal{F},$$

(c) for any $\varepsilon > 0$ one can find $r > 0$ such that

$$\int_{\mathbb{R}^d} \big|u(x+h) - u(x)\big|^2 \,\mathrm{d}x < \varepsilon^2 \textit{ for all } u \in \mathcal{F} \textit{ and } h \in \mathbb{R}^d \textit{ with } |h| < r.$$

Proof. We will only prove the "if" direction (only this part is of interest for the subsequent applications).

Let $\mathcal{F} \subset L^2(\mathbb{R}^d)$ satisfy (a–c). Let $\varepsilon > 0$ and choose R as in (b) and r as in (c). Consider the open hypercube $Q := \big(-\frac{r}{2}, \frac{r}{2}\big)^d$ and let $Q_1, \dots, Q_N \subset \mathbb{R}^d$ be suitably chosen translations of Q such that:

$$Q_j \text{ are mutually disjoint}, \qquad \overline{B}_R(0) \subset S := \overline{Q}_1 \cup \dots \cup \overline{Q}_N,$$

then it follows from (b) that

$$\int_{S^\complement} |u(x)|^2 \,\mathrm{d}x < \varepsilon^2 \text{ for all } u \in \mathcal{F}. \tag{2.15}$$

Denote by W the subspace spanned by the characteristic functions 1_{Q_j} and let P be the orthogonal projector on W in $L^2(\mathbb{R}^d)$,

$$Pu(x) = \begin{cases} \dfrac{1}{|Q_j|} \displaystyle\int_{Q_j} u(x)\,\mathrm{d}x, & x \in Q_j, \\ 0, & \text{otherwise.} \end{cases}$$

Let $u \in \mathcal{F}$, then

$$\begin{aligned} \|u - Pu\|^2 &= \int_{S^\complement} |u - Pu|^2 \,\mathrm{d}x + \int_S |u - Pu|^2 \,\mathrm{d}x \\ &= \int_{S^\complement} |u|^2 \,\mathrm{d}x + \int_S |u - Pu|^2 \,\mathrm{d}x \\ \text{use (2.15)} \quad &\le \varepsilon^2 + \sum_{j=1}^N \int_{Q_j} \big|u(x) - Pu(x)\big|^2 \,\mathrm{d}x \\ \text{(Definition of } P) \quad &= \varepsilon^2 + \sum_{j=1}^N \int_{Q_j} \Big|\frac{1}{|Q_j|} \int_{Q_j} \big(u(x) - u(z)\big)\,\mathrm{d}z\Big|^2 \,\mathrm{d}x. \end{aligned} \tag{2.16}$$

Using the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}
\Big|\frac{1}{|Q_j|}\int_{Q_j}\big(u(x)-u(z)\big)\,\mathrm{d}z\Big|^2 &\le \frac{1}{|Q_j|^2}\int_{Q_j}1^2\,\mathrm{d}z\int_{Q_j}\big|u(x)-u(z)\big|^2\,\mathrm{d}z\\
&= \frac{1}{|Q_j|}\int_{Q_j}\big|u(x)-u(z)\big|^2\,\mathrm{d}z\\
\text{(one has } z-x\in 2Q \text{ for } x,z\in Q_j) \quad &\le \frac{1}{|Q_j|}\int_{2Q}\big|u(x)-u(x+y)\big|^2\,\mathrm{d}y.
\end{aligned}$$

The substitution of this inequality into (2.16) gives

$$\begin{aligned}
\|u-Pu\|^2 &\le \varepsilon^2+\sum_{j=1}^N\int_{Q_j}\frac{1}{|Q_j|}\int_{2Q}\big|u(x)-u(x+y)\big|^2\,\mathrm{d}y\,\mathrm{d}x\\
\text{(Fubini, } |Q_i|=|Q|) \quad &= \varepsilon^2+\frac{1}{|Q|}\sum_{j=1}^N\int_{2Q}\int_{Q_j}\big|u(x)-u(x+y)\big|^2\,\mathrm{d}x\,\mathrm{d}y\\
(Q_j \text{ mutually disjoint}) \quad &= \varepsilon^2+\frac{1}{|Q|}\int_{2Q}\int_{Q_1\cup\dots\cup Q_N}\big|u(x)-u(x+y)\big|^2\,\mathrm{d}x\,\mathrm{d}y\\
&\le \varepsilon^2+\frac{1}{|Q|}\int_{2Q}\int_{\mathbb{R}^d}\big|u(x)-u(x+y)\big|^2\,\mathrm{d}x\,\mathrm{d}y\\
\text{use (c)} \quad &\le \varepsilon^2+\frac{1}{|Q|}\int_{2Q}\varepsilon^2\,\mathrm{d}y=(1+2^d)\varepsilon^2.
\end{aligned}$$

Therefore, for any $u\in\mathcal{F}$ one has $\|u-Pu\|\le\sqrt{1+2^d}\varepsilon$, and the triangle inequality gives $\|u\|\le\sqrt{1+2^d}\varepsilon+\|Pu\|$. If $u,v\in\mathcal{F}$ with $\|Pu-Pv\|<\varepsilon$, then using the linearity of P we obtain $\|u-v\|\le(\sqrt{1+2^d}+1)\varepsilon$.

The subspace W is finite-dimensional, $\mathcal{F}$ is bounded by (a), and P is a bounded operator. It follows that $P(\mathcal{F})$ is a bounded subset of a finite-dimensional vector space, hence it is relatively compact. In particular, one can find a finite ε-net in $P(\mathcal{F})$, and its preimage is then a finite $(\sqrt{1+2^d}+1)\varepsilon$-net in $\mathcal{F}$. As ε is arbitrary, the relative compactness of $\mathcal{F}$ follows. □

We are going to apply the result to some differential operators, and the following observation will be useful:

Lemma 2.35. *The unit ball $\mathcal{F}$ in $H^1(\mathbb{R}^d)$ satisfies (c) in Theorem 2.34.*

Proof. For any $u\in C_c^\infty(\mathbb{R}^d)$ and $h\in\mathbb{R}^d$ one has

$$\begin{aligned}
\int_{\mathbb{R}^d}\big|u(x+h)-u(x)\big|^2\,\mathrm{d}x &= \int_{\mathbb{R}^d}\Big|\int_0^1\frac{d}{dt}u(x+th)\,\mathrm{d}t\Big|^2\,\mathrm{d}x\\
&= \int_{\mathbb{R}^d}\Big|\int_0^1 h\cdot\nabla u(x+th)\,\mathrm{d}t\Big|^2\,\mathrm{d}x \le h^2\int_{\mathbb{R}^d}\int_0^1\big|\nabla u(x+th)\big|^2\,\mathrm{d}t\ \mathrm{d}x\\
&\le h^2\int_0^1\int_{\mathbb{R}^d}\big|\nabla u(x+th)\big|^2\,\mathrm{d}x\,\mathrm{d}t = h^2\|\nabla u\|^2_{L^2},
\end{aligned}$$

which then extends by density to the whole of $H^1(\mathbb{R}^d)$. Hence, for any $u \in \mathcal{F}$ one has

$$\int_{\mathbb{R}^d} \big|u(x+h) - u(x)\big|^2 \,\mathrm{d}x \le h^2,$$

and (c) holds by taking $r := \varepsilon$ for any $\varepsilon > 0$. $\square$

Let $\Omega \subset \mathbb{R}^d$ be a non-empty open set. Recall that the associated Dirichlet and Neumann Laplacians T_D and T_N are defined as the self-adjoint operators in $L^2(\Omega)$ generated by the closed sesquilinear forms

$$\begin{aligned} t_D(u,u) &= \int_\Omega |\nabla u|^2 \,\mathrm{d}x, \quad D(t_D) = H_0^1(\Omega),\\ t_N(u,u) &= \int_\Omega |\nabla u|^2 \,\mathrm{d}x, \quad D(t_N) = H^1(\Omega). \end{aligned}$$

Recall that the scalar product $H_0^1(\Omega)$ is induced by the scalar product of $H^1(\Omega)$.

Proposition 2.36 (Compact embeddings of $\boldsymbol{H_0^1}$). *If Ω is bounded, then the embedding $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ is compact.*

Proof. We first denote $\widetilde{H}^1(\mathbb{R}^d) := \{u \in H^1(\mathbb{R}^d) : u = 0 \text{ a.e. outside } \Omega\} \subset H^1(\mathbb{R}^d)$. and let $\widetilde{\mathcal{F}}$ be the unit ball in $\widetilde{H}^1(\mathbb{R}^d)$ (with respect to the induced scalar product). We would like to show that $\widetilde{\mathcal{F}}$ is relatively compact in $L^2(\mathbb{R}^d)$, for which we check the conditions (a–c) in the Riesz-Kolmogorov theorem (Theorem 2.34). If $v \in \widetilde{\mathcal{F}}$, then $\|v\|_{L^2(\mathbb{R}^d)} = \|v\|_{H^1(\Omega)} \le 1$, which shows (a). If $R > 0$ with $\Omega \subset B_R(0)$, then

$$\int_{|x|\ge R} |v(x)|^2 \,\mathrm{d}x = 0 \text{ for any } v \in \widetilde{F},$$

which shows (b). Finally, $\widetilde{\mathcal{F}}$ is contained in the unit ball of $H^1(\mathbb{R}^d)$, and (c) holds by Lemma 2.35. It follows that the embedding $j : \widetilde{H}^1(\mathbb{R}^d) \hookrightarrow L^2(\mathbb{R}^d)$ is compact.

For $u \in L^2(\Omega)$ denote by $\widetilde{u}$ its extension by zero the whole of $\mathbb{R}^d$, then the linear map $k : H_0^1(\Omega) \ni u \mapsto \widetilde{u} \in \widetilde{H}^1(\mathbb{R}^d)$ is an isometry (Proposition 1.67). Furthermore, consider the operator of restriction $\iota : L^2(\mathbb{R}^d) \to L^2(\Omega)$, $\iota u(x) = u(x)$ for any $x \in \Omega$ and any $u \in L^2(\mathbb{R}^d)$, then ι is clearly bounded.

The operator $\iota j k : H_0^1(\Omega) \to L^2(\Omega)$ is then compact (a composition of two continuous operators and a compact operator), and this operator is exactly the embedding of $H_0^1(\Omega)$ into $L^2(\Omega)$. $\square$

Using Theorem 2.33 we arrive at:

Corollary 2.37 (Dirichlet Laplacians with compact resolvents). *Let $\Omega \subset \mathbb{R}^d$ be a bounded open set, then the Dirichlet Laplacian in Ω, viewed as a self-adjoint operatorin $L^2(\Omega)$, has compact resolvent.*

There is no literal extension of Corollary 2.37 for Neumann Laplacians: there are bounded domains Ω (with "bad" boundaries) such that the respective Neumann Laplacians are not with compact resolvents (and the embedding of $H^1(\Omega)$ in $L^2(\Omega)$ is not compact.) Anyway, under some additional assumptions the result still holds true.

Definition 2.38. We say that an open set $\Omega \subset \mathbb{R}^d$ has the *extension property*, if there exists a bounded linear operator $E : H^1(\Omega) \to H^1(\mathbb{R}^d)$ such that $Eu(x) = u(x)$ for all $u \in H^1(\Omega)$ and all $x \in \Omega$. Such an operator E is usually called an *extension operator*.

The following result will be used without detailed proof:

Proposition 2.39. *Any bounded open set with Lipschitz boundary has the extension property.*

Proof idea. If Ω is a half-space, $\Omega = (0,\infty) \times \mathbb{R}^{d-1}$, then one explicitly constructs a bounded extension operator $E : H^1(\Omega) \to H^1(\mathbb{R}^d)$ by $(Eu)(x_1, x') = u(|x_1|, x')$.

If Ω is bounded, with Lipschitz boundary, one covers $\partial\Omega$ by finitely many balls B_j and finds bi-Lipschitz maps $\Phi_j : B_j \to B_{r_j}(0)$ such that

$$\Phi_j(\Omega \cap B_j) = \{(x_1, x') \in B_{r_j}(0) : x_1 > 0\},$$

and takes a partition of unity χ_j with $\operatorname{supp} \chi_j \subset B_j$. If $u \in H^1(\Omega)$, then the functions $v_j := (\chi_j u) \circ \Phi_j^{-1}$ can be viewed as H^1-functions on the half-space, and they can be extended to a $H^1(\mathbb{R}^d)$-function as above. One uses Φ_j in the reverse direction and takes the sum over j. (One can show that if Φ is bi-Lipschitz, then a function u is in H^1 if and only if $u \circ \Phi$ is in H^1.) $\square$

Proposition 2.40 (Compact embeddings of $\boldsymbol{H^1}$). *If Ω is a bounded open set with the extension property (for example, with Lipschitz boundary), then the embedding $H^1(\Omega) \hookrightarrow L^2(\Omega)$ is compact.*

Proof. Let $E : H^1(\Omega) \to H^1(\mathbb{R}^d)$ be a bounded extension operator. Pick any $\chi \in C_c^\infty(\mathbb{R}^d)$ with $\chi = 1$ in Ω $B \supset \Omega$ and let B be an open ball containing $\operatorname{supp} \chi$, then the operator

$$E_0 : H^1(\Omega) \ni u \mapsto \chi E u \in H_0^1(B)$$

is bounded. The embedding $j : H_0^1(B) \hookrightarrow L^2(B)$ is compact by Proposition 2.36, as B is bounded. Furthermore, let $\iota : L^2(B) \to L^2(\Omega)$ be the operator of restriction, $\iota u(x) = u(x)$ for all $x \in \Omega$, which is bounded. The embedding $J : H^1(\Omega) \hookrightarrow L^2(\Omega)$ can be decomposed as $J = \iota j E_0$, and it is compact due to the compactness of j. $\square$

By applyng Theorem 2.33 we show:

Corollary 2.41 (Neumann Laplacians with compact resolvents). *If Ω is a bounded open set with the extension property (for example, with Lipschitz boundary), then the Neumann Laplacian in Ω, viewed as a self-adjoint operator in $L^2(\Omega)$, has compact resolvent.*

The compactness results can also be applied to the Schrödinger operators as follows:

Theorem 2.42 (Schrödinger operators with growing potentials). *Let $V \in L^2_{\mathrm{loc}}(\mathbb{R}^d)$ be real-valued, semibounded from below. Denote $w(r) := \inf_{|x|\geq r} V(x)$. If*

$$\lim_{r\to+\infty} w(r) = +\infty,$$

then the Schrödinger operator $T = -\Delta + V$ (defined as the Friedrichs extension, see Example 1.80) has compact resolvent.

Proof. Without loss of generality we assume $V \geq 0$ (otherwise consider $T + c$ with a suitable $c \in \mathbb{R}$). Recall (Example 1.80) that the sesquilinear form for T is

$$t(u,u) = \int_{\mathbb{R}^d} \big(|\nabla u|^2 + V|u|^2\big)\,\mathrm{d}x,$$

whose domain $D(t) = H^1_V(\mathbb{R}^d)$ is equiped with the norm $\|u\|^2_{H^1_V} = \|u\|^2_{H^1} + \|\sqrt{V}u\|_{L^2}$. It is sufficient to show that the unit ball $\mathcal{F}_V$ in $H^1_V(\mathbb{R}^d)$ is relatively compact. For that we show that all assumptions in the Riesz-Kolmogorov criterion (Theorem 2.34) are satisfied.

- The condition (a) holds due to $\|u\|_{L_2} \leq \|u\|_{H^1_V} \leq 1$ for any $u \in \mathcal{F}_V$.
- If $u \in \mathcal{F}_V$, then

$$\begin{aligned}\int_{|x|\geq R} |u(x)|^2\,\mathrm{d}x &\leq \frac{1}{w(R)} \int_{|x|\geq R} V(x)|u(x)|^2\,\mathrm{d}x \\ &\leq \frac{\|\sqrt{V}u\|^2_{L^2}}{w(R)} \leq \frac{\|u\|^2_{H^1_V}}{w(R)} \leq \frac{1}{w(R)} \xrightarrow{R\to+\infty} 0,\end{aligned}$$

  and (b) follows.
- Due to $\|u\|_{H^1_V} \geq \|u\|_{H^1}$, the set $\mathcal{F}_V$ is contained in the unit ball of $H^1(\mathbb{R}^d)$, and (c) holds by Lemma 2.35. □

Remark 2.43. The assumptions in Theorems 2.42 (Schrödinger operators) and Corollaries 2.37 and 2.41 are only sufficient conditions: it is known that they are not necessary (this will be discussed in the exercises).

3 Spectral theorem

To be provided with a certain motivation, let T be either a compact self-adjoint operator or a semibounded self-adjoint operator with compact resolvent in a Hilbert space $\mathcal{H}$. As shown in the previous sections, there exists an orthonormal basis $(e_n)_{n\in\mathbb{N}}$ in $\mathcal{H}$ and real numbers λ_n such that, with

$$Tx = \sum_{n\in\mathbb{N}} \lambda_n \langle e_n, x\rangle e_n \quad \text{for all } x \in D(T),$$

and the domain $D(T)$ is characterized by

$$D(T) = \Big\{ x \in \mathcal{H} : \sum_{n\in\mathbb{N}} \lambda_n^2 \big|\langle e_n, x\rangle\big|^2 < \infty \Big\}.$$

Recall that if we introduce the map $U : \mathcal{H} \to \ell^2(\mathbb{N})$ defined by $Ux =: (x_n)$, $x_n = \langle e_n, x\rangle$, then the operator UTU^* becomes the discrete multiplication operator $(x_n) \mapsto (\lambda_n x_n)$.

If f is a bounded function on $\mathbb{R}$, then one can define a linear operator $f(T) \in \mathcal{B}(\mathcal{H})$ by

$$f(T)x = \sum_{n\in\mathbb{N}} f(\lambda_n)\langle e_n, x\rangle e_n,$$

then the map $f \mapsto f(T)$ satisfies a number of properties. For example, $(fg)(T) = f(T)g(T)$, $\overline{f}(T) = f(T)^*$ etc. The existence of such a construction allows one to show that some complicated equations have solutions (in a suitable class) and even to write rather explicit expressions for the solutions. For example, one can easily show that the initial value problem

$$-ix'(t) = Tx(t),\ x(0) = y \in D(T), \quad x : \mathbb{R} \to D(T),$$

has a solution that can be written as $x(t) = f_t(T)y$ with $f_t(x) = e^{itx}$. Informally speaking, for a large class of equations involving the operator T one may first assume that T is a real constant and obtain a formula for the solution, and then one can give this formula an operator-valued meaning using the above map $f \mapsto f(T)$.

At this point, the definition of $f(T)$ only makes sense for a restricted class of self-adjoint operators T (admitting an orthonormal eigenbasis). The aim of the present section is to develop a similar theory for the general self-adjoint operators. Namely, we will show that:

- each self-adjoint operator is unitarily equivalent to a multiplication operator in some measure space (spectral theorem),
- the operators $f(T)$ are uniquely defined for any self-adjoint T and a large class of functions f (functional calculus).

In order to cover both bounded and unbounded self-adjoint operators, it will be convenient to prove analogous results first for the unitary operators (which are easier to deal with as they are bounded) and then to pass to self-adjoint operators using a suitable transform.

3.1 Spectral theorem for unitary operators

We will intensively use the unit circle

$$\mathbb{S} := \{z \in \mathbb{C} : \; |z| = 1\}.$$

One can consider $\mathbb{S}$ as a one-dimensional submanifold, and the integration and the smoothness over $\mathbb{S}$ are then well-defined: if one uses $\theta \mapsto e^{i\theta}$ as a local chart, then for $f : \mathbb{S} \to \mathbb{C}$ one has:

- $f \in C^k(\mathbb{S})$ if and only if $\theta \to f(e^{i\theta})$ is a 2π-periodic C^k-function on $\mathbb{R}$.
- and
$$\int_{\mathbb{S}} f \,\mathrm{d}s := \int_0^{2\pi} f(e^{i\theta})\,\mathrm{d}\theta,$$
which naturally defines the spaces $L^p(\mathbb{S})$.

If $f \in L^1(\mathbb{S})$, its *Fourier coefficients* are the numbers

$$\widehat{f}(n) := \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta})e^{-in\theta}\,\mathrm{d}\theta, \quad n \in \mathbb{Z}.$$

If $f \in C^k(\mathbb{S})$ for some $k \in \mathbb{N}$, then using the integration by parts we have

$$\begin{aligned} c_k := \Big\| \frac{d^k}{d\theta^k} f(e^{i\theta}) \Big\|_{L^1(0,2\pi)} &\ge \Big| \int_0^{2\pi} \Big(\frac{d^k}{d\theta^k} f(e^{i\theta}) \Big) e^{-in\theta}\,\mathrm{d}\theta \Big| \\ &= \Big| (in)^k \int_0^{2\pi} f(e^{i\theta})e^{-in\theta}\,\mathrm{d}\theta \Big| = 2\pi |n|^k \widehat{f}(n), \end{aligned}$$

which shows that $\widehat{f}(n) = O(|n|^{-k})$ for large n. In particular, for $f \in C^\infty(\mathbb{S})$ the last estimate holds with any $k \in \mathbb{N}$, and it is known from basic analysis courses that

$$f(e^{i\theta}) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)e^{in\theta}, \quad \theta \in \mathbb{R},$$

while the series on the right-hand side converges uniformly in θ. By denoting $z := e^{i\theta}$ we have then

$$f(z) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)z^n, \quad z \in \mathbb{S},$$

with the uniform convergence on the right-hand side.

Let $U : \mathcal{H} \to \mathcal{H}$ be a unitary operator. Recall that this means that U is bijective and that $\|Ux\| = \|x\|$ for all $x \in \mathcal{H}$, which is equivalent to $U^* = U^{-1}$. It is easy to check that $\operatorname{spec} U \subset \mathbb{S}$: the inclusion $\operatorname{spec} U \subset \{z : |z| \le 1\}$ follows from Prop. 2.15, and if $|z| < 1$, then $U - z = U(I - zU^*)$, and due to $\|zU^*\| = |z| < 1$ the both operators on the right-hand side are bijective with bounded inverses.

Our idea is to define operators $f(U)$, for a large class of functions $f : \mathbb{S} \to \mathbb{C}$, by

$$f(U) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)U^n. \tag{3.1}$$

Remark that for $f \in C^\infty(\mathbb{S})$ the series on the right-hand of (3.1) converges with respect to the operator norm and defines a continuous linear operator. In particular, for the constant function $f \equiv 1$ the operator $f(U)$ is the identity map. We would like to extend the definition to a larger class of functions.

Theorem 3.1 (Continuous functional calculus for unitary operators). *The map $C^\infty(\mathbb{S}) \ni f \mapsto f(U) \in \mathcal{B}(\mathcal{H})$ defined by (3.1) extends uniquely to a linear map $C^0(\mathbb{S}) \to \mathcal{B}(\mathcal{H})$ such that for any $f, g \in C^0(\mathbb{S})$ one has*

(a) $f(U)^* = \overline{f}(U)$,

(b) $f(U)g(U) = (fg)(U)$,

(c) if $f \geq 0$, *then* $f(U) \geq 0$,

(d) $\|f(U)\| \leq \|f\|_\infty$.

Proof. We first establish all the properties for functions $f, g \in C^\infty(\mathbb{S})$.

(a) Let $f \in C^\infty(\mathbb{S})$, then $\overline{f} \in C^\infty(\mathbb{S})$. One has

$$\overline{\widehat{f}(n)} = \widehat{\overline{f}}(-n) \text{ for any } n \in \mathbb{Z},$$

$$\begin{aligned} f(U)^* = \Big(\sum_{n\in\mathbb{Z}} \widehat{f}(n) U^n\Big)^* &= \sum_{n\in\mathbb{Z}} \overline{\widehat{f}(n)} U^{-n} \\ &= \sum_{n\in\mathbb{Z}} \overline{\widehat{f}(-n)} U^n = \sum_{n\in\mathbb{Z}} \widehat{\overline{f}}(n) U^n = \overline{f}(U). \end{aligned}$$

(b) Let $f, g \in C^\infty(\mathbb{S})$, then

$$\begin{aligned} f(U)g(U) &= \Big(\sum_{m\in\mathbb{Z}} \widehat{f}(m) U^m\Big)\Big(\sum_{m'\in\mathbb{Z}} \widehat{g}(m') U^{m'}\Big) \\ &= \sum_{n\in\mathbb{Z}} c_n U^n \text{ with } c_n := \sum_k \widehat{f}(k)\widehat{g}(n-k). \end{aligned}$$

At the same time

$$(fg)(e^{i\theta}) = \Big(\sum_{m\in\mathbb{Z}} \widehat{f}(m) e^{im\theta}\Big)\Big(\sum_{m'\in\mathbb{Z}} \widehat{g}(m') e^{im'\theta}\Big) = \sum_{n\in\mathbb{Z}} c_n e^{in\theta},$$

while the series converge uniformly, which shows that $c_n = \widehat{fg}(n)$.

(c) Let $f \in C^\infty(\mathbb{S})$ with $f \geq 0$. For $\varepsilon > 0$ consider the function $h_\varepsilon := \sqrt{f + \varepsilon} \in C^\infty(\mathbb{S})$. Then $h_\varepsilon(U)$ is defined by the above series. As h_ε is real-valued, the operator $h_\varepsilon(U)$ is self-adjoint by (a), and $f = h_\varepsilon^2 - \varepsilon$, which implies $f(U) = h_\varepsilon(U)^2 - \varepsilon I$ by (b). Then for any $v \in \mathcal{H}$ we have

$$\begin{aligned} \langle v, f(U)v\rangle &= \big\langle v, \big(h_\varepsilon(U)^2 - \varepsilon I\big)v\big\rangle \\ &= \langle h_\varepsilon(U)v, h_\varepsilon(U)v\rangle - \varepsilon\langle v, v\rangle = \|h_\varepsilon(U)v\|^2 - \varepsilon\|v\|^2 \geq -\varepsilon\|v\|^2, \end{aligned}$$

and by sending ε to 0 we obtain $\langle v, f(U)v\rangle \geq 0$.

(d) Let $f \in C^\infty(\mathbb{S})$ and $M := \|f\|_\infty$, then $M^2 - |f|^2 \ge 0$, and by (c) this means that for any $v \in \mathcal{H}$ we have $\big\langle v, (M^2 - |f|^2)v\big\rangle \ge 0$. We transform, using (a) and (b),

$$\big\langle v, (M^2 - |f|^2)v\big\rangle = M^2\|v\|^2 - \langle v, f(U)^* f(U)v\rangle = M^2\|v\|^2 - \|f(U)v\|^2,$$

and obtain $\|f(U)v\|^2 \le M^2\|v\|^2$.

Hence, (a)–(d) are proved for the C^∞-functions. The extension to $C^0(\mathbb{S})$ is done using the density: as $C^\infty(\mathbb{S})$ is dense in $C^0(\mathbb{S})$ with respect to $\|\cdot\|_\infty$, for any $f \in C^0(\mathbb{S})$ there exists $(f_n) \subset C^\infty(\mathbb{S})$ with $\|f - f_n\|_\infty \to 0$. Then (f_n) is Cauchy in $C^0(\mathbb{S})$ and then $\big(f_n(U)\big)$ is Cauchy in $\mathcal{B}(\mathcal{H})$. As $\mathcal{B}(\mathcal{H})$ is a Banach space, there exists $A := \lim f_n(U)$. One then routinely checks that A is independent of the choice of f_n, so one can define $f(U) := A$, and then one routinely checks (a)–(d) using the passage to the limit. □

For the next step we recall the Riesz representation theorem: if X is a compact metric space and $\beta : C^0(X) \to \mathbb{C}$ is a linear functional with $\beta(f) \ge 0$ for $f \ge 0$, then there exists a unique Borel measure μ on X such that

$$\beta(f) = \int_X f \,\mathrm{d}\mu \text{ for any } f \in C^0(X).$$

It follows from the construction of μ that $C^0(X)$ is dense in $L^2(X, \mu)$.

Lemma 3.2. *Let $v \in \mathcal{H}$ and let μ_v be the unique Borel measure on $\mathbb{S}$ defined by*

$$\langle v, f(U)v\rangle = \int_{\mathbb{S}} f \,\mathrm{d}\mu_v.$$

Remark that the right-hand side is linear in f and non-negative for non-negative f due to Theorem 3.1. Then the map $C^0(\mathbb{S}) \ni f \mapsto f(U)v \in \mathcal{H}$ has a unique continuous extension to an isometry $W_v : L^2(\mathbb{S}, \mathrm{d}\mu_v) \to \mathcal{H}$. This isometry satisfies $UW_v = W_v M_z$, where M_z is the multiplication operator $(M_z f)(z) = zf(z)$ in $L^2(\mathbb{S}, \mu_v)$.

Proof. Let $f, g \in C^0(\mathbb{S})$. We have

$$\begin{aligned}\langle W_v f, W_v g\rangle &= \big\langle f(U)v, g(U)v\big\rangle = \big\langle v, f(U)^* g(U)v\big\rangle \\ &= \big\langle v, (\overline{f}g)(U)v\big\rangle = (\text{Definition of } \mu_v) = \int_{\mathbb{S}} \overline{f} g \,\mathrm{d}\mu_v = \langle f, g\rangle_{L^2(X,\mu_v)}\end{aligned}$$

The subspace $C^0(\mathbb{S})$ is dense in $L^2(\mathbb{S}, \mu_v)$ and W_v preserves the scalar product, so it extends by density in a unique way to an isometry $L^2(\mathbb{S}, \mu_v) \to \mathcal{H}$.

By Theorem 3.1(b), for any $f \in C^0(\mathbb{S})$ we have

$$(M_z f)(U) = \big((z \mapsto z) \cdot f\big)(U) = Uf(U), \quad M_z f(U)v = Uf(U)v,$$

which means that $M_z W_v = UW_v$ on $C^0(\mathbb{S})$, and this equality extends by density to the whole $L^2(\mathbb{S}, \mu_v)$. □

Remark 3.3 (Countable direct sums). If $(\mathcal{H}_n)_{n\in N}$ is a finite or countable family of Hilbert spaces, their direct sum $\mathcal{H} = \bigoplus_{n\in N} \mathcal{H}_n$ is defined by

$$\mathcal{H} = \Big\{x = (x_n)_{n\in N} : \ x_n \in \mathcal{H}_n, \ \|x\|^2_{\mathcal{H}} := \sum_{n\in N} \|x_n\|^2_{\mathcal{H}_n} < \infty\Big\},$$

and one can check (Exercise) that it is a Hilbert space. If U_n are unitary operators in $\mathcal{H}_n$, then $U : (x_n) \mapsto (U_n x_n)$ is clearly a unitary operator in $\mathcal{H}$. Furthermore, if T_n are self-adjoint operators in $\mathcal{H}_n$, then one easily checks (Exercise) that the direct sum $T = \bigoplus_{n\in N} T_n$ defined by

$$T : (x_n) \mapsto (T_n x_n), \quad D(T) = \Big\{(x_n) \in \mathcal{H} : \ x_n \in D(T_n), \ \sum_{n\in N} \|T x_n\|^2_{\mathcal{H}_n} < \infty\Big\},$$

is a self-adjoint operator in $\mathcal{H}$.

Now we arrive to the main result of the subsection:

Theorem 3.4 (Spectral theorem for unitary operators). *Let U be a unitary operator in a separable Hilbert space $\mathcal{H}$. Then there exist a subset $N \subset \mathbb{N}$, finite Borel measures ν_n on $\mathbb{S}$, $n \in N$, and a unitary map*

$$W : L^2(Y, \nu) \to \mathcal{H}, \quad Y = \mathbb{S} \times N, \quad \nu(A \times \{n\}) = \nu_n(A) \text{ for } A \subset \mathbb{S},$$

such that $W^{-1}UW = M_\rho$ with the function $\rho : Y \ni (y, n) \mapsto y \in \mathbb{S} \subset \mathbb{C}$.

Remark 3.5. In a less formal language, the set Y is consists of $|N|$ disjoint copies of $\mathbb{S}$, and the restriction of ν to the nth copy coincides with ν_n.

There are versions of the spectral theorem for non-separable Hilbert spaces, then one should use the axiom of choice and pay more attention to proper definitions of direct sums for uncountable measure spaces and Hilbert spaces.

Proof. As $\mathcal{H}$ is separable, there exists a dense subset $\{w_j\}_{j\in\mathbb{N}}$. We use Lemma 3.2 for the vector w_1 and obtain a measure ν_1 on $\mathbb{S}$ and an isomorphism $W_1 : L^2(\mathbb{S}, \nu_1) \to \mathcal{H}_1$, where $\mathcal{H}_1$ is a closed subspace of $\mathcal{H}$. If $\mathcal{H}_1 = \mathcal{H}$, the proof is finished. Otherwise we note that by construction $U(\mathcal{H}_1) = \mathcal{H}_1$, and then the unitarity of U implies $U(\mathcal{H}_1^\perp) = \mathcal{H}_1^\perp$, i.e. U can be viewed as a unitary operator in $\mathcal{H}_1$. We now find the first j with $w_j \notin \mathcal{H}_1$ and denote $v_2 :=$ the orthogonal projection of w_j on $\mathcal{H}_1^\perp$, then we use Lemma 3.2 for v_2, which gives rise to a measure ν_2 on $\mathbb{S}$ and an isomorphism $W_2 : L^2(\mathbb{S}, \nu_2) \to \mathcal{H}_2$, where $\mathcal{H}_2$ is a closed subspace of $\mathcal{H}_1^\perp$. If $\mathcal{H}_2 = \mathcal{H}_1^\perp$, the proof is finished, otherwise we take the first w_j which is not in $\mathcal{H}_1 \oplus \mathcal{H}_2$ and continue with its orthogonal projection on $(\mathcal{H}_1 \oplus \mathcal{H}_2)^\perp$ etc.

If the procedure stops after finitely many steps, the set N is finite, otherwise $N = \mathbb{N}$. By construction we have

$$\mathcal{H} = \bigoplus_{n\in N} \mathcal{H}_n, \quad W_n^{-1} U|_{\mathcal{H}_n} W_n = M_{z,n},$$

while $M_{z,n}$ acts as $(M_{z,n}f)(z) = zf(z)$ in $L^2(\mathbb{S}, \nu_n)$. One easily sees that

$$\Theta \ni L^2(Y, \nu) \ni x \mapsto (x_n) \in \bigoplus_{n\in N} L^2(\mathbb{S}, \nu_n), \quad x_n(z) = x(z, n),$$

is an isomorphism, and we arrive at the conclusion by taking

$$W := \Theta^{-1}\big(\bigoplus_{n\in N} W_n\big)\Theta. \qquad \square$$

3.2 Spectral theorem for self-adjoint operators

In this section we will prove similar results for self-adjoint operators. The passage between unitary and self-adjoint operators uses the following simple observation.

Lemma 3.6. *If T is self-adjoint, then $U := I - 2i(T + i)^{-1}$ is unitary.*

Proof. Remark first that $\operatorname{spec} T \subset \mathbb{R}$, and $(T \pm i)^{-1}$ are well-defined continuous operators. For arbitrary $x \in \mathcal{H}$ and $y \in D(T)$ we have

$$\begin{aligned}\langle x, y\rangle &= \big\langle (T + i)(T + i)^{-1}x, y\big\rangle = \big\langle (T + i)^{-1}x, (T - i)y\big\rangle \\ &= \big\langle x, \big[(T + i)^{-1}\big]^*(T - i)y\big\rangle,\end{aligned}$$

which shows that $[(T + i)^{-1}]^* = (T - i)^{-1}$. It follows that $U^* = I + 2i(T - i)^{-1}$ and $UU^* = U^*U$ as $(T - i)^{-1}$ and $(T + i)^{-1}$ commute (Theorem 2.4). Finally,

$$\begin{aligned}U^*U &= I - 2i(T + i)^{-1} + 2i(T - i)^{-1} + 4(T - i)^{-1}(T + i)^{-1} \\ &= I - 2i\Big[(T + i)^{-1} - (T - i)^{-1} + 2i(T - i)^{-1}(T + i)^{-1}\Big],\end{aligned}$$

and the expression in the square brackets is zero due to the resolvent identitites (Theorem 2.4). $\square$

Remark 3.7. The expression for U can be rewritten as $U = (T - i)(T + i)^{-1}$, which is an operator-valued version of the transform

$$c : \mathbb{R} \ni x \mapsto \frac{x - i}{x + i} \in \mathbb{S}. \tag{3.2}$$

Remark that $\mathbb{R}$ contains the spectra of all self-adjoint operators and the unit circle contains the spectra of all unitary operators. The above transform is called the *Cayley transform.*

We are now going to state the main result. In a short form, it says that any self-adjoint operator is unitarily equivalent to a multilplication operator on a measure space. In fact, one can even assume some special structure of the measure space and other additional properties:

Theorem 3.8 (Spectral theorem: Self-adjoint operators as multiplication operators). *Let T be a self-adjoint operator in a separable Hilbert space $\mathcal{H}$. Then there exist a subset $N \subset \mathbb{N}$, finite Borel measures μ_n on $\mathbb{R}$, $n \in N$, and a unitary map*

$$\Theta : L^2(X, \mu) \to \mathcal{H}, \quad X = \mathbb{R} \times N, \quad \mu(A \times \{n\}) = \mu_n(A) \textit{ for } A \subset \mathbb{R},$$

such that $\Theta^{-1} T \Theta = M_h$, where the function h is given by $h : X \ni (x, n) \mapsto x \in \mathbb{R}$.

Proof. Consider the unitary operator $U := I - 2i(T + i)^{-1}$. Using the spectral theorem for unitary operators (Theorem 3.4) we construct a subset $N \subset \mathbb{N}$, finite Borel measures ν_n on $\mathbb{S}$, $n \in N$, and a unitary map

$$W : L^2(Y, \nu) \to \mathcal{H}, \quad Y = \mathbb{S} \times N, \quad \nu(A \times \{n\}) = \nu_n(A) \text{ for } A \subset \mathbb{S},$$

such that $W^{-1} U W = M_\rho$ with $\rho : (y, n) \mapsto y$.

As $I - U = 2i(T + i)^{-1}$ is an injective operator, the operator $I - M_\rho \equiv M_{1-\rho}$ is also injective in $L^2(Y, \nu)$. In other words, $0 \notin \operatorname{spec_p} M_{1-z}$, which is equivalent to $1 \notin \operatorname{spec_p} M_\rho$, and by Proposition 2.7 one obtains $\nu\big(\rho^{-1}(1)\big) = 0$. We have

$$\nu\big(\rho^{-1}(1)\big) = \{1\} \times N = \bigcup_{n \in N} \{(1, n)\}, \quad \nu(\{1\} \times N) = \sum_{n \in N} \nu\big(\{1\} \times \{n\}\big) = \sum_{n \in N} \nu_n\big(\{1\}\big),$$

and we conclude that $\nu_n\big(\{1\}\big) = 0$ for all $n \in N$.

Denote by η the inverse of the Cayley transform c from (3.2),

$$\eta := c^{-1} : \mathbb{S} \ni y \mapsto i \frac{1 + y}{1 - y} \in \mathbb{R},$$

which is defined ν_n-a.e. on $\mathbb{S}$ for any $n \in \mathbb{N}$. Define μ_n to be the pushfoward $\eta_* \nu_n$, more precisely,

$$\mu_n(A) := \nu_n\big(\eta^{-1}(A)\big) \equiv \nu_n\big(c(A)\big), \quad A \subset \mathbb{R},$$

then μ_n are finite Borel measures on $\mathbb{R}$. We define X and μ as in the formulation of the theorem and remark that the pullback operator

$$\Phi : L^2(X, \mu) \to L^2(Y, \nu), \quad (\Phi f)(y, n) = f\big(\eta(y), n\big),$$

is unitary by construction. Now we define $\Theta := W \Phi$, which is by construction a unitary operator $L^2(X, \mu) \to \mathcal{H}$. Let us show that this operator satisfies all the required conditions.

Proof of $\Theta D(M_h) \subset D(T)$. Let $f \in D(M_h) \subset L^2(X, \mu)$, then $g := (h + i) f \in L^2(X, \mu)$. By the definition of Φ, for $(y, n) \in \mathbb{S} \times \mathbb{N}$ we have

$$(\Phi g)(y, n) = \Big(h\big(\eta(y), n\big) + i\Big)(\Phi f)(y, n) = \big(\eta(y) + i\big) \Phi f(y, n).$$

We have

$$\eta(y) + i = i \frac{1 + y}{1 - y} + i = \frac{i(1 + y) + i(1 - y)}{1 - y} = \frac{2i}{1 - y},$$

which yields $(1-y)(\Phi g)(y,n) = 2i(\Phi f)(y,n)$. The last equality can be rewritten as $(I - M_\rho)\Phi g = 2i\Phi f$. Now we apply W on the both parts: using $WM_\rho = UW$ we arrive at $(I-U)\Theta g = 2i\Theta f$. Using the definition of U we have

$$I - U = I - \big(I - 2i(T+i)^{-1}\big) = 2i(T+i)^{-1},$$

which results in

$$\Theta f = (T+i)^{-1}\Theta g \in D(T). \tag{3.3}$$

Proof of $\Theta^{-1}D(T) \subset D(M_h)$. Let $v \in D(T)$ and $w := (T+i)v$. we have then

$$v = (T+i)^{-1}w = \frac{1}{2i}\,(I-U)w.$$

We have $W^{-1}UW = M_\rho$, which gives $I - U = WM_{1-\rho}W^{-1}$ and then

$$W^{-1}v = \frac{1}{2i}\,M_{1-\rho}W^{-1}w \quad \Rightarrow \quad \Phi^{-1}W^{-1}v = \frac{1}{2i}\,\Phi^{-1}M_{1-\rho}W^{-1}w. \tag{3.4}$$

Remark that for $g \in L^2(Y,\nu)$ and $(x,n) \in X$ we have

$$(\Phi^{-1}M_{1-\rho}g)(x,n) = \big(M_{1-\rho}g\big)\big(c(x),n\big) = \big(1 - c(x)\big)g\big(c(x),n\big),$$

and with the help of

$$1 - c(x) = 1 - \frac{x-i}{x+i} = \frac{x+i-(x-i)}{x+i} = \frac{2i}{x+i}$$

one rewrites it as

$$(\Phi^{-1}M_{1-\rho}g)(x,n) = \frac{2i}{x+i}\,g\big(c(x),n\big) \equiv \frac{2i}{h(x,n)+i}(\Phi^{-1}g)(x,n).$$

This computaton shows that $\Phi^{-1}M_{1-\rho} = 2i(M_h+i)^{-1}\Phi^{-1}$. Using this identity in (3.4) we arrive at $\Theta^{-1}v = (M_h+i)^{-1}\Theta^{-1}w \in D(M_h)$.

Proof of $\Theta^{-1}T\Theta = M_h$. We have just shown that $\Theta D(M_h) = D(T)$. If $f \in D(M_h)$ and $g = (M_h+i)f$, then in (3.3) we already saw that $(T+i)^{-1}\Theta g = \Theta f$. Applying $(T+i)$ one the both sides and using the definition of g we arrive at

$$\Theta(M_h+i)f = (T+i)\Theta f \quad \Rightarrow \quad \Theta M_h f = T\Theta f.$$

As $f \in D(M_h)$ was arbitrary, this shows the sought equality. $\square$

Remark 3.9. It should be clear that the representation obtained in Theorem 3.8 is not unique. If T is a self-adjoint operator admitting an orthonormal basis $(e_n)_{n\in\mathbb{N}}$ of eigenvectors with eigenvalues λ_n, then it fits into the general scheme by setting

$$N := \mathbb{N}, \quad \mu_n := \delta_{\lambda_n}, \quad \Theta f := \sum_{n\in\mathbb{N}} f(\lambda_n, n) e_n,$$

where δ_a is the point measure defined by

$$\delta_a(A) = \begin{cases} 1, & a \in A, \\ 0, & a \notin A. \end{cases}$$

If all λ_n are distinct, one can "put them on the same copy of $\mathbb{R}$" and consider, for example,

$$N := \{1\}, \quad \mu_1 := \sum_{n\in\mathbb{N}} a_n \delta_{\lambda_n},$$

where (a_n) is an arbitrary sequence of positive numbers with

$$\sum_{n\in\mathbb{N}} a_n < \infty,$$

and in that case the associated unitary transform is given by

$$\Theta f := \sum_{n\in\mathbb{N}} \sqrt{a_n}\, f(\lambda_n, 1) e_n.$$

If one does not require the finiteness of μ_1 one can even take simply $a_n \equiv 1$. Indeed, there are many other ways to "distribute" λ_n among several copies of $\mathbb{R}$. By looking at this example one easily unbderstands that the minimal cardinality of the set N coincides with $\max_{n\in\mathbb{N}} \dim\ker(T - \lambda_n)$.

For general self-adjoint T, the minimal cardinality of N in Theorem 3.8 is usually referred to as the *spectral multiplicity* of T.

Remark 3.10 (Norms of multiplication operators). In view of Theorem 3.8, many questions of the general spectral theory are reduced to the study of multilplication operators, so let us complement the respective constructions from Example 2.6. If (X, μ) is a measure space and $f : X \to \mathbb{R}$ is measurable function, then one denotes

$$\operatorname*{ess\,sup}_{x\in X} f(x) := \inf\big\{M \in \mathbb{R} : f \le M \ \mu\text{-a.e.}\big\} \equiv \sup(\operatorname{ess\,ran} f).$$

It is clear that $\operatorname{ess\,sup}_{x\in X} f(x) \le \sup_{x\in X} f(x)$, and neglecting the values of f on any zero measure subset has no effect for $\operatorname{ess\,sup} f$.

Let M_f be the multiplication operator in $L^2(X, \mu)$ defined as in Example 2.6 with a bounded measurable function f, then one easily shows that

$$\|M_f\| = \operatorname*{ess\,sup}_{x\in X} \big|f(x)\big| \equiv \sup_{\lambda\in\operatorname{spec} M_f} |\lambda|. \tag{3.5}$$

In fact, if $|f| \le M$ μ-a.e., then for any $\phi \in L^2(X,\mu)$, $\phi \neq 0$, one has

$$\|M_f\phi\|^2 = \int_X \big|f(x)\phi(x)\big|^2\,\mathrm{d}\mu(x) \le M^2 \int_X \big|\phi(x)\big|^2\,\mathrm{d}\mu(x) = M^2\|\phi\|^2, \quad \frac{\|M_f\phi\|}{\|\phi\|} \le M,$$

and taking first sup over ϕ and then inf over all possible M one arrives at the inequality $\|M_f\| \le \operatorname{ess\,sup}_{x\in X}\big|f(x)\big|$. On the other hand, let $M := \operatorname{ess\,sup}_{x\in X}\big|f(x)\big|$. If $M = 0$, then $M_f = 0$, and (3.5) is true. If $M > 0$, then for any $m \in \mathbb{N}$ the subset

$$\widetilde{S}_m := \big\{x \in X : \ |f(x)| > M - 2^{-m}\big\}$$

has non-zero measure. Choose $S_m \subset \widetilde{S}_m$ with $0 < \mu(S_m) < \infty$ and denote by ϕ_m the indicator function of S_m, then one has, for all m with $2^{-m} < M$

$$\begin{aligned}\|M_f\phi_m\|^2 &= \int_X \big|f(x)\phi_m(x)\big|^2\,\mathrm{d}\mu(x) = \int_{S_m} \big|f(x)\big|^2\,\mathrm{d}\mu(x)\\ &\ge (M-2^{-m})^2 \int_{S_m} 1\,\mathrm{d}\mu(x) = (M-2^{-m})^2\|\phi_m\|^2,\end{aligned}$$

showing $\|M_f\| \ge M - 2^{-m}$ for any sufficiently large m, and sending m to ∞ one finishes the proof of (3.5).

Due to the spectral theorem we can now define the operators $f(T)$ for a large class of functions f. The main idea is very simple: if T is unitarily equivalent to M_h, then $f(T)$ must be unitarily equivalent to $M_{f\circ h}$.

Recall that the elements of the σ-algebra generated by open subsets are called *Borel subsets*. A function $f : \mathbb{R} \to \mathbb{C}$ is called Borel, if the preimage of any Borel subset is Borel. We denote

$$\mathcal{B}_\infty(\mathbb{R}) := \big\{f : \mathbb{R} \to \mathbb{C} : \ f \text{ is a bounded Borel function}\big\}.$$

Theorem 3.11 (Functional calculus for self-adjoint operators). *Let T be a self-adjoint operator in a separable Hilbert space $\mathcal{H}$ and let μ, Θ and h be as in Theorem 3.8. Then the linear map $\mathcal{B}_\infty(\mathbb{R}) \ni f \mapsto f(T) \in \mathcal{B}(\mathcal{H})$ given by*

$$f(T) := \Theta M_{f\circ h}\Theta^{-1} \tag{3.6}$$

has the following properties:

(a) it is a $*$-homomorphism, i.e.

$$(fg)(T) = f(T)g(T), \quad f(T)^* = \overline{f}(T), \quad f, g \in \mathcal{B}_\infty(\mathbb{R}),$$

(b) for any $f \in \mathcal{B}_\infty(\mathbb{R})$ one has

$$\|f(T)\| \le \sup_{\lambda\in\operatorname{spec} T}\big|f(\lambda)\big|,$$

with equality if f is continuous,

(c) if $f_n, f \in \mathcal{B}_\infty(\mathbb{R})$ such that $f_n \to f$ pointwise and $\|f_n\|_\infty$ is bounded, then $f_n(T) \to f(T)$ in the strong sense, i.e. $f_n(T)v \to f(T)v$ for any $v \in \mathcal{H}$.

(d) for any $z \in \mathbb{C} \setminus \mathbb{R}$ and the functions $r_z : \mathbb{R} \ni x \mapsto (x-z)^{-1} \in \mathbb{C}$ there holds $r_z(T) = (T-z)^{-1}$.

Moreover, any linear map $\mathcal{B}_\infty(\mathbb{R}) \ni f \mapsto f(T) \in \mathcal{B}(\mathcal{H})$ satisfying the above properties (a)–(d) is given by (3.6) *(in other words, the functional calculus is unique).*

Proof. (a) The both properties easily follow from the definition of the multiplication operators and the unitarity of Θ:

$$\begin{aligned}
f(T)g(T) &= \Theta M_{f\circ h}\Theta^{-1}\Theta M_{g\circ h}\Theta^{-1} = \Theta M_{f\circ h}M_{g\circ h}\Theta^{-1}\\
&= \Theta M_{(f\circ h)(g\circ h)}\Theta^{-1} = \Theta M_{(fg)\circ h}\Theta^{-1} = (fg)(T),\\
f(T)^* &= \big(\Theta M_{f\circ h}\Theta^{-1}\big)^* = (\Theta^{-1})^*(M_{f\circ h})^*\Theta^*\\
&= \Theta M_{\overline{f\circ h}}\Theta^{-1} = \Theta M_{\overline{f}\circ h}\Theta^{-1} = \overline{f}(T).
\end{aligned}$$

(b) The unitarity of Θ and Proposition 2.7 (spectrum of multiplication operators) show that

$$\operatorname{spec} T = \operatorname{ess\,ran} h, \quad \operatorname{spec} f(T) = \operatorname{ess\,ran} f \circ h.$$

The first equality implies that the set $H := h^{-1}(\operatorname{spec} T)$ satisfies $\mu(H^{\mathsf{C}}) = 0$. As discussed in Remark 3.10, one has $\|f(T)\| = \operatorname{ess\,sup}_{x\in X} |f\circ h(x)|$, and we estimate

$$\begin{aligned}
\operatorname*{ess\,sup}_{x\in X} \big|f\circ h(x)\big| &= \operatorname*{ess\,sup}_{x\in H} \big|f\circ h(x)\big| = \operatorname{ess}\sup_{x\in h^{-1}(\operatorname{spec} T)} \big|f\circ h(x)\big|\\
&\le \sup_{x\in h^{-1}(\operatorname{spec} T)} \big|f\circ h(x)\big| = \sup_{\lambda\in\operatorname{spec} T} \big|f(\lambda)\big|,
\end{aligned}$$

which shows the first claim.

Now assume additionally that f is continuous. Let

$$\lambda \in \operatorname{spec} T, \quad \varepsilon > 0, \quad J_\varepsilon := \big(f(\lambda)-\varepsilon, f(\lambda)+\varepsilon\big),$$

then the set $I_\varepsilon := f^{-1}(J_\varepsilon)$ is open and contains an open interval $K_\delta := (\lambda-\delta, \lambda+\delta)$ with some $\delta > 0$. Due to $\lambda \in \operatorname{spec} T \equiv \operatorname{ess\,ran} h$ we have

$$\mu\big(\{x \in X: \ h(x) \in K_\delta\}\big) > 0.$$

Due to $\{x \in X: \ f\big(h(x)\big) \in J_\varepsilon\} = \{x \in X: \ h(x) \in I_\varepsilon\} \supset \{x \in X: \ h(x) \in K_\delta\}$ we obtain

$$\mu\big(\{x \in X: \ f\big(h(x)\big) \in J_\varepsilon\}\big) > 0.$$

As $\varepsilon > 0$ was arbitrary, this shows that $f(\lambda) \in \operatorname{ess\,ran} f\circ h \equiv \operatorname{spec} M_{f\circ h}$, and then, with the help of (3.5),

$$\big|f(\lambda)\big| \le \sup_{z\in\operatorname{spec} M_{f\circ h}} |z| = \|M_{f\circ h}\| = \|f(T)\|.$$

This shows that $\sup_{\lambda\in\operatorname{spec} T} |f(\lambda)| \le \|f(T)\|$ and completes the proof of (b).

For (c), let $K > 0$ be such that $\|f_n\| \le K$ for all n, then also $|f| \le K$. Let $v \in \mathcal{H}$ and $\phi := \Theta^{-1} v \in L^2(X,\mu)$, then the convergence $f_n(T)v \to f(T)v$ is equivalent to $M_{f_n \circ h}\phi \mapsto M_{f\circ h}\phi$. We have

$$\begin{aligned}
\big\|M_{f_n\circ h}\phi - M_{f\circ h}\phi\big\|^2 &= \int_X \big|M_{f_n\circ h}\phi(x) - M_{f\circ h}\phi(x)\big|^2 \,\mathrm{d}\mu(x)\\
&= \int_X \Big| f_n\big(h(x)\big)\phi(x) - f\big(h(x)\big)\phi(x)\Big|^2 \,\mathrm{d}\mu(x)\\
&= \int_X \underbrace{\Big| f_n\big(h(x)\big) - f\big(h(x)\big)\Big|^2 |\phi(x)|^2}_{=:F_n(x)} \,\mathrm{d}\mu(x).
\end{aligned}$$

Due to the initial assumptions on f_n and f we have:

- $|F_n| \le 4K^2|\phi|^2 \in L^1(X,\mu)$,
- $\lim_{n\to\infty} F_n(x) = 0$ for a.e. $x \in X$ (while $f_n(y)$ converge to $f(y)$ for any $y \in X$, the function ϕ is defined μ-a.e. only),

and the dominated convergence theorem shows that $\|M_{f_n\circ h}\phi - M_{f\circ h}\phi\| \to 0$.

(d) Actually this part was already implicitly covered in the proof of the spectral theorem, but we repeat the constructions. Let $v \in \mathcal{H}$ and $\varphi := \Theta^{-1}v$. Then $\psi := (r_z \circ h)\varphi \in D(M_h)$ and $(M_h - z)\psi = \varphi$, i.e. $(M_h - z)M_{r_z}\varphi = \varphi$, i.e.

$$\underbrace{\Theta^{-1}(T-z)\Theta}_{=M_h - z} M_{r_z}\Theta^{-1}v = \Theta^{-1}v \quad \Rightarrow \quad \Theta^{-1}(T-z)r_z(T)v = \Theta^{-1}v.$$

By applying $(T-z)^{-1}\Theta$ on the both sides we arrive at $r_z(T)v = (T-z)^{-1}v$.

It remains to show the uniqueness. This parts uses a number of facts from the measure and integration theory, so we only present the main steps. Let $\Phi_1, \Phi_2 : \mathcal{B}_\infty(\mathbb{R}) \to \mathcal{B}(\mathcal{H})$ be two linear maps satisfying (a)–(d). From (d) it follows that Φ_1 and Φ_2 coincide on the linear combinations $\sum c_j r_{z_j}$ with $c_j \in \mathbb{C}$ and $z_j \in \mathbb{C}\setminus\mathbb{R}$. Using the Stone-Weierstrass theorem one easily sees that these linear combinations are dense (with respect to $\|\cdot\|_\infty$) in the space $C^0_\infty(\mathbb{R})$ of bounded continuous functions on $\mathbb{R}$ that vanish at infinity, and then due to (b) the maps Φ_1 and Φ_2 agree on $C^0_\infty(\mathbb{R})$. Now consider

$$\mathcal{A} := \big\{ f \in \mathcal{B}_\infty(\mathbb{R}) : \ \Phi_1(f) = \Phi_2(f)\big\},$$

then $\mathcal{A}$ is an algebra of functions by (a), and we have just shown that $C^0_\infty(\mathbb{R}) \subset \mathcal{A}$. Furthemore, due to (c) the set $\mathcal{A}$ is closed under pointwise limits of uniformly bounded sequences. The characteristic function of any closed interval can be realized as the pointwise limit of a uniformly bounded sequence of bounded continuous functions vanishing at infinity, which shows that the characteristic function of any closed interval belongs to $\mathcal{A}$. We now remark that the supports of functions in $\mathcal{A}$ form a σ-algebra. We have just shown that this σ-algebra contains all closed intervals, hence, it contains all Borel subsets. It then follows that $\mathcal{A}$ contains the characteristic functions of all Borel subsets and then also all simple Borel functions (i.e. Borel functions taking only a finite number of values). As each bounded Borel functions can be realized as the pointwise limit of a uniformly bounded sequence of simple Borel functions, one arrives at $\mathcal{A} = \mathcal{B}_\infty(\mathbb{R})$. $\square$

Remark 3.12. Remark that the approach used in the part (c) of the proof is very typical and powerful, as it reduces the convergence in an abstract Hilbert space $\mathcal{H}$ to the use of the dominated convergence in a measure space: this passage is made possible due to the spectral theorem. Similar arguments will be used later at many places.

Remark 3.13. Remark that the assertion (b) means that the only the values of f on $\operatorname{spec} T$ are of importance for the definition of $f(T)$: if $f, g \in \mathcal{B}_\infty(\mathbb{R})$ with $f = g$ on $\operatorname{spec} T$, then

$$\|f(T) - g(T)\| = \|(f-g)(T)\| \le \sup_{\lambda \in \operatorname{spec} T} \big|(f-g)(\lambda)\big| = 0,$$

i.e. $f(T) = g(T)$. Denote

$$\mathcal{B}_\infty(\operatorname{spec} T) := \big\{ f : \operatorname{spec} T \to \mathbb{C} : \quad f \text{ is Borel and bounded} \big\}$$

and for $f \in \mathcal{B}_\infty(\operatorname{spec} T)$ let $\widetilde{f} \in \mathcal{B}_\infty(\mathbb{R})$ be any extension of f to the whole of $\mathbb{R}$ (for example, one can simply take the extension by zero), and one sets

$$f(T) := \widetilde{f}(T),$$

then one directly arrives the following technical improvement:

Corollary 3.14. *All assertions of Theorem 3.11 hold for $\mathcal{B}_\infty(\mathbb{R})$ replaced by $\mathcal{B}_\infty(\operatorname{spec} T)$.*

The second assertion in (b) (the exact norm for continuous functions) follows from the fact that any bounded continuous function on $\operatorname{spec} T$ can be extended to a bounded continuous function on $\mathbb{R}$ with the same sup-norm (Tietze extension theorem).

Remark 3.15. For practical computations one does not need to have the canonical representation from Theorem 3.8 to construct the Borel functional calculus. It is sufficient to represent $T = U M_H U^{-1}$, where $U : \mathcal{H} \to L^2(X, d\mu)$ is a unitary operator and M_H is the multiplcation operator by some function H in some measure space (X, μ). Then for any $f \in \mathcal{B}_\infty(\mathbb{R})$ one can put $f(T) = U M_{f\circ H} U^{-1}$, and one easily checks that all the required properties are satisfied. For example, if T is the free Laplacian in $L^2(\mathbb{R}^d)$, then we know already that $T = \mathcal{F}^{-1} M_H \mathcal{F}$ where M_H is the multiplication by $|\xi^2|$ in $L^2(\mathbb{R}^d)$ and $\mathcal{F}$ is the Fourier transform. Then $f(T) = \mathcal{F}^{-1} M_F \mathcal{F}$ with $F : \xi \mapsto f(|\xi|^2)$.

For example, if $f(x) = \cos\sqrt{x}$ for $x \in [0, \infty)$ and $\varphi \in C_c^\infty(\mathbb{R}^d)$, then

$$f(T)\varphi(x) \equiv \big(\cos\sqrt{-\Delta}\,\big)\varphi(x) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{i\xi\cdot x} \big(\cos|\xi|\big)\, \widehat{\varphi}(\xi)\, \mathrm{d}\xi.$$

In fact, one can also define $f(T)$ for unbounded functions f using the same expressions: the functional calculus for unbounded functions is also unique if one assumes that suitable convergences of functions imply suitable convergences of the associated operators.

3.3 Some direct applications of the spectral theorem

In a sense, the rest of the course will consist of various applications of the spectral theorem and the functional calculus. Nevertheless, let us discuss the most immediate consequences (some of them can be proved by other methods, but the use of the spectral theorem gives a particularly transparent proof). We will use without further comments the objects appearing in the formulation of the spectral theorem.

Proposition 3.16 (The norm of the resolvent). *For any $z \in \operatorname{res} T$ one has* $\|(T-z)^{-1}\| = \dfrac{1}{\operatorname{dist}(z, \operatorname{spec} T)}$.

Proof. One has $(T-z)^{-1} = r(T)$ for the bounded continuous function

$$r : \operatorname{spec} T \ni x \mapsto (x-a)^{-1} \to \mathbb{C},$$

hence,

$$\begin{aligned}\|(T-z)^{-1}\| &= \sup_{\lambda \in \operatorname{spec} T} |r(\lambda)| = \sup_{\lambda \in \operatorname{spec} T} \frac{1}{|\lambda - z|} \\ &= \frac{1}{\inf_{\lambda \in \operatorname{spec} T} |\lambda - z|} = \frac{1}{\operatorname{dist}(z, \operatorname{spec} T)}.\end{aligned}$$ □

The above norm equality is often used in the following form:

Corollary 3.17 (Distance to spectrum). *Let $0 \neq v \in D(T)$ and $z \in \mathbb{C}$, then*

$$\operatorname{dist}(z, \operatorname{spec} T) \le \frac{\big\|(T-z)v\big\|}{\|v\|}.$$

Proof. If $z \in \operatorname{spec} T$, then the left-hand side is zero, and the inequality is valid. Assume now that $z \notin \operatorname{spec} T$ and use Proposition 3.16:

$$\|v\| = \|(T-z)^{-1}(T-z)v\| \le \|(T-z)^{-1}\| \, \|(T-\lambda)v\| = \frac{1}{\operatorname{dist}(z, \operatorname{spec} T)} \|(T-z)v\|.$$ □

Remark 3.18. The above norm equality for the resolvent is one of the basic tools to approximate the spectra of self-adjoint operators. For non-self-adjoint operators the estimate fails even in the finite-dimensional case. For example, take $\mathcal{H} = \mathbb{C}^2$ and

$$T = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

then $\operatorname{spec} T = \{0\}$, and for $z \neq 0$ we have

$$(T-z)^{-1} = -\frac{1}{z^2} \begin{pmatrix} z & 1 \\ 0 & z \end{pmatrix}.$$

For the vectors $e_1 = (1,0)$ and $e_2 = (0,1)$ one has $\langle e_1, (T-z)^{-1} e_2 \rangle = -z^{-2}$, which shows that the norm of the resolvent near $z = 0$ is of order $z^{-2} = \operatorname{dist}(z, \operatorname{spec} T)^{-2}$.

Theorem 3.19 (Bounds for spectra imply bounds for operators). *Let T be a self-adjoint operator and $c \in \mathbb{R}$, then*

(a) $T \geq c$ if and only $\operatorname{spec} T \subset [c, \infty)$.

(b) T is bounded with $\|T\| \leq c$ if and only if $\operatorname{spec} \subset [-c, c]$.

Proof. (a) If $T \geq c$, then the inclusion $\operatorname{spec} T \subset [c, \infty)$ was already shown in Theorem 2.23.

Now assume that $\operatorname{spec} T \subset [c, \infty)$ for some $c \in \mathbb{R}$. Then $\operatorname{ess\,ran} h \subset [c, \infty)$, in particular, $h \geq c$ μ-a.e., and for any $\varphi \in D(M_h)$ one has

$$\begin{aligned}\langle \varphi, M_h \varphi \rangle_{L^2(X,\mu)} &= \int_X \overline{\varphi(x)}\, h(x) \varphi(x)\, \mathrm{d}\mu(x)\\ &= \int_X h(x) |\varphi(x)|^2\, \mathrm{d}\mu(x) \geq \int_X c |\varphi(x)|^2\, \mathrm{d}\mu(x) = c \|\varphi\|^2_{L^2(X,\mu)}.\end{aligned}$$

If $v \in D(T)$, then $\varphi := \Theta^{-1} v \in D(M_h)$ and

$$\begin{aligned}\langle v, Tv \rangle_{\mathcal{H}} &= \langle \Theta\varphi, T\Theta\varphi \rangle = \langle \varphi, \Theta^{-1} T \Theta \varphi \rangle_{L^2(X,\mu)} \equiv \langle \varphi, M_h \varphi \rangle_{L^2(X,\mu)}\\ &\geq c \|\varphi\|^2_{L^2(X,\mu)} = c \|\Theta^{-1} v\|^2_{L^2(X,\mu)} = c \|v\|^2_{\mathcal{H}},\end{aligned}$$

which means that $T \geq c$.

(b) If T is bounded, then $\operatorname{spec} T \in \big[-\|T\|, \|T\|\big] \subset [-c, c]$ by Proposition 2.15.

Now assume that $\operatorname{spec} T \subset [-c, c]$. This means that $\operatorname{ess\,ran} h \equiv \operatorname{spec} T \subset [-c, c]$ and then $|h| \leq c$ μ-a.e., and then M_h is bounded with $\|M_h\| \leq c$. Then $T = \Theta M_h \Theta^{-1}$ is also bounded and has the same norm as M_h. □

In what follows we will use very frequently $f(T)$ for f :=indicator function of a set. Such operators have a special name:

Definition 3.20 (Spectral projectors). Let $\Omega \subset \mathbb{R}$ be a Borel subset. The *spectral projector* of T on Ω is the operator $P_T(\Omega) := 1_\Omega(T)$, where 1_Ω is the indicator function of Ω.

Let us summarize the most important properties of the spectral projectors.

Proposition 3.21 (Properties of spectral projectors). *For any self-adjoint operator T acting in a separable Hilbert space $\mathcal{H}$ the following assertions hold true:*

(a) Let $\Omega \subset \mathbb{R}$ be a Borel subset, then

(a.1) $P_T(\Omega)$ is an orthogonal projector.

(a.2) $P_T(\Omega) D(T) \subset D(T)$,

(a.3) $T P_T(\Omega) = P_T(\Omega) T$ on $D(T)$.

(b) If $\Omega, \widetilde{\Omega} \subset \mathbb{R}$ are Borel subsets, then

(b.1) if $\Omega \cap \widetilde{\Omega} = \emptyset$, then

– $\operatorname{ran} P_T(\Omega) \perp \operatorname{ran} P_T(\widetilde{\Omega})$,
– $P_T(\Omega \cup \widetilde{\Omega}) = P_T(\Omega) + P_T(\widetilde{\Omega})$,
– $\operatorname{ran} P_T(\Omega \cup \widetilde{\Omega}) = \operatorname{ran} P_T(\Omega) + \operatorname{ran} P_T(\widetilde{\Omega})$.

(b.2) if $\Omega \subset \widetilde{\Omega}$, *then* $\operatorname{ran} P_T(\Omega) \subset \operatorname{ran} P_T(\widetilde{\Omega})$,

(b.3) $P_T(\Omega) + P_T(\Omega^{\mathsf{C}}) = I$. *This means that* $\operatorname{ran} P_T(\Omega^{\mathsf{C}}) = \operatorname{ran} P_T(\Omega)^{\perp}$.

(c) $P_T\big((a,b)\big) = 0$ *if and only if* $\operatorname{spec} T \cap (a,b) = \emptyset$.

(d) for any $\lambda \in \mathbb{R}$ *there holds* $\operatorname{ran} P_T(\{\lambda\}) = \ker(T - \lambda)$.

(e) $\operatorname{spec} T = \{\lambda \in \mathbb{R} : P_T\big((\lambda - \varepsilon, \lambda + \varepsilon)\big) \neq 0$ *for all* $\varepsilon > 0\}$.

(f) For any $\lambda \in \mathbb{R}$ *and* $\varepsilon > 0$ *then*

(f.1) $\operatorname{ran} P_T\big((\lambda - \varepsilon, \lambda + \varepsilon)\big) \in D(T)$ *and* $\big\|(T - \lambda) P_T\big((\lambda - \varepsilon, \lambda + \varepsilon)\big)\big\| \le \varepsilon$,

(f.2) for any $\varphi \in \big[I - P_T\big((\lambda - \varepsilon, \lambda + \varepsilon)\big)\big] D(T)$ *there holds* $\|(T - \lambda)\varphi\| \ge \varepsilon \|\varphi\|$,

(f.3) for any $\varphi \in P_T\big([\lambda, \infty)\big) D(T)$ *there holds* $\langle \varphi, T\varphi \rangle \ge \lambda \|\varphi\|^2$.

(g) For any $\lambda \in \mathbb{R}$ *one has* $P_T(\{\lambda\}) = \mathrm{s} - \lim_{\varepsilon \to} P_T\big((\lambda - \varepsilon, \lambda + \varepsilon)\big)$.[5].

Proof. Due to the spectral theorem and the functional calculus, without loss of generality one may assume that $\mathcal{H} = L^2(X, \mu)$ and $T = M_h$ with X, μ, h as in the spectral theorem (Theorem 3.8), then $P_T(\Omega) = M_{1_\Omega \circ h}$.

To prove (a.1) we remark that $1_\Omega^2 = 1_\Omega$ and $\overline{1_\Omega} = 1_\Omega$, which gives

$$P_T(\Omega)^2 = P_T(\Omega), \quad P_T(\Omega)^* = P_T(\Omega).$$

The first equality means that $P_T(\Omega)$ is a projector, and the second one means that this projector is orthogonal. If $g : X \to \mathbb{C}$ is bounded and measurable, then the explicit description of $D(M_h)$ shows that $M_g D(M_h) \subset D(M_h)$ with $M_g M_h \varphi = M_h M_g \varphi$ for any $\varphi \in D(M_h)$. Taking $g := 1_\Omega \circ h$ one shows (a.2) and (a.3).

(b.1) If $\varphi \in \operatorname{ran} P_T(\Omega)$, then $\varphi = (1_\Omega \circ h)\varphi$, and $\varphi(x) = 0$ for all μ-a.e. $x \notin \Omega$. Analogously, if $\psi \in \operatorname{ran} P_T(\widetilde{\Omega})$, then $\psi(x) = 0$ for μ-a.e. $x \notin \widetilde{\Omega}$. If $\Omega \cap \widetilde{\Omega} = \emptyset$, then $\overline{\varphi(x)}\psi(x) = 0$ μ-a.e., which implies $\langle \varphi, \psi \rangle_{L^2(X,\mu)} = 0$. This shows the orthogonality $\operatorname{ran} P_T(\Omega) \perp \operatorname{ran} P_T(\widetilde{\Omega})$. The second identity follows from

$$\begin{aligned} P_T(\Omega) + P_T(\widetilde{\Omega}) &= 1_\Omega(T) + 1_{\widetilde{\Omega}}(T) = (1_\Omega + 1_{\widetilde{\Omega}})(T) \\ &= 1_{\Omega \cup \widetilde{\Omega}}(T) = P_T(\Omega \cup \widetilde{\Omega}), \end{aligned}$$

and the third identity follows from the general properties of projectors: if P and P' are orthogonal projectors with $\operatorname{ran} P \perp \operatorname{ran} P'$, then $\operatorname{ran}(P + P') = \operatorname{ran} P + \operatorname{ran} P'$.

[5]Recall that if A_n, A are bounded linear operators in $\mathcal{H}$, then $A = \mathrm{s} - \lim A_n$ means that $Av = \lim A_n v$ for any $v \in \mathcal{H}$ (one says that A_n converges strongly to A).

(b.2) Using (a.1) and (b) we have

$$\begin{aligned}\operatorname{ran} P_T(\widetilde{\Omega}) &= \operatorname{ran}\big(P_T(\Omega) + P_T(\widetilde{\Omega}\setminus\Omega)\big)\\ &= \operatorname{ran} P_T(\Omega) + \operatorname{ran} P_T(\widetilde{\Omega}\setminus\Omega) \supset \operatorname{ran} P_T(\Omega).\end{aligned}$$

(b.3) follows from (b.1) and the observation that $I = 1_{\mathbb{R}}(T) = P_T(\mathbb{R})$.

(c) The condition $P_T\big((a,b)\big) = 0$ is equivalent to $1_{(a,b)}\circ h = 0$ μ-e.a., which in turn means that $(a,b)\cap \operatorname{ess\,ran} h = \emptyset$, and it remains to recall that $\operatorname{ess\,ran} h = \operatorname{spec} M_h$.

(d) For $\varphi \in L^2(X,\mu)$ one has $h\varphi = \lambda\varphi$ if and only if $\varphi(x) = 0$ for $h(x) \neq \lambda$, i.e. if $\varphi = (1_{\{\lambda\}}\circ h)\varphi$, which exactly means $\varphi = P_T(\{\lambda\})\varphi$. As $P_T(\{\lambda\})$ is an orthogonal projector, the set of such φ coincides with $\operatorname{ran} P_T(\{\lambda\})$.

(e) The condition $P_T\big((\lambda-\varepsilon,\lambda+\varepsilon)\big) \neq 0$ for all $\varepsilon > 0$ is equivalent $1_{(\lambda-\varepsilon,\lambda+\varepsilon)}\circ h \neq 0$ for all $\varepsilon > 0$, which in turn means that

$$\mu\{x \in X : h(x) \in (\lambda-\varepsilon,\lambda+\varepsilon)\} > 0 \text{ for all } \varepsilon > 0,$$

which is exactly the condition $\lambda \in \operatorname{ess\,ran} h \equiv \operatorname{spec} M_h$.

(f.1) Let $\phi \in L^2(X,\mu)$ and $\varphi := P_T\big((\lambda-\varepsilon,\lambda+\varepsilon)\big)\phi = (1_{(\lambda-\varepsilon,\lambda+\varepsilon)}\circ h)\phi$, then $\varphi(s,n) = 0$ for all $(s,n)\in X$ with $s \notin (\lambda-\varepsilon,\lambda+\varepsilon)$. Therefore, if $x := (s,n) \in X$ with $\varphi(x) \neq 0$, then $h(x) \equiv s \in (\lambda-\varepsilon,\lambda+\varepsilon)$, i.e. one has the inequalities $|h(x)-\lambda| < \varepsilon$, as then $|h(x)| \le |\lambda| + \varepsilon$. It follows that

$$\begin{aligned}\int_X \big|h(x)\varphi(x)\big|^2\,\mathrm{d}\mu(x) &= \int_{\{x\in X:\,\varphi(x)\neq 0\}} \big|h(x)\varphi(x)\big|^2\,\mathrm{d}\mu(x)\\ &\le \big(|\lambda|+\varepsilon\big)^2 \int_{\{x\in X:\,\varphi(x)\neq 0\}} \big|\varphi(x)\big|^2\,\mathrm{d}\mu(x) = \big(|\lambda|+\varepsilon\big)^2\|\varphi\|^2 < \infty,\end{aligned}$$

which shows that $\varphi \in D(M_h)$. Similarly,

$$\begin{aligned}\|(M_h-\lambda)\varphi\|^2 &= \int_X \big|h(x)-\lambda\big|\,\big|\varphi(x)\big|^2\,\mathrm{d}\mu(x)\\ &= \int_{\{x\in X:\,\varphi(x)\neq 0\}} \big|h(x)-\lambda\big|\,\big|\varphi(x)\big|^2\,\mathrm{d}\mu(x)\\ &\le \varepsilon^2 \int_{\{x\in X:\,\varphi(x)\neq 0\}} \big|\varphi(x)\big|^2\,\mathrm{d}\mu(x)\\ &= \varepsilon^2 \int_{\{x\in X:\,\varphi(x)\neq 0\}} \big|\phi(x)\big|^2\,\mathrm{d}\mu(x)\\ &\le \varepsilon^2 \int_X \big|\phi(x)\big|^2\,\mathrm{d}\mu(x) \le \varepsilon^2\|\phi\|^2.\end{aligned}$$

(f.2) Let $\phi \in D(M_h)$ and $\varphi := \big[I - P_T\big((\lambda-\varepsilon,\lambda+\varepsilon)\big)\big]\phi = (1 - 1_{(\lambda-\varepsilon,\lambda+\varepsilon)}\circ h)\phi$, then $\varphi(s,n) = 0$ for all $(s,n)\in X$ with $s \in (\lambda-\varepsilon,\lambda+\varepsilon)$. It follows that if $(s,n)\in X$

with $\varphi(s,n)\neq 0$, then $h(s,n)\equiv s\notin(\lambda-\varepsilon,\lambda+\varepsilon)$, i.e. $|h(s,n)-\lambda|\geq\varepsilon$, and then

$$\begin{aligned}\|(M_h-\lambda)\varphi\|^2 &= \int_X \big|h(x)-\lambda\big|\,\big|\varphi(x)\big|^2\,\mathrm{d}\mu(x)\\ &= \int_{\{x\in X:\,\varphi(x)\neq 0\}} \big|h(x)-\lambda\big|\,\big|\varphi(x)\big|^2\,\mathrm{d}\mu(x)\\ &\geq \varepsilon^2\int_{\{x\in X:\,\varphi(x)\neq 0\}} \big|\varphi(x)\big|^2\,\mathrm{d}\mu(x)=\varepsilon^2\|\varphi\|^2.\end{aligned}$$

(f.3) Let $\phi\in D(M_h)$ and $\varphi:=P_T\big[\lambda,\infty)\big)\phi\equiv(1-1_{[\lambda,\infty)}\circ h)\phi$, then $\varphi(s,n)=0$ for all $s<\lambda$, and then $h(s,n)\equiv s\geq\lambda$ for all $(s,n)\in X$ with $\varphi(s,n)\neq 0$, and

$$\begin{aligned}\langle\varphi,M_h\varphi\rangle &= \int_X h(x)|\varphi(x)|^2\,\mathrm{d}\mu(x)\\ &= \int_{\{x\in X:\,\varphi(x)\neq 0\}} h(x)|\varphi(x)|^2\,\mathrm{d}\mu(x)\\ &\geq \lambda\int_{\{x\in X:\,\varphi(x)\neq 0\}} |\varphi(x)|^2\,\mathrm{d}\mu(x)=\lambda\|\varphi\|^2.\end{aligned}$$

(g) Remark that $\|1_{(\lambda-\varepsilon,\lambda+\varepsilon)}\|_\infty\leq 1$ and $1_{(\lambda-\varepsilon,\lambda+\varepsilon)}\to 1_{\{\lambda\}}$ pointwise. By the functional calculus (Theorem 3.11) it follows that

$$\operatorname*{s-lim}_{\varepsilon\to 0^+} P_T\big((\lambda-\varepsilon,\lambda+\varepsilon)\big)=\operatorname*{s-lim}_{\varepsilon\to 0^+} 1_{(\lambda-\varepsilon,\lambda+\varepsilon)}(T)=1_{\{\lambda\}}(T)=P_T(\{\lambda\}). \qquad \square$$

Remark 3.22. In order to have a better "feeling" of the spectral projectors, let us consider the most simple case. Let $\Omega:=(a,b)\subset\mathbb{R}$ and assume that the spectrum of T in Ω consists of $N<\infty$ eigenvalues $a<\lambda_1<\cdots<\lambda_N<b$. By Prop. 3.21(b,c) one has

$$\begin{aligned}P_T(\Omega)\big) &= \underbrace{P_T\big((a,\lambda_1)\big)}_{=0}+P_T(\{\lambda_1\})+\underbrace{P_T\big((\lambda_1,\lambda_2)\big)}_{=0}+P_T(\{\lambda_2\})+\dots\\ &=\sum_{k=1}^N P_T(\{\lambda_k\}).\end{aligned}$$

By (d), each operator $P_T(\{\lambda_n\})$ is the orthogonal projector on $\ker(T-\lambda_k)$: if $n_k:=\dim\ker(T-\lambda_k)\in\mathbb{N}\cup\{\infty\}$ and $(\varphi_j^k)_{j\in\{1,\dots,n_k\}}$ is an orthonormal basis in $\ker(T-\lambda_k)$, then for any $v\in\mathcal{H}$ one has

$$P_T(\{\lambda_k\})\,v=\sum_{j=1}^{n_k}\langle\varphi_j^k,v\rangle\varphi_j^k,\quad P_T(\Omega)\,v=\sum_{k=1}^N\sum_{j=1}^{n_k}\langle\varphi_j^k,v\rangle\varphi_j^k,$$

and $\operatorname{span}\big((\varphi_j^k)_{k\in\{1,N\},\,j\in\{1,\dots,n_k\}}\big)=\operatorname{ran}P_T(\Omega)$, in particlar, $\dim\operatorname{ran}P_T(\Omega)$ is exactly the number of eigenvalues of T in Ω.

Recall that Weyl sequences were defined in Proposition 2.5.

Corollary 3.23 (Spectrum and Weyl sequences). *Let T be self-adjoint, then $\lambda \in \operatorname{spec} T$ if and only if there exists a Weyl sequence for λ.*

Proof. If there exist a Weyl sequence for λ, then $\lambda \in \operatorname{spec} T$ by Proposition 2.5.

On the other hand, let $\lambda \in \operatorname{spec} T$. For $n \in \mathbb{N}$ denote $I_n := (\lambda - 2^{-n}, \lambda + 2^{-n})$, then by (d) one has $P_T(I_n) \neq 0$. By Prop. 3.21(a) one can find u_n with $\|u_n\| = 1$ and $u_n = P_T(I_n)u_n$. By Prop. 3.21(f) one has $u_n \in D(T)$ and $\|(T - \lambda)u_n\| \leq 2^{-n}\|u_n\|$, which shows that (u_n) is a Weyl sequence for λ. □

Proposition 3.24 (Isolated points of the spectrum). *Let λ be an isolated point of the spectrum of a self-adjoint operator T, then:*

(a) λ is an eigenvalue of T,

(b) there exists $c > 0$ such that $\|(T - \lambda)u\| \geq c\|u\|$ for all $u \in D(T) \cap \ker(T - \lambda)^\perp$.

Proof. (a) By Prop. 3.21(d) it is sufficient to show that $P_T(\{\lambda\}) \neq 0$. Let $\varepsilon > 0$ such that $(\lambda - \varepsilon, \lambda + \varepsilon) \cap \operatorname{spec} T = \{\lambda\}$, then

$$(\lambda - \varepsilon, \lambda) \cap \operatorname{spec} T = (\lambda, \lambda + \varepsilon) \cap \operatorname{spec} T = \emptyset.$$

By Prop. 3.21(e) we have $P_T\big((\lambda - \varepsilon, \lambda + \varepsilon)\big) \neq 0$, and using to Prop. 3.21(b,c) we arrive at

$$P_T(\{\lambda\}) = \underbrace{P_T\big((\lambda - \varepsilon, \lambda)\big)}_{=0} + P_T(\{\lambda\}) + \underbrace{P_T\big((\lambda, \lambda + \varepsilon)\big)}_{=0} = P_T\big((\lambda - \varepsilon, \lambda + \varepsilon)\big) \neq 0.$$

(b) Recall that if P is an orthogonal projector on some closed subspace V, then $P^\perp := I - P$ is the orthogonal projection on $V^\perp$. As already seen, $P := P_T(\{\lambda\})$ is the orthogonal projector on $\ker(T - \lambda)$, then $P^\perp := I - P_T(\{\lambda\})$ is the orthogonal projector on $\ker(T-\lambda)^\perp$. Therefore, for any $u \in D(T)$ one has $u \perp \ker(T-\lambda)$ if and only if $u = P^\perp u$. As seen in (a), for some $\varepsilon > 0$ one has $P_T(\{\lambda\}) = P_T\big((\lambda-\varepsilon, \lambda+\varepsilon)\big)$.

Now let $u \in D(T)$ with $u \perp \ker(T-\lambda)$, then $u = P^\perp u \equiv \big[I - P_T\big((\lambda-\varepsilon, \lambda+\varepsilon)\big)\big]u$, and $\|(T - \lambda)u\| \geq \varepsilon\|u\|$ due to Prop. 3.21(e.2). □

4 Classification of spectra and perturbations

4.1 Discrete and essential spectra

In this section, let T be a self-adjoint operator in a separable Hilbert space $\mathcal{H}$. Up to now we just distinguished between the whole spectrum ($\operatorname{spec} T$) and the point spectrum ($\operatorname{spec}_{\mathrm{p}} T$), i.e. the set of eigenvalues. Let us introduce another classification of spectra, which is useful when studying various perturbations.

Definition 4.1 (Discrete spectrum, essential spectrum). The *discrete spectrum* $\operatorname{spec}_{\mathrm{disc}} T$ of T is defined by

$$\operatorname{spec}_{\mathrm{disc}} T := \big\{\lambda \in \operatorname{spec} T : \exists \varepsilon > 0 \text{ with } \dim \operatorname{ran} P_T\big((\lambda-\varepsilon, \lambda+\varepsilon)\big) < \infty\big\},$$

and *essential spectrum* $\operatorname{spec}_{\mathrm{ess}} T$ is

$$\operatorname{spec}_{\mathrm{ess}} T := \operatorname{spec} T \setminus \operatorname{spec}_{\mathrm{disc}} T.$$

By definition, the discrete spectrum and the essential spectrum are disjoint. Let us find equivalent characterizations for both of them.

Theorem 4.2 (Characterization of the discrete spectrum).
Let $\lambda \in \mathbb{R}$, then the following two conditions are equivalent:

(a) $\lambda \in \operatorname{spec}_{\text{disc}} T$,

(b) λ *is an eigenvalue of* T *of finite multiplicity* *<u>and</u>* *an isolated point of* $\operatorname{spec} T$.
(Such eigenvalues are usually called discrete eigenvalues*).*

Proof. (a⇒b) Let $\lambda \in \operatorname{spec}_{\text{disc}} T$, then:

- for any $\varepsilon > 0$ one has $P_T\big((\lambda-\varepsilon, \lambda+\varepsilon)\big) \neq 0$ by Prop. 3.21(e),
- for some $\varepsilon_0 > 0$ one has $N := \dim \operatorname{ran} P_T\big(\lambda-\varepsilon_0, \lambda+\varepsilon_0)\big) < \infty$.

For $0 < \varepsilon < \varepsilon' < \varepsilon_0$ one has $\operatorname{ran} P_T\big((\lambda-\varepsilon, \lambda+\varepsilon)\big) \subset \operatorname{ran} P_T\big((\lambda-\varepsilon', \lambda+\varepsilon')\big)$ (Prop. 3.21(b)). In particular, the function

$$\kappa : (0, \varepsilon_0) \ni \varepsilon \mapsto \dim \operatorname{ran} P_T\big((\lambda-\varepsilon, \lambda+\varepsilon)\big) \in \{1, \dots, N\}$$

is non-decreasing. It follows that there exist $\varepsilon_1 \in (0, \varepsilon_0)$ and $k \in \{1, \dots, N\}$ such that $\kappa(\varepsilon) = k$ for all $\varepsilon \in (0, \varepsilon_1)$, and then it follows that $V := \operatorname{ran} P_T\big((\lambda-\varepsilon, \lambda+\varepsilon)\big)$ is the same for all $\varepsilon \in (0, \varepsilon_1)$ with $\dim V = k \geq 1$.

Let P be the orthogonal projector on V, then $P \neq 0$. As orthogonal projectors are unquely determined by their ranges, it follows $P = P_T\big((\lambda-\varepsilon, \lambda+\varepsilon)\big)$ for all $\varepsilon \in (0, \varepsilon_1)$, in particular, $P = \operatorname{s-lim}_{\varepsilon \to 0+} P_T\big((\lambda-\varepsilon, \lambda+\varepsilon)\big) = P_T(\{\lambda\})$ by Prop. 3.21, and then $\lambda \in \operatorname{spec}_{\text{p}} T$ by Prop. 3.21(d).

Now pick any $\varepsilon \in (0, \varepsilon_1)$, then due to $P_T\big((\lambda-\varepsilon, \lambda+\varepsilon)\big) = P_T(\{\lambda\})$ and

$$P_T\big((\lambda-\varepsilon, \lambda+\varepsilon)\big) = P_T(\{\lambda\}) + P_T\big((\lambda-\varepsilon, \lambda) \cup (\lambda, \lambda+\varepsilon)\big)$$

we obtain $P_T\big((\lambda-\varepsilon, \lambda) \cup (\lambda, \lambda+\varepsilon)\big) = 0$. Using Prop. 3.21(b) again we obtain

$$\{0\} = \operatorname{ran} P_T\big((\lambda-\varepsilon, \lambda) \cup (\lambda, \lambda+\varepsilon)\big) = \operatorname{ran} P_T\big((\lambda-\varepsilon, \lambda)\big) + \operatorname{ran} P_T\big((\lambda, \lambda+\varepsilon)\big),$$

which means $\operatorname{ran} P_T\big((\lambda-\varepsilon, \lambda)\big) = \operatorname{ran} P_T\big((\lambda, \lambda+\varepsilon)\big) = \{0\}$, and by Prop. 3.21(c) this means $\operatorname{spec} T \cap (\lambda-\varepsilon, \lambda) = \operatorname{spec} T \cap (\lambda, \lambda+\varepsilon) = \emptyset$, i.e. λ is an isolated point of $\operatorname{spec} T$.

(b⇒a) Let λ be an eigenvalue of finite multiplicity and an isolated point of the spectrum. One has $\dim \operatorname{ran} P_T(\{\lambda\}) = \dim \ker(T-\lambda) < \infty$ by Prop. 3.21(d), and for some $\varepsilon > 0$ one has $(\lambda-\varepsilon, \lambda) \cap \operatorname{spec} T = (\lambda, \lambda+\varepsilon) \cap \operatorname{spec} T = \emptyset$. Using Prop. 3.21(c) we obtain

$$P_T\big((\lambda-\varepsilon, \lambda+\varepsilon)\big) = \underbrace{P_T\big((\lambda-\varepsilon, \lambda)\big)}_{=0} + P_T(\{\lambda\}) + \underbrace{P_T\big((\lambda, \lambda+\varepsilon)\big)}_{=0} = P_T(\{\lambda\}),$$

and $\dim \operatorname{ran} P_T\big((\lambda-\varepsilon, \lambda+\varepsilon)\big) = \dim \operatorname{ran} P_T(\{\lambda\}) < \infty$. □

Theorem 4.3 (Characterization of the essential spectrum).
Let $\lambda \in \operatorname{spec} T$, then $\lambda \in \operatorname{spec}_{\text{ess}} T$ if and only if at least one of the following three conditions holds:

- $\lambda \notin \operatorname{spec}_{\mathrm{p}} T$,
- λ *is an accumulation point of* $\operatorname{spec} T$,
- $\dim\ker(T-\lambda) = \infty$.

Furthermore, $\operatorname{spec}_{\mathrm{ess}} T$ *is a closed set.*

Proof. The first part just describes the points of the spectrum which are not eigenvalues of finite multiplicity or not isolated. For the second part we note that $\operatorname{spec}_{\mathrm{ess}} T$ is obtained from the closed set $\operatorname{spec} T$ by removing some isolated points. Each isolated point is a relatively open subset, and any set of isolated points is again relatively open (as the union of arbitrarily many open sets is open), so removing such a set from the closed set $\operatorname{spec} T$ gives a closed set. $\square$

Example 4.4 (Essential spectrum for compact operators). If T is a compact self-adjoint operator in an infinite-dimensional space $\mathcal{H}$, then one easily sees that $\operatorname{spec}_{\mathrm{ess}} T = \{0\}$.

Namely, by Theorem 2.24 for any $\varepsilon > 0$ the set $\operatorname{spec} T \setminus (-\varepsilon, \varepsilon)$ consists of a finite number of eigenvalues of finite multiplicity, hence we have: $\operatorname{spec}_{\mathrm{ess}} T \setminus (-\varepsilon, \varepsilon) = \emptyset$ and $\dim\operatorname{ran} P_T\big(\mathbb{R} \setminus (-\varepsilon, \varepsilon)\big) < \infty$ (see Remark 3.22). On the other hand, by Prop. 3.21(b.3) one has

$$\underbrace{\dim\operatorname{ran} P_T\big(\mathbb{R} \setminus (-\varepsilon, \varepsilon)\big)}_{<\infty} + \dim\operatorname{ran} P_T\big((-\varepsilon, \varepsilon)\big) = \dim\operatorname{ran} \mathcal{H} = \infty$$

which gives $\dim\operatorname{ran} P_T\big((-\varepsilon, \varepsilon)\big) = \infty$ for any $\varepsilon > 0$, i.e. $0 \in \operatorname{spec}_{\mathrm{ess}} T$.

Definition 4.5 (Purely discrete and purely essential spectra). Let T be a self-adjoint operator and $I \subset \mathbb{R}$ some open interval. We say that T has

- a *purely discrete spectrum in* I if $\operatorname{spec}_{\mathrm{ess}} T \cap I = \emptyset$,
- a *purely essential spectrum in* I if $\operatorname{spec}_{\mathrm{disc}} T \cap I = \emptyset$.

If $\operatorname{spec}_{\mathrm{ess}} T = \emptyset$, then we say simply that the spectrum of T is *purely discrete*, and for $\operatorname{spec}_{\mathrm{disc}} T = \emptyset$ we says that the spectrum of T is *purely essential.*

Example 4.6. The spectrum of the free Laplacian in $L^2(\mathbb{R}^d)$ has $[0, +\infty)$, and it is purely essential, as it has no isolated points.

Example 4.7. If T is a self-adjoint operator with compact resolvent, it has no essential spectrum (each point of the spectrum is an eigenvalue of finite multilplicity, and the spectrum has no accumulation points). In fact, one can easily show that a self-adjoint operator has purely discrete spectrum if and only if it has compact resolvent.

The main difference between the discrete and the essential spectra comes from their behavior with respect to perturbations. This will be discussed in the following sections.

4.2 Relatively bounded perturbations

We have seen since the beginning of the course that one needs to pay attention to the domains of unbounded operators. The aim of the present subsection is to describe some classes of operators in which such problems can be avoided using the idea that if one perturbs a "good" operator by adding a "small" operator, then the result is again "goood".

Recall that a linear operator T is essentially self-adjoint if its closure is a self-adjoint operator (or , equivalently, that the adjoint T^* is symmetric). In addition, we say that a linear operator T is essentially self-adjoint on a subspace $\mathcal{D} \subset D(T)$, if the closure of the restriction $T|_{\mathcal{D}}$ of T on $\mathcal{D}$ is a self-adjoint operator. We already know that an essentially self-adjoint operator has a unique self-adjoint extension.

The following simple result will be used several times:

Lemma 4.8. *Let T be a symmetric operator and $\lambda \in \mathbb{R} \setminus \{0\}$, then:*

(a) $\overline{\operatorname{ran}(T + i\lambda)} = \operatorname{ran}(\overline{T} + i\lambda)$,

(b) $\operatorname{ran}(T + i\lambda)$ *is closed if and only if T is closed.*

Proof. It is clear that (b) is a consequence of (a). To prove (a) we remark first that for any $x \in D(T)$ we have:

$$\begin{aligned} \big\|(T + i\lambda)x\big\|^2 &= \big\langle (T + i\lambda)x, (T \pm i\lambda)x \big\rangle \\ &= \langle Tx, Tx\rangle + \lambda^2 \langle x, x\rangle + i\lambda\big(\underbrace{\langle Tx, x\rangle - \langle x, Tx\rangle}_{=0 \text{ as } T \text{ is symmetric}} \big) \\ &= \|Tx\|^2 + \lambda^2 \|x\|^2. \end{aligned} \tag{4.1}$$

Let $y \in \overline{\operatorname{ran}(T + i\lambda)}$ and $y_n \in \operatorname{ran}(T + i\lambda)$ with $y_n \to y$, then (y_n) is a Cauchy sequence in $\mathcal{H}$. One has $y_n = (T + i\lambda)x_n$ with some $x_n \in D(T)$, and due to (4.1) the sequence (x_n) is also Cauchy in $\mathcal{H}$ and converges to some $x \in D(T)$. As T is closable, then the operator $T + i\lambda$ is also closable, which shows that $x \in D(\overline{T})$ and $y = \overline{(T + i\lambda)}x \equiv (\overline{T} + i\lambda)x$, i.e. $y \in \operatorname{ran}(\overline{T} + i\lambda)$.

Now let $y \in \operatorname{ran}(\overline{T} + i\lambda)$, then $y = (\overline{T} + i\lambda)x$ for some $x \in D(\overline{T})$. This means that there are $x_n \in D(T)$ such that $x \to x$ in $\mathcal{H}$ and $y_n := (T + i\lambda)x_n \to y$ in $\mathcal{H}$ Due to $y_n \in D(T)$ one obtains $y \in \overline{\operatorname{ran}(T + i\lambda)}$. $\square$

Theorem 4.9 (Self-adjointness criterion). *Let T be a closed densely defined symmetric operator in a Hilbert space $\mathcal{H}$ and $\lambda > 0$, then the following three assertions are equivalent:*

(A) T is self-adjoint,

(B) $\ker(T^* + i\lambda) = \ker(T^* - i\lambda) = \{0\}$,

(C) $\operatorname{ran}(T + i\lambda) = \operatorname{ran}(T - i\lambda) = \mathcal{H}$.

Proof. (A⇒B) is clear, as a self-adjoint operator cannot have non-real eigenvalues.

(B⇒C). One has $\overline{\operatorname{ran}(T \mp i\lambda)} = \ker(T^* \pm i\lambda)^\perp = \mathcal{H}$ (Prop. 2.16), and the subspaces $\operatorname{ran}(T \pm i\lambda)$ are closed by Lemma 4.8(b).

(C⇒A). Let $\varphi \in D(T^*)$. Due to the surjectivity of $T - i\lambda$ one can find $\psi \in D(T)$ with $(T - i\lambda)\psi = (T^* - i\lambda)\varphi$. As $T \subset T^*$, we have $(T^* - i\lambda)(\psi - \varphi) = 0$. We have

$$\ker(T^* - i\lambda) = \operatorname{ran}(T + i\lambda)^\perp = \mathcal{H}^\perp = \{0\},$$

which means that $\varphi = \psi \in D(T)$ and then $T^*\varphi = T\varphi$. This shows $T^* \subset T$, and T is self-adjoint. □

By combining Theorem 4.9 with Lemma 4.8(a) we immediately obtain:

Theorem 4.10 (Essential self-adjointness criterion). *Let T be a densely defined symmetric operator in a Hilbert space $\mathcal{H}$ and $\lambda > 0$, then the following three assertions are equivalent:*

(A) *T is essentially self-adjoint,*

(B) $\ker(T^* + i\lambda) = \ker(T^* - i\lambda) = \{0\}$,

(C) $\operatorname{ran}(T + i\lambda)$ *and* $\operatorname{ran}(T - i\lambda)$ *are dense in* $\mathcal{H}$.

Remark that for semibounded operators we have an alternative version with less conditions to check, which is proved in a very similar way (exercise):

Theorem 4.11 (Essential self-adjointness for semibounded operators). *Let T be a densely symmetric operator in a Hilbert space $\mathcal{H}$ with $T \geq 0$ and let $a > 0$, then:*

- $\overline{\operatorname{ran}(T + a)} = \operatorname{ran}(\overline{T} + a)$,
- $\operatorname{ran}(T + a)$ *is closed if and only if T is closed,*

and the following three assertions are equivalent:

(a) *T is essentially self-adjoint,*

(b) $\ker(T^* + a) = \{0\}$,

(c) $\operatorname{ran}(T + a)$ *is dense in* $\mathcal{H}$.

Now we would like to apply the above assertions to the study of some perturbations of self-adjoint operators.

Definition 4.12 (Relative boundedness). Let A be a self-adjoint operator in a Hilbert space $\mathcal{H}$ and B be a linear operator with $D(A) \subset D(B)$. Assume that there exist $a, b > 0$ such that

$$\|Bf\| \leq a\|Af\| + b\|f\| \text{ for all } f \in D(A),$$

then B is called *relatively bounded* with respect to A or, for short, *A-bounded*. The infimum of all possible values a is called the *relative bound* of B with respect to A. If the relative bound is equal to 0 (i.e. if for any $a > 0$ one can find $b > 0$ such that the above estimate holds), then B is called *infinitesimally small* with respect to A. □

Theorem 4.13 (Kato-Rellich). *Let A be a self-adjoint operator in $\mathcal{H}$ and let B be a symmetric operator in $\mathcal{H}$ which is A-bounded with a relative bound < 1, then*

(i) the operator $A + B$ with domain $D(A + B) = D(A)$ is self-adjoint.

Moreover:

(ii) if A is essentially self-adjoint on some subspace $\mathcal{D} \subset D(A)$, then $A + B$ is essentially self-adjoint on $\mathcal{D}$,

(iii) if A is semibounded from below, then also $A + B$ is semibounded from below.

Proof. We will only prove (i), and (ii) and (iii) will be discussed as exercises.

By assumption, one can find $a \in (0, 1)$ and $b > 0$ such that

$$\|Bu\| \le a\|Au\| + b\|u\| \text{ for all } \quad u \in D(A). \tag{4.2}$$

Remark that $A + B$ with domain $D(A + B) = D(A)$ is at least symmetric. The proof of (i) is now decomposed in three steps.

Step 1. Let $\lambda > 0$, then as in (4.1) one obtains

$$\|(A + B \pm i\lambda)u\|^2 = \|(A + B)u\|^2 + \lambda^2\|u\|^2 \text{ for all } u \in D(A).$$

Therefore, for all $u \in D(A)$ one can estimate

$$\begin{aligned} 2\big\|(A + B \pm i\lambda)u\big\| &\ge \big\|(A + B)u\big\| + \lambda\|u\| \\ &\ge \|Au\| - \|Bu\| + \lambda\|u\| = (1 - a)\|Au\| + (\lambda - b)\|u\|. \end{aligned} \tag{4.3}$$

Let us pick some $\lambda > b$.

Step 2. Let us show that $A + B$ with $D(A + B) = D(A)$ is a closed operator.

Let $(u_n) \subset D(A)$ and $f_n := (A + B)u_n$ such that both u_n and f_n converge in $\mathcal{H}$. By (4.3), the sequence Au_n is Cauchy. As A is closed, the sequence u_n converge to some $u \in D(A)$ and Au_n converge to Au. By (4.2), the sequence Bu_n is Cauchy and converges to some $v \in \mathcal{H}$. For any $h \in D(A)$ one has $\langle v, h\rangle = \lim\langle Bu_n, h\rangle = \lim\langle u_n, Bh\rangle = \langle u, Bh\rangle = \langle Bu, h\rangle$. As $D(A)$ is dense, it follows that $v = Bu$, i.e. $\lim Bu_n = Bu$. So finally $(A + B)u_n = Au_n + Bu_n$ converge to $(A + B)u$, which shows that $A + B$ is closed.

Step 3. Let us show that the operators $A + B \pm i\lambda : D(A) \to \mathcal{H}$ are bijective at least for large λ. As previously, for any $u \in D(A)$ we have $\|(A \pm i\lambda)u\|^2 = \|Au\|^2 + \lambda^2\|u\|^2$ and then

$$\begin{aligned} \|Bu\| &\le a\|Au\| + b\|u\| \\ &\le a\big\|(A \pm i\lambda)u\big\| + \frac{b}{|\lambda|}\big\|(A \pm i\lambda)u\big\| = \big(a + \frac{b}{|\lambda|}\big)\big\|(A \pm i\lambda)u\big\|. \end{aligned}$$

Let $v \in \mathcal{H}$, then $u := (A \pm i\lambda)^{-1}v \in D(A)$, and the preceding inequality gives

$$\|B(A \pm i\lambda)^{-1}v\| \leq \Big(a + \frac{b}{|\lambda|}\Big)\|v\|.$$

As $a \in (0,1)$, we can choose λ sufficiently large to have $a + b/|\lambda| < 1$, then it follows that $\big\|B(A \pm i\lambda)^{-1}\big\| < 1$.

Now we represent $A + B \pm i\lambda = \big(I + B(A \pm i\lambda)^{-1}\big)(A \pm i\lambda)$. Recall that the operators $A \pm i\lambda : D(A) \to \mathcal{H}$ are bijective. If λ is sufficiently large, then the operators $I + B(A \pm i\lambda)^{-1} : \mathcal{H} \to \mathcal{H}$ are bijections due to $\big\|B(A \pm i\lambda)^{-1}\big\| < 1$. Therefore, the operators $A+B \pm i\lambda$ are bijective, in particular, $\operatorname{ran}(A+B \pm i\lambda) = \mathcal{H}$. By Theorem 4.9 the operator $A + B$ is self-adjoint. $\square$

4.3 Essential self-adjointness of Schrödinger operators

The Kato-Rellich theorem is one of the tools used to simplify the consideration of the Schrödinger operators.

Theorem 4.14 (Kato-Rellich for Schrödinger operators). *Let $d \in \mathbb{N}$,*

$$p = 2 \text{ for } d \le 3, \qquad p > \frac{d}{2} \text{ for } d \ge 4,$$

and $V \in L^p(\mathbb{R}^d) + L^\infty(\mathbb{R}^d)$ real-valued. Then the operator $T = -\Delta + V$ with domain $D(T) = H^2(\mathbb{R}^d)$ is self-adjoint in $\mathcal{H} = L^2(\mathbb{R}^d)$. Moreover, it is essentially self-adjoint on $C_c^\infty(\mathbb{R}^d)$ and semibounded from below.

Proof. We give the proof only for the dimension $d \le 3$ (some comments on $d \ge 4$ are given at the end). For all $f \in C_c^\infty(\mathbb{R}^d)$, $\lambda > 0$, $x \in \mathbb{R}^d$ we have, using the Fourier inversion formula and the Cauchy-Schwarz inequality,

$$\begin{aligned}
\big|f(x)\big| &= \left|\frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{i\xi\cdot x} \widehat{f}(\xi)\,\mathrm{d}\xi\right| \le \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \big|\widehat{f}(\xi)\big|\,\mathrm{d}\xi \\
&= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \frac{1}{|\xi|^2+\lambda} \big(|\xi|^2+\lambda\big)\big|\widehat{f}(\xi)\big|\,\mathrm{d}\xi \\
&\le \frac{1}{(2\pi)^{d/2}} \left\|\frac{1}{|\xi|^2+\lambda}\right\|_{L^2} \left\|\big(|\xi|^2+\lambda\big)\widehat{f}\right\|_{L^2} \\
&\le \frac{1}{(2\pi)^{d/2}} \left\|\frac{1}{|\xi|^2+\lambda}\right\|_{L^2} \Big(\underbrace{\big\||\xi|^2\widehat{f}\big\|_{L^2}}_{=\|\Delta f\|^2_{L^2}} + \lambda \underbrace{\|\widehat{f}\,\|_{L^2}}_{=\|f\|^2_{L^2}}\Big) \\
&= a_\lambda \|\Delta f\|_{L^2} + b_\lambda \|f\|_{L^2}
\end{aligned}$$

with

$$a_\lambda = \frac{c_\lambda}{(2\pi)^{d/2}}, \quad b_\lambda = \frac{\lambda c_\lambda}{(2\pi)^{d/2}}, \quad c_\lambda := \left\|\frac{1}{|\xi|^2+\lambda}\right\|_{L^2}.$$

Remark that the condition $c_\lambda < \infty$ is equivalent to $d \le 3$. Therefore,

$$\|f\|_\infty \le a_\lambda \|\Delta f\| + b_\lambda \|f\|, \tag{4.4}$$

which extends by density to all $f \in H^2(\mathbb{R}^d)$.

Let $V = V_2 + V_\infty$ with $V_2 \in L^2(\mathbb{R}^d)$ and $V_\infty \in L^\infty(\mathbb{R}^d)$. Using (4.4), for any $f \in H^2(\mathbb{R}^d)$ and any $\lambda > 0$ we have

$$\begin{aligned}
\|Vf\|_{L^2} &= \|(V_2+V_\infty)f\|_{L^2} \le \|V_2 f\|_{L^2} + \|V_\infty f\|_{L^2} \\
&\le \|V_2\|_{L^2}\|f\|_\infty + \|V_\infty\|_\infty \|f\|_{L^2} \le \widetilde{a}_\lambda \|\Delta f\| + \widetilde{b}_\lambda \|f\|, \\
\widetilde{a}_\lambda &:= \|V_2\|_{L^2} a_\lambda, \quad \widetilde{b}_\lambda = \|V_2\|_{L^2} b_\lambda + \|V_\infty\|_\infty.
\end{aligned}$$

As $\lim_{\lambda\to+\infty} a_\lambda = 0$, the operator of multilplication by V is infinitesimally small with respect to the free Laplacian, and all the claims follow from the Kato-Rellich theorem (Theorem 4.13).

The above proof uses that the function $\mathbb{R}^d \ni \xi \mapsto \big(|\xi|^2+\lambda\big)^{-1}$ belongs to $L^2(\mathbb{R}^d)$ for $d \le 3$. For $d \ge 4$ it does not work, and one must use additional results stating that $H^k(\mathbb{R}^d)$ are continuously embedded in $L^q(\mathbb{R}^d)$ for suitable combinations of k, q, d. $\square$

Example 4.15 (Coulomb potential). Consider $\mathcal{H} = L^2(\mathbb{R}^3)$ and $T = -\Delta + V$ with the Coulomb potential $V(x) = q/|x|$, $q \in \mathbb{R}$. Let Ω be any ball centered at the origin, then

$$1_\Omega V \in L^2(\mathbb{R}^3), \quad (1-1_\Omega)V \in L^\infty(\mathbb{R}^3), \quad V = 1_\Omega V + (1-1_\Omega)V \in L^2(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$$

and by Theorem 4.14 the operator T is self-adjoint on $H^2(\mathbb{R}^3)$ and essentially self-adjoint on $C_c^\infty(\mathbb{R}^3)$.

Let us mention some other results (of non-perturbative nature) allowing to show the essential self-adjointness for a larger class of potentials .

Theorem 4.16 (Essential self-adjointness for locally bounded potentials). *Let $\mathcal{H} = L^2(\mathbb{R}^d)$ and let $V \in L^\infty_{\mathrm{loc}}(\mathbb{R}^d)$ be real-valued such that $T = -\Delta + V$ with domain $D(T) = C_c^\infty(\mathbb{R}^d)$ is semibounded from below (remark that we require the semiboundness of T and not the semiboundedness of V). Then T is essentially self-adjoint on $C_c^\infty(\mathbb{R}^d)$.*

Proof. By adding a constant to the potential V one may assume that $T \ge 1$. In other words, using the integration by parts,

$$\int_{\mathbb{R}^d} \Big(\big|\nabla u(x)\big|^2 + V(x)\big|u(x)\big|^2\Big)\,\mathrm{d}x \ge \int_{\mathbb{R}^d} \big|u(x)\big|^2\,\mathrm{d}x \tag{4.5}$$

for all $u \in C_c^\infty(\mathbb{R}^d)$, and this extends by density at least to all $u \in H^1_{\mathrm{comp}}(\mathbb{R}^d)$, where H^1_{comp} stands for H^1 functions vanishing outside a compact set. By Theorem 4.11 it is sufficient to show that $\operatorname{ran} T$ is dense in $\mathcal{H}$.

Let $f \perp \operatorname{ran} T$, which means that $f \in L^2(\mathbb{R}^d)$ and $\big\langle f, (-\Delta+V)u\big\rangle = 0$ for all $u \in C_c^\infty(\mathbb{R}^d)$. Note that T preserves the real-valuedness, and we can suppose without loss of generality that f is real-valued (otherwise consider its real and imaginary parts separately). We have at least $(-\Delta+V)f = 0$ weakly, and then $\Delta f = Vf \in L^2_{\mathrm{loc}}(\mathbb{R}^d)$ (this uses the assumption $V \in L^\infty_{\mathrm{loc}}$). The elliptic regularity theorem (Theorem 1.54) shows that $f \in H^2_{\mathrm{loc}}(\mathbb{R}^d)$.

For any real-valued $\varphi, u \in C_c^\infty(\mathbb{R}^d)$ one has

$$\begin{aligned}\int_{\mathbb{R}^d} \nabla(\varphi f)\cdot\nabla(\varphi u)\,\mathrm{d}x &= \int_{\mathbb{R}^d} (f\nabla\varphi + \varphi\nabla f)\cdot(u\nabla\varphi + \varphi\nabla u)\,\mathrm{d}x\\ &= \int_{\mathbb{R}^d} \Big(|\nabla\varphi|^2 fu + \varphi\nabla\varphi\cdot(f\nabla u + u\nabla f) + \varphi^2\nabla f\cdot\nabla u\Big)\,\mathrm{d}x\\ &= \int_{\mathbb{R}^d} \Big(|\nabla\varphi|^2 fu + \varphi\nabla\varphi\cdot(f\nabla u - u\nabla f)\Big)\,\mathrm{d}x\\ &\quad + \int_{\mathbb{R}^d} \Big(2\varphi\nabla\varphi\, u + \varphi^2\nabla u)\cdot\nabla f\,\mathrm{d}x.\end{aligned}$$

Using the definition of weak derivatives one transforms the last summand:

$$\begin{aligned}\int_{\mathbb{R}^d} \Big(2\varphi\nabla\varphi\, u + \varphi^2\nabla u\Big)\cdot\nabla f\,\mathrm{d}x &= \int_{\mathbb{R}^d} \Big(\nabla(\varphi^2)\, u + \varphi^2\nabla u\Big)\cdot\nabla f\,\mathrm{d}x\\ &= \int_{\mathbb{R}^d} \nabla(\varphi^2 u)\cdot\nabla f\,\mathrm{d}x = \int_{\mathbb{R}^d} \varphi^2 u\,(-\Delta f)\,\mathrm{d}x,\end{aligned}$$

and then

$$\int_{\mathbb{R}^d} \nabla(\varphi f)\cdot\nabla(\varphi u)\,\mathrm{d}x = \int_{\mathbb{R}^d} \Big(|\nabla\varphi|^2 fu + \varphi\nabla\varphi\cdot(f\nabla u - u\nabla f)\Big)\,\mathrm{d}x + \int_{\mathbb{R}^d} \varphi^2 u\,(-\Delta f)\,\mathrm{d}x.$$

It follows that for any $u\in C_c^\infty(\mathbb{R}^d)$ one has

$$\begin{aligned}&\int_{\mathbb{R}^d} \nabla(\varphi f)\cdot\nabla(\varphi u)\,\mathrm{d}x + \int_{\mathbb{R}^d} V\,\varphi f\,\varphi u\,\mathrm{d}x\\ &\quad= \int_{\mathbb{R}^d} \Big(|\nabla\varphi|^2 fu + \varphi\nabla\varphi\cdot(f\nabla u - u\nabla f)\Big)\,\mathrm{d}x + \int_{\mathbb{R}^d} \varphi^2 u\,\underbrace{(-\Delta+V)f}_{=0}\,\mathrm{d}x\\ &\quad= \int_{\mathbb{R}^d} \Big(|\nabla\varphi|^2 fu + \varphi\nabla\varphi\cdot(f\nabla u - u\nabla f)\Big)\,\mathrm{d}x,\end{aligned}$$

which then extends by density at least to all $u\in H^1_{\mathrm{loc}}(\mathbb{R}^d)$. In particular, one can take $u:=f$, then $f\nabla u - u\nabla f\equiv 0$ and

$$\int_{\mathbb{R}^d} \Big(\big|\nabla(\varphi f)\big|^2 + V|\varphi f|^2\Big)\,\mathrm{d}x = \int_{\mathbb{R}^d} |\nabla\varphi|^2 f^2\,\mathrm{d}x.$$

Using (4.5) for $u:=\varphi f$ we arrive at

$$\int_{\mathbb{R}^d} |\nabla\varphi|^2 f^2\,\mathrm{d}x \ge \int_{\mathbb{R}^d} |\varphi f|^2\,\mathrm{d}x \tag{4.6}$$

for all real-valued $\varphi\in C_c^\infty(\mathbb{R}^d)$.

Let $\varphi\in C_c^\infty(\mathbb{R}^d)$ be real-valued such that $\varphi(x)=1$ for all $|x|\le 1$ and $\varphi(x)=0$ for all $|x|\ge 2$. Consider $\varphi_n: x\mapsto\varphi(x/n)$, then (4.6) gives

$$\int_{\mathbb{R}^d} \big|\nabla\varphi_n\big|^2 f^2\,\mathrm{d}x \ge \int_{\mathbb{R}^d} \varphi_n^2 f^2\,\mathrm{d}x,\quad n\in\mathbb{N}.$$

Recall that $f\in L^2(\mathbb{R}^d)$ and $\|\nabla\varphi_n\|_\infty = \frac{1}{n}\|\nabla\varphi\|_\infty$. Furthermore, if $N\ge\mathbb{N}$, then for all $n\ge N$ one has $\varphi_n=1$ on $B_N(0)$, and then

$$\begin{aligned}\int_{B_N(0)} f^2\,\mathrm{d}x \le \int_{\mathbb{R}^d} \varphi_n^2 f^2\,\mathrm{d}x &\le \int_{\mathbb{R}^d} \big|\nabla\varphi_n\big|^2 f^2\,\mathrm{d}x\\ &\le \|\nabla\varphi_n\|_\infty^2\|f\|_{L^2}^2 \le \frac{1}{n^2}\,\|\nabla\varphi\|_\infty^2\|f\|_{L^2}^2 \xrightarrow{n\to\infty} 0,\end{aligned}$$

which shows that $f=0$ in $B_N(0)$ a.e. As $N\in\mathbb{N}$ is arbitrary, one has $f=0$. □

Remark 4.17. Theorem 4.16 still holds if one replaces the assumption $V\in L^\infty_{\mathrm{loc}}(\mathbb{R}^d)$ by the weaker assumption $V\in L^2_{\mathrm{loc}}(\mathbb{R}^d)$, but then one needs more advanced machineries for the proof.

4.4 Stability of the essential spectrum

We have seen above that the spectrum of a self-adjoint operator can be characterized using Weyl sequences (Proposition 2.5). The following theorem gives a description of the essential spectrum in a similar spirit.

Theorem 4.18 (Weyl criterion for the essential spectrum). *For any self-adjoint operator T in a Hilbert space $\mathcal{H}$ the condition $\lambda \in \operatorname{spec}_{\mathrm{ess}} T$ is equivalent to the existence of a sequence $(u_n) \subset D(T)$ satisfying the following three properties:*

(a) $\|u_n\| \geq c$ *with some* $c > 0$*,*

(b) u_n *converge weakly to* 0*,*

(c) $(T - \lambda)u_n$ *converge to* 0 *in* $\mathcal{H}$*.*

Such a sequence will be called a singular Weyl sequence *for* λ*. Moreover, it will be shown in the proof that the conditions (a) and (b) above can be replaced by the single condition*

(AB) u_n *form an orthonormal sequence.*

Proof. Let $W(T)$ be the set of all λ for which one can find a singular Weyl sequence, then we need to show that $W(T) = \operatorname{spec}_{\mathrm{ess}} T$. Remark first that

$$W(T) \subset \operatorname{spec} T \subset \mathbb{R},$$

because any singular Weyl sequence is also a Weyl sequence.

(a) We first show the inclusion $W(T) \subset \operatorname{spec}_{\mathrm{ess}} T$. Let $\lambda \in W(T)$ and let (u_n) be an associated singular Weyl sequence. We know already that $\lambda \in \operatorname{spec} T$, so assume by contradiction that $\lambda \in \operatorname{spec}_{\mathrm{disc}} T$ and let $\Pi := P_T(\{\lambda\})$ be the orthogonal projector on $\ker(T - \lambda)$. Recall (Prop. 3.24) that one can find $c > 0$ such that

$$\|(T - \lambda)(I - \Pi)u\| \geq c\|(I - \Pi)u\| \text{ for all } u \in D(T). \tag{4.7}$$

The finite-rank operator Π is compact, hence, Πu_n converge to 0 in $\mathcal{H}$. Therefore, the vectors $w_n := (1 - \Pi)u_n$ satisfy $\|w_n\| \geq \frac{1}{2}$ for large n. On the other hand, the vectors $(T - \lambda)w_n = (1 - \Pi)(T - \lambda)u_n$ converge to 0 in $\mathcal{H}$, which contradicts (4.7).

(b) Now we pass to the proof of $\operatorname{spec}_{\mathrm{ess}} T \subset W(T)$. Let $\lambda \in \operatorname{spec}_{\mathrm{ess}} T$, then there are two possibilities.

(b: Possibility 1) Assume that λ is an isolated point of $\operatorname{spec} T$. Then it is an eigenvalue of infinite multiplicity, and any infinite orthonormal family in $\ker(T - \lambda)$ forms a singular Weyl sequence and $\lambda \in W(T)$.

(b: Possibility 2) Assume that λ is an accumulation point of $\operatorname{spec} T$. Consider the intervals $I_\varepsilon := (\lambda - \varepsilon, \lambda + \varepsilon)$, then for $0 < \varepsilon' < \varepsilon$ one has $I_{\varepsilon'} \subset I_\varepsilon$, which yields $\operatorname{ran} P_T(I_{\varepsilon'}) \subset \operatorname{ran} P_T(I_\varepsilon)$.

Now we prove the following <u>claim</u>: for any $\varepsilon > 0$ there exists $\varepsilon' \in (0, \varepsilon)$ such that $\operatorname{ran} P_T(I_{\varepsilon'}) \neq \operatorname{ran} P_T(I_\varepsilon)$. Assume the opposite, i.e. that for some $\varepsilon > 0$ one has $\operatorname{ran} P_T(I_{\varepsilon'}) = \operatorname{ran} P_T(I_\varepsilon)$ for all $\varepsilon' \in (0, \varepsilon)$. Then Prop. 3.21(g) gives

$$P_T(I_\varepsilon) = \lim_{\varepsilon' \to 0} P_T(I_{\varepsilon'}) = P_T(\{\lambda\}),$$

and due to $P_T(I_\varepsilon) = P_T(\{\lambda\}) + P_T(I_\varepsilon \setminus \{\lambda\})$ we have $P_T(I_\varepsilon \setminus \{\lambda\}) = 0$, and

$$\begin{aligned}\{0\} &= \operatorname{ran} P_T(I_\varepsilon \setminus \{\lambda\}) \equiv \operatorname{ran} P_T\big((\lambda-\varepsilon,\lambda) \cup (\lambda,\lambda+\varepsilon)\big)\\ &= \operatorname{ran} P_T\big((\lambda-\varepsilon,\lambda)\big) + \operatorname{ran} P_T\big((\lambda,\lambda+\varepsilon)\big),\end{aligned}$$

which shows $\operatorname{ran} P_T\big((\lambda-\varepsilon,\lambda)\big) = \operatorname{ran} P_T\big((\lambda,\lambda+\varepsilon)\big) = \{0\}$, and it follows that $\operatorname{spec} T \cap (\lambda-\varepsilon,\lambda) = \operatorname{spec} T \cap (\lambda,\lambda+\varepsilon) = \emptyset$, i.e. that λ is an isolated point of $\operatorname{spec} T$. This contradiction proves the above claim.

It follows from the claim that there is a strictly decreasing sequence $\varepsilon_n > 0$ with $\lim_{n\to\infty} \varepsilon_n = 0$ such that for the intervals $J_n := (\lambda-\varepsilon_n, \lambda+\varepsilon_n)$ one has $\operatorname{ran} P_T(J_{n+1}) \subset \operatorname{ran} P_T(J_n)$ and $\operatorname{ran} P_T(J_{n+1}) \neq \operatorname{ran} P_T(J_n)$ for any $n \in \mathbb{N}$. In view of $J_{n+1} \subset J_n$ and of the orthogonal decomposition (Prop. 3.21(b))

$$\operatorname{ran} P_T(J_n) = \operatorname{ran} P_T(J_{n+1}) + \operatorname{ran} P_T(J_n \setminus J_{n+1})$$

we have $P_T(J_n \setminus J_{n+1}) \neq 0$ for any n, and we can find $u_n \in \mathcal{H}$ with $\|u_n\| = 1$ and $P_T(J_n \setminus J_{n+1})u_n = u_n$. As the sets $J_n \setminus J_{n+1}$ are mutually disjoint, the vectors u_n form an orthonormal sequence (Prop. 3.21(b)) and, in particular, converge weakly to 0. Due to $u_n \in \operatorname{ran} P_T(J_n \setminus J_{n+1}) \subset \operatorname{ran} P_T(J_n)$ we have by Prop. 3.21(f)

$$u_n \in D(T), \qquad \|(T-\lambda)u_n\| \le \varepsilon_n \|u_n\| = \varepsilon_n \overset{n\to\infty}{\longrightarrow} 0,$$

This shows that (u_n) is a singular Weyl sequence for λ and $\lambda \in W(T)$. □

The following theorem provides a starting point to the study of perturbations of self-adjoint operators.

Theorem 4.19 (Stability of the essential spectrum). *Let A and B be self-adjoint operators such that $K(z) := (A-z)^{-1} - (B-z)^{-1}$ is a compact operator for some $z \in \operatorname{res} A \cap \operatorname{res} B$, then $\operatorname{spec}_{\mathrm{ess}} A = \operatorname{spec}_{\mathrm{ess}} B$.*

Proof. Let $\lambda \in \operatorname{spec}_{\mathrm{ess}} A$ and let (u_n) be an associated singular Weyl sequence. Without loss of generality assume that $\|u_n\| = 1$ for all n. We have

$$\lim \Big((A-z)^{-1} - \frac{1}{\lambda - z}\Big)u_n = \lim \frac{1}{z-\lambda}(A-z)^{-1}(A-\lambda)u_n = 0. \tag{4.8}$$

As $K(z)$ is compact, the sequence $K(z)u_n$ converges to 0 in $\mathcal{H}$, and

$$\begin{aligned}\lim \frac{1}{z-\lambda}(B-\lambda)(B-z)^{-1}u_n &= \lim \Big((B-z)^{-1} - \frac{1}{\lambda-z}\Big)u_n\\ &= \lim \Big((A-z)^{-1} - \frac{1}{\lambda-z}\Big)u_n - \lim K(z)u_n = 0.\end{aligned}$$

Now denote $v_n := (B-z)^{-1}u_n$, then the preceding computations already shows that $(B-\lambda)v_n$ converge to 0 in $\mathcal{H}$. As u_n weakly converge to 0, for any $h \in \mathcal{H}$ one has

$$\langle h, v_n\rangle = \langle h, (B-z)^{-1}u_n\rangle = \big\langle \big((B-z)^{-1}\big)^* h, u_n\big\rangle \overset{n\to\infty}{\longrightarrow} 0,$$

showing that v_n converge weakly to 0. In addition we have

$$v_n - \frac{1}{\lambda - z}\, u_n = (B-z)^{-1}u_n - \frac{1}{\lambda - z}\, u_n$$
$$= \underbrace{(A-z)^{-1}u_n - \frac{1}{\lambda - z}\, u_n}_{\to 0 \text{ by } (4.8)} - \underbrace{K(z)u_n}_{\to 0} \to 0,$$

which gives

$$\lim \|v_n\| = \lim \frac{1}{|\lambda - z|} \underbrace{\|u_n\|}_{=1} = \frac{1}{|\lambda - z|} > 0.$$

Therefore, the vectors v_n (with sufficiently large n) form a singular Weyl sequence for B and λ, and $\lambda \in \operatorname{spec}_{\text{ess}} B$ due to the Weyl criterion (Theorem 4.18).

The above argument shows $\operatorname{spec}_{\text{ess}} A \subset \operatorname{spec}_{\text{ess}} B$, and by interchanging the roles of A and B one obtains $\operatorname{spec}_{\text{ess}} A \supset \operatorname{spec}_{\text{ess}} B$. $\square$

Remark 4.20. (a) a simple algebra based on the resolvent identities shows that if $K(z)$ in Theorem 4.19 is compact for <u>some</u> $z \in \operatorname{res} A \cap \operatorname{res} B$, then it is compact for <u>all</u> $z \in \operatorname{res} A \cap \operatorname{res} B$.

(b) The simplest situation in which the assumption of Theorem 4.19 is satisfied is $B = A + K$ with a compact self-adjoint operator K: in this case one has that

$$(A+i)^{-1} - (B+i)^{-1} = (B+i)^{-1}K(A+i)^{-1} = \text{bounded}\cdot\text{compact}\cdot\text{bounded} = \text{compact}.$$

But this does not exhaust all possibilities: the assumption can be satisfied even for $B = A + K$ with some unbounded K: this is what we are going to discuss.

Definition 4.21 (Relatively compact operators). Let A be a self-adjoint operator in a Hilbert space $\mathcal{H}$, and let B a <u>closable</u> linear operator in $\mathcal{H}$ with $D(A) \subset D(B)$. We say that B is *relatively compact with respect to* A (or simply A-*compact*) if $B(A-z)^{-1}$ is compact for some $z \in \operatorname{res} A$.

Remark 4.22. It follows from the resolvent identitites that if $B(A-z)^{-1}$ is compact for <u>some</u> $z \in \operatorname{res} A$, then it is compact for <u>all</u> $z \in \operatorname{res} A$. If B is compact, then it is relatively compact with respect to any A.

We first show that relatively compact perturbations are covered by the Kato-Rellich theorem:

Lemma 4.23. *If B is relatively compact with respect to A, then it is also infinitesimally small with respect to A.*

Proof. We show first that

$$\lim_{\lambda \to +\infty} \left\| B(A - i\lambda)^{-1} \right\| = 0 \tag{4.9}$$

Assume that (4.9) is false, then one can find a constant $\alpha > 0$, $0 \neq u_n \in \mathcal{H}$ and $\lambda_n > 0$ with $\lim \lambda_n = +\infty$ such that $\left\| B(A - i\lambda)^{-1}u_n \right\| > \alpha \|u_n\|$ for all n. Consider

$v_n := (A - i\lambda_n)^{-1}u_n \in D(A)$, then $\|Bv_n\| > \alpha\|(A - i\lambda)v_n\|$. Taking the square on both sides and using $\big\|(A - i\lambda_n)v_n\big\|^2 = \|Av_n\|^2 + \lambda_n^2\|v_n\|^2$ we arrive at

$$\|Bv_n\|^2 > \alpha^2\|Av_n\|^2 + \alpha^2\lambda_n^2\|v_n\|^2.$$

Without loss of generality assume the normalization $\|Bv_n\| = 1$, then:

(i) the sequence Av_n is bounded and (ii) v_n converge to 0.

Let $z \in \operatorname{res} A$, then $(A - z)v_n$ is also bounded and, therefore, it contains a weakly convergent subsequence $(A - z)v_{n_k}$. As $B(A - z)^{-1}$ is a compact operator, the vectors $B(A - z)^{-1}(A - z)v_{n_k} \equiv Bv_{n_k}$ converge in $\mathcal{H}$ to some w with $\|w\| = 1$. As v_{n_k} converge to 0, the closability of B implies $w = 0$. This contradiction shows that (4.9) is true.

Let $a > 0$, then one can find $\lambda > 0$ such that $\|B(A - i\lambda)^{-1}u\| \leq a\|u\|$ for all $u \in \mathcal{H}$. Taking $u := (A - i\lambda)v$ with an arbitrary $v \in D(A)$ we see that

$$\|Bv\| \leq a\|(A - i\lambda)v\| \leq a\|Av\| + a\lambda\|v\| \text{ for all } v \in D(A).$$

As $a > 0$ is arbitrary, we get the result. □

Now we arrive at the main result of this section:

Theorem 4.24 (Essential spectrum under relatively compact perturbations). *Let A be a self-adjoint operator in a Hilbert space $\mathcal{H}$ and let B be symmetric and A-compact, then:*

- *the operator $A + B$ with $D(A + B) = D(A)$ is self-adjoint,*
- *if A is semibounded, so $A + B$ is also semibounded,*
- $\operatorname{spec}_{\text{ess}}(A + B) = \operatorname{spec}_{\text{ess}} A$.

Proof. The self-adjointness/semiboundedness of $A + B$ follow by the Kato-Rellich theorem 4.13 (which is applicable due to Lemma 4.23). One has

$$(A + i)^{-1} - (A + B + i)^{-1} = \underbrace{(A + B + i)^{-1}}_{\text{bounded}}\underbrace{B(A + i)^{-1}}_{\text{compact}} = \text{compact},$$

and $\operatorname{spec}_{\text{ess}}(A + B) = \operatorname{spec}_{\text{ess}} A$ by Theorem 4.19. □

4.5 Essential spectra for Schrödinger operators

Definition 4.25 (Potentials of Kato class). Let $d \in \mathbb{N}$ and

$$p = 2 \text{ for } d \le 3, \qquad p > \frac{d}{2} \text{ for } d \ge 4$$

(i.e. p are the same as in the Kato-Rellich theorem for Schrödinger operators, Theorem 4.14). A measurable function $V : \mathbb{R}^d \to \mathbb{R}$ belongs to the **Kato class** on $\mathbb{R}^d$ if for any $\varepsilon > 0$ one can find real-valued $V_p \in L^p(\mathbb{R}^d)$ and $V_\infty \in L^\infty(\mathbb{R}^d)$ with $V = V_p + V_\infty$ and $\|V_\infty\|_\infty < \varepsilon$.

Theorem 4.26 (Essential spectrum for Kato class potentials). *If V is a Kato class potential in $\mathbb{R}^d$, then V is compact with respect to the free Laplacian in $L^2(\mathbb{R}^d)$, and the essential spectrum of $T = -\Delta + V$ is equal to $[0, \infty)$.*

Proof. First remark that V is covered by Theorem 4.14, so T is uniquely defined. We give the proof for $d \le 3$ only.

Let $\mathcal{F}$ and T_0 be the Fourier transform and the free Laplacian in $L^2(\mathbb{R}^d)$. Then $T_0 = \mathcal{F}^{-1} M_{|\xi|^2} \mathcal{F}$ and for any $z \notin \operatorname{spec} T_0$ one has

$$(T_0 - z)^{-1} = \mathcal{F}^{-1} (M_{|\xi|^2} - z)^{-1} \mathcal{F} \equiv \mathcal{F}^{-1} M_{\frac{1}{|\xi|^2 - z}} \mathcal{F},$$

i.e. for any $f \in L^2(\mathbb{R}^d)$ one has

$$\mathcal{F}(T_0 - z)^{-1} f(\xi) = \frac{1}{|\xi^2| - z}\, \widehat{f}(\xi). \tag{4.10}$$

The function $\xi \mapsto (|\xi^2| - z)^{-1}$ is in $L^2(\mathbb{R}^d)$ and can be written as $(2\pi)^{\frac{d}{2}} \widehat{h}_z$ for some function $h_z \in L^2(\mathbb{R}^d)$. Then (4.10) takes the form

$$\mathcal{F}(T_0 - z)^{-1} f = (2\pi)^{\frac{d}{2}} \widehat{h}_z \widehat{f}(\xi),$$

which means $(T_0 - z)^{-1} f = h_z * f$, where $*$ is the convolution product,[6] and

$$(T_0 - z)^{-1} f = \int_{\mathbb{R}^d} h_z(x - y) f(y)\, \mathrm{d}y.$$

Let $\varepsilon > 0$ and let $V = V_2 + V_\infty$ with $V_2 \in L^2(\mathbb{R}^d)$ and $\|V_\infty\|_\infty < \varepsilon$. Then $V_2(T - z)^{-1}$ is an integral operator,

$$V_2(T_0 - z)^{-1} f(x) = \int_{\mathbb{R}^d} K(x, y) f(y)\, \mathrm{d}y, \quad K(x, y) = V_2(x) h_z(x - y).$$

One has

$$\begin{aligned}\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \big|K(x, y)\big|^2 \mathrm{d}x\, \mathrm{d}y &= \int_{\mathbb{R}^d} \big|V_2(x)\big|^2 \int_{\mathbb{R}^d} \big|h_z(x - y)\big|^2 \mathrm{d}y\, \mathrm{d}x \\ \text{(substitution } s := x - y) &= \int_{\mathbb{R}^d} \big|V_2(x)\big|^2 \int_{\mathbb{R}^d} \big|h_z(s)\big|^2 \mathrm{d}s\, \mathrm{d}x = \|V_2\|_{L^2}^2 \|h_z\|_{L^2}^2 < \infty,\end{aligned}$$

[6] One easily checks the identity $\widehat{f * g} = (2\pi)^{d/2} \widehat{f}\, \widehat{g}$ for "good" functions f and g, which is then suitably extended using appropriate limit passages.

which means that $V_\varepsilon(T_0 - z)^{-1}$ is a Hilbert-Schmidt operator and, therefore, is compact, see Subsection 2.3. At the same time we have the estimate

$$\|V_\infty(T_0 - z)^{-1}\| \leq \|V_\infty\|_\infty \|(T_0 - z)^{-1}\| \leq \varepsilon \|(T - z)^{-1}\|.$$

Therefore, $V(T_0 - z)^{-1} = V_2(T_0 - z)^{-1} + V_\infty(T_0 - z)^{-1}$, the first summand is a compact operator and the second summand is a bounded operator whose norm can be made arbitrarily small, which shows that $V(T_0 - z)^{-1}$ is also a compact operator. It follows by Theorem 4.24 that $\operatorname{spec}_{\mathrm{ess}} T = \operatorname{spec}_{\mathrm{ess}} T_0 = [0, \infty)$. $\square$

Example 4.27 (Coulomb potential). The previous theorem easily applies e.g. to the operators $-\Delta + \frac{q}{|x|}$. For any $R > 0$ one has

$$\frac{1}{|x|} = \frac{1_{|x|<R}}{|x|} + \frac{1_{x \geq |R|}}{|x|}.$$

the first summand is in $L^2(\mathbb{R}^3)$ and the second summand is bounded and its sup-norm can be made arbitrarily small if one takes R sufficiantly large. So the essential spectrum of $-\Delta + q/|x|$ is always the same as for the free Laplacian, i.e. coincides with $[0, +\infty)$.

5 Min-max principle and applications

5.1 Min-max principle

The min-max principle is one of the main tools in the spectral analysis of semibounded self-adjoint operators. It has numerours applications, in particular, it allows one to obtain various inequalitites for the eigenvalues and to compare the spectra of operators acting in different Hilbert spaces: we will see some applications in the subsequent chapters.

Throughout this subsection let T be a lower semibounded self-adjoint operator in an *infinite-dimensional* Hilbert space $\mathcal{H}$. We denote

$$\Sigma \equiv \Sigma(T) := \begin{cases} \inf \operatorname{spec}_{\text{ess}} T, & \text{if } \operatorname{spec}_{\text{ess}} T \neq \emptyset, \\ +\infty, & \text{otherwise.} \end{cases}$$

Theorem 5.1 (Min-max principle: operator version). *Introduce the min-max numbers of T as follows:*

$$\Lambda_n = \Lambda_n(T) = \inf_{\substack{V \subset D(T) \\ \dim V = n}} \sup_{\varphi \in V,\, \varphi \neq 0} \frac{\langle \varphi, T\varphi \rangle}{\langle \varphi, \varphi \rangle}, \quad n \in \mathbb{N},$$

then the sequence (Λ_n) is non-decreasing, and we are in one and only one of the following situations:

(a) For any $n \in \mathbb{N}$ there holds $\Lambda_n < \Sigma$.

Then T has infinitely many discrete eigenvalues in $(-\infty, \Sigma)$, and $\Lambda_n(T)$ is exactly the n-th eigenvalue of T, if one counts them in the non-decreasing order and takes the multiplicities into account.

(b) There exists $N \in \mathbb{N}_0$ such that $\Lambda_n < \Sigma$ for all $n \leq N$ with and $\Lambda_{N+1} \geq \Sigma$.

Then T has exactly N discrete eigenvalues in $(-\infty, \Sigma)$, and the number Λ_n is exactly the nth eigenvalue of T for each $n \in \{1, \dots, N\}$, while $\Lambda_n = \Sigma$ for all $n \geq N + 1$.

Remark that the case $N = 0$ is possible: this means than $\Lambda_n \geq \Sigma$ for all $n \in \mathbb{N}$, and T has no discrete eigenvalues in $(-\infty, \Sigma)$.

Proof. For any $W \subset D(T)$ with $\dim W = n + 1$ one can find $V \subset D(T)$ with $\dim V = n$ and $V \subset W$, and then

$$\sup_{\varphi \in W,\, \varphi \neq 0} \frac{\langle \varphi, T\varphi \rangle}{\langle \varphi, \varphi \rangle} \geq \sup_{\varphi \in V,\, \varphi \neq 0} \frac{\langle \varphi, T\varphi \rangle}{\langle \varphi, \varphi \rangle} \geq \inf_{\substack{V \subset W \\ \dim V = n}} \sup_{\varphi \in V,\, \varphi \neq 0} \frac{\langle \varphi, T\varphi \rangle}{\langle \varphi, \varphi \rangle}.$$

It follows that

$$\Lambda_{n+1} = \inf_{\substack{W \subset D(T) \\ \dim W = n+1}} \sup_{\varphi \in W,\, \varphi \neq 0} \frac{\langle \varphi, T\varphi \rangle}{\langle \varphi, \varphi \rangle} \geq \inf_{\substack{W \subset D(T) \\ \dim W = n+1}} \inf_{\substack{V \subset W \\ \dim V = n}} \sup_{\varphi \in V,\, \varphi \neq 0} \frac{\langle \varphi, T\varphi \rangle}{\langle \varphi, \varphi \rangle}$$

$$\geq \inf_{\substack{V \subset D(T) \\ \dim V = n}} \sup_{\varphi \in V,\, \varphi \neq 0} \frac{\langle \varphi, T\varphi \rangle}{\langle \varphi, \varphi \rangle} = \Lambda_n.$$

By assumption, the spectrum of T in $(-\infty,\Sigma)$ is purely discrete, and the discrete eigenvalues in $(-\infty,\Sigma)$ may only accumulate to Σ (as any accumulation point of the spectrum belongs to the essential spectrum). Hence, all these eigenvalues can be enumerated in the non-decreasing order (one counts according to the mutliplicitieis): we denote them by E_k and denote by v_k associated eigenvectors, $k \in \{1,\dots,N\}$ and $N \in \mathbb{N}\cup\{\infty\}$. We may assume without loss of generality that v_k form an orthonormal family, i.e. $\langle v_j, v_k\rangle = \delta_{j,k}$.

Let $n \in \{1,\dots,N\}$ and $V_n = \operatorname{span}(v_1,\dots,v_n)$, then $V_n \subset D(T)$ with $\dim V_n = n$. For any $\varphi \in V_n$ we have

$$\begin{aligned}\langle\varphi, T\varphi\rangle &= \Big\langle \sum_{j=1}^n \langle v_j,\varphi\rangle v_j, T\sum_{j=1}^n \langle v_j,\varphi\rangle v_j\Big\rangle = \Big\langle \sum_{j=1}^n \langle v_j,\varphi\rangle v_j, \sum_{j=1}^n \langle v_j,\varphi\rangle E_j v_j\Big\rangle\\ &= \sum_{j=1}^n E_j |\langle\varphi, v_j\rangle|^2 \le E_n \sum_{j=1}^n |\langle\varphi, v_j\rangle|^2 = E_n\|\varphi\|^2,\end{aligned}$$

and it follows that $\Lambda_n \le E_n$.

Now let $n \in \{1,\dots,N\}$, V be any n-dimensional subspace of $D(T)$, P be the orthogonal projector on V_{n-1}. If one has the strict inequality $E_{n-1} < E_n$, then $E_1,\dots,E_{n-1}$ exhausts the whole spectrum of T in $(-\infty, E_n)$ and one has the equality $P = P_T\big((-\infty,E_n)\big)$, see Remark 3.22. If $E_{n-1} = E_n$, then choose the largest $k \le n-1$ with $E_k < E_n$, then $E_1,\dots,E_k$ exhausts the whole spectrum of T in $(-\infty,E_n)$, and $\operatorname{ran} P_T\big((-\infty,E_n)\big) = V_k \subset V_{n-1} = \operatorname{ran} P$. Therefore, in all cases one has the inclusion $\operatorname{ran} P_T\big((-\infty,E_n)\big) \subset \operatorname{ran} P$.

Remark that due to $\dim V = n > n-1 = \dim\operatorname{ran} P$ one can find a non-zero $\psi \in V$ with $P\psi = 0$. The condition $P\psi = 0$ implies $\psi \perp \operatorname{ran} P_T\big((-\infty,E_n)\big)$, which is equivalent to $\psi \in \operatorname{ran} P_T\big([E_n,+\infty)\big)$, see Prop. 3.21(b.3). Then by Prop. 3.21(g) we arrive at $\langle\psi, T\psi\rangle \ge E_n\|\psi\|^2$. Therefore,

$$\sup_{\varphi\in V,\,\varphi\neq 0} \frac{\langle\varphi, T\varphi\rangle}{\langle\varphi,\varphi\rangle} \ge \frac{\langle\psi, T\psi\rangle}{\langle\psi,\psi\rangle} \ge E_n.$$

As V was an arbitrary n-dimensional subspace of $D(T)$, this proves $\Lambda_n \ge E_n$.

The above discussion shows that $\Lambda_n = E_n$ for any $n \in \{1,\dots,N\}$. Now consider two cases:

(Case 1) $N = \infty$, then we have already $\Lambda_n = E_n$ for all $n \in \mathbb{N}$.

(Case 2) $N < \infty$. We have already $\Lambda_n = E_n$ for any $n \in \{1,\dots,N\}$. Let us show that $\Lambda_n = \Sigma$ for all $n \ge N+1$.

First remark that $\dim\operatorname{ran} P_T\big((-\infty,\Sigma)\big) = N$ (see Remark 3.22). Let $n \ge N+1$ and V be any n-dimensional subspace of $D(T)$, then there is a non-zero $\psi \in V$ with $\psi \perp \operatorname{ran} P_T\big((-\infty,\Sigma)\big)$, which means that $\psi \in \operatorname{ran} P_T\big([\Sigma,\infty)\big)$ and implies $\langle\psi, T\psi\rangle \ge \Sigma\|\psi\|^2$. Then

$$\sup_{\varphi\in V,\,\varphi\neq 0} \frac{\langle\varphi, T\varphi\rangle}{\langle\varphi,\varphi\rangle} \ge \frac{\langle\psi, T\psi\rangle}{\langle\psi,\psi\rangle} \ge \Sigma,$$

and $\Lambda_n \ge \Sigma$ due to the arbitrary choice of V.

On the other hand, let $\varepsilon > 0$, then $W = \operatorname{ran} P_T\big((\Sigma-\varepsilon, \Sigma+\varepsilon)\big)$ if an infinite-diemsnional subspace of $D(T)$ (as $\Sigma \in \operatorname{spec}_{\mathrm{ess}} T$) and contains an infinite orthonormal family $(u_j)_{j\in\mathbb{N}}$. Then

$$W_n := \operatorname{span}(u_1, \dots, u_n) \subset \operatorname{ran} P_T\big((\Sigma-\varepsilon, \Sigma+\varepsilon)\big) \subset D(T)$$

is an n-dimensional subspace. For any $u \in W_n$ we have $\|(T-\Sigma)u\| \le \varepsilon\|u\|$ and

$$\big|\langle u, Tu\rangle - \Sigma\langle u, u\rangle\big| = \big|\langle u, (T-\Sigma)u\rangle\big| \le \|u\|\,\big\|(T-\Sigma)u\big\| \le \varepsilon\|u\|^2,$$

in particular, $\langle u, Tu\rangle \le (\Sigma+\varepsilon)\|u\|^2$. Therefore,

$$\Lambda_n = \inf_{\substack{V\subset D(T)\\ \dim V=n}} \sup_{\substack{u\in V\\ u\neq 0}} \frac{\langle u, Tu\rangle}{\langle u, u\rangle} \le \sup_{\substack{u\in W_n\\ u\neq 0}} \frac{\langle u, Tu\rangle}{\langle u, u\rangle} \le \Sigma+\varepsilon.$$

As $\varepsilon > 0$ was arbitrary, one has $\Lambda_n \le \Sigma$ for any $n \in \mathbb{N}$. Together with the above estimates one obtains $\Lambda_n = \Sigma$ for all $n \ge N+1$.

We now see that the case 1 corresponds to the situation (a) of the claim, while the case 2 corresponds to the situation (b) of the claim, and this covers all possible situations. $\square$

For some operators the domain is not given explicitly (or has a very complcated description), and would be preferrable to work with the sesquilinear form instead of the operator. Hence, the following reformulation is useful:

Theorem 5.2 (Min-max principle: form version). *Let t be the closed sesquilinear form of T and $\mathcal{D} \subset D(t)$ be a dense subspace with respect to $\langle\cdot,\cdot\rangle_t$. Then for any $n \in \mathbb{N}$ one has*

$$\Lambda_n(T) = \inf_{\substack{V\subset\mathcal{D}\\ \dim V=n}} \sup_{\varphi\in V,\,\varphi\neq 0} \frac{t(\varphi,\varphi)}{\langle\varphi,\varphi\rangle}. \tag{5.1}$$

Proof. Denote by $\mu_n(\mathcal{D})$ the quantity on the right-hand side of (5.1), then the standard density argument shows that $\mu_n(\mathcal{D}) = \mu_n\big(D(t)\big)$. We know by Theorem 1.61 that $D(T)$ is dense in $D(t)$, therefore, $\mu_n\big(D(T)\big) = \mu_n\big(D(t)\big)$. For $\varphi \in D(T)$ we have $t(\varphi,\varphi) = \langle\varphi, T\varphi\rangle$, which shows that $\mu_n\big(D(T)\big) = \Lambda_n(T)$. $\square$

It is clearly of interest to have more candidates for $\mathcal{D}$ in Theorem 5.2 (suitable $\mathcal{D}$ are often referred to as *test subspaces*). The following assertion can be useful:

Proposition 5.3. *If T is essentially self-adjoint on some subspace $\mathcal{D} \subset D(T)$, then $\mathcal{D}$ is dense in $D(t)$ and can be used as a test subspace in* (5.1).

Proof. Without loss of generality assume $T \ge 1$, then $\langle u, v\rangle_t = t(u,v) = \langle u, Tv\rangle$ for all $u, v \in D(T)$. Let $u \in D(T)$, then there exist $u_n \in \mathcal{D}$ with $u_n \to u$ and $Tu_n \to Tu$ in $\mathcal{H}$, and

$$\|u-u_n\|_t^2 = \big\langle u-u_n, T(u-u_n)\big\rangle \le \|u-u_n\|\,\|Tu-Tu_n\| \xrightarrow{n\to\infty} 0.$$

This shows that any $u \in D(T)$ can be approximated by vectors from $\mathcal{D}$ in $D(t)$. As $D(T)$ is dense in $D(t)$, the claim follows. $\square$

For what follows it will be convenient to introduce the following notation: if T has N eigenvalues in $(-\infty, \Sigma)$, then for $n \in \{1, \ldots, N\}$ one denotes

$E_n(T) :=$ the n-th eigenvalue of T (if enumerated in the non-decreasing order with multiplicities taken into account).

The following assertions are obvious cosequences of Theorem 5.1:

Corollary 5.4 (Min-max and existence of eigenvalues). *The following relations hold:*

(a) $\lim_{n\to\infty} \Lambda_n(T) = \Sigma(T)$,

(b) for each $N \in \mathbb{N}$ *the following two assertions are equivalent:*

- T *has at least* N *eigenvalues in* $\big(-\infty, \Sigma(T)\big)$,
- $\Lambda_N(T) < \Sigma(T)$.

If one of these conditions is satisfied, then for any $n \in \{1, \ldots, N\}$ *one has* $E_n(T) = \Lambda_n(T)$.

5.2 Comparison of operators

It is important that with the help of the min-max principle it is sometimes possible to compare the spectral properties of two operators even if they acs in different Hilbert spaces. The assertions of this section are simple corollaries of the min-max principle, but they will play a central role in what follows.

Corollary 5.5 (Min-max for perturbations). *Let T be a lower semibounded self-adjoint operators in $\mathcal{H}$ and $A = A^* \in \mathcal{B}(\mathcal{H})$, then*

$$\big|\Lambda_n(T+A) - \Lambda_n(T)\big| \le \|A\| \text{ for any } n \in \mathbb{N}.$$

Proof. For any $u \in D(T)$ with $u \neq 0$ one has $\big\langle u, (T+A)u\big\rangle = \langle u, Tu\rangle + \langle u, Au\rangle$ and $\big|\langle u, Au\rangle\big| \le \|A\|\,\|u\|^2$, therefore,

$$\frac{\langle u, (T+A)u\rangle}{\|u\|^2} \le \frac{\langle u, Tu\rangle}{\|u\|^2} + \|A\|.$$

Therefore, for any subspace $F \subset D(T)$ one has

$$\sup_{\varphi\in F,\,\varphi\neq 0} \frac{\langle \varphi, (T+A)\varphi\rangle}{\langle \varphi, \varphi\rangle} \le \sup_{\varphi\in F,\,\varphi\neq 0} \frac{\langle \varphi, T\varphi\rangle}{\langle \varphi, \varphi\rangle} + \|A\|,$$

and taking the infimum over all subspaces F with $\dim F = n$ first on the lef-hand side and then on the right-hand side one shows $\Lambda_n(T+A) \le \Lambda_n(T) + \|A\|$. Writing $T = (T+A) - A$ one obtains $\Lambda_n(T) \le \Lambda_n(T+A) + \|A\|$. □

Corollary 5.6. *If T is a lower semibounded self-adjoint operator with compact resolvent and $A = A^* \in \mathcal{B}(\mathcal{H})$, then*

$$\big|E_n(T+A) - E_n(T)\big| \le \|A\| \text{ for any } n \in \mathbb{N}.$$

Proof. From the equality

$$(T+A+i)^{-1} = \underbrace{(T+i)^{-1}}_{\text{compact}} - \underbrace{(T+A+i)^{-1}A}_{\text{bounded}} \cdot \underbrace{(T+i)^{-1}}_{\text{compact}}$$

it follows that $T+A$ has compact resolvent, and then $E_n = \Lambda_n$ for any n and for both T and $T+A$, and the inequality follows from Corollary 5.5. □

Corollary 5.7 (Min-max spectral estimates). *Let T and $\widetilde{T}$ be lower semibounded seml-adjoint operators in Hilbert spaces $\mathcal{H}$ and $\widetilde{\mathcal{H}}$ such that*

$$\Lambda_n(T) \le \Lambda_n(\widetilde{T}) \text{ for all } n \in \mathbb{N}, \tag{5.2}$$

then the following assertions hold true:

(a) $\Sigma(T) \le \Sigma(\widetilde{T})$,

(b) if for some $\lambda \in \mathbb{R}$ the following two assumptions are satisfied:

- *the spectra of both T and $\widetilde{T}$ in $(-\infty,\lambda)$ are purely discrete,*
- *$\widetilde{T}$ has N eigenvalues (counting the multiplicity) in $(-\infty,\lambda)$,*

then:

- *T has at least N eigenvalues in $(-\infty,\lambda)$,*
- *for any $n\in\{1,\dots,N\}$ one has $E_n(T)\le E_n(\widetilde{T})$.*

(c) if for some $\lambda\in\mathbb{R}$ the spectrum of T in $(-\infty,\lambda)$ is purely discrete and consists of N eigenvalues (counting with multiplicities), then:

- *the spectrum of $\widetilde{T}$ in $(-\infty,\lambda)$ is also purely discrete,*
- *$\widetilde{T}$ has at most N eigenvalues (counting the multiplicity) in $(-\infty,\lambda)$.*

Proof. The claim (a) follows from Corollary 5.4(a). For (b) we remark that under the assumptions made one has $\lambda<\min\big\{\Sigma(T),\Sigma(\widetilde{T})\big\}$ and $\Lambda_N(\widetilde{T})<\lambda$, and it follows (Corollary 5.4(b)) that $E_n(\widetilde{T})=\Lambda_n(\widetilde{T})$ for any $n\le N$. Due to (5.2) we have $\Lambda_N(T)\le\Lambda_N(\widetilde{T})<a<\Sigma(T)$, which means that T has at least N eigenvalues in $\big(-\infty,\Sigma(T)\big)$, and then $E_n(T)=\Lambda_n(T)$ for any $n\le N$. Finally, for any $n\le N$ we have $E_n(T)=\Lambda_n(T)\le\Lambda_n(\widetilde{T})=E_n(\widetilde{T})$.

(c) By assumption we have $\lambda\le\Sigma(T)\equiv\lim_{n\to\infty}\Lambda_n(T)$. Due to (5.2) we also have $\lambda\le\lim_{n\to\infty}\Lambda_n(\widetilde{T})$, which shows that the spectrum of $\widetilde{T}$ in $(-\infty,\lambda)$ is purely discrete. We are now in the situation of (b), and the number of eigenvalues of $\widetilde{T}$ in $(-\infty,\lambda)$ cannot exceed N. □

Let let us discuss situations in which the main assumption (5.2) of Corollary 5.7 is satisfied. One of the most frequently used constructions is as follows:

Definition 5.8 (Comparison of operators via an identification map). Let T and $\widetilde{T}$ be lower semibounded self-adjoint operators in Hilbert spaces $\mathcal{H}$ and $\widetilde{\mathcal{H}}$, generated by closed sesqulinear forms t and $\widetilde{t}$. Let $\widetilde{\mathcal{D}}\subset D(\widetilde{t}\,)$ be a dense subspace (with respect to $\|\cdot\|_{\widetilde{t}}$) and assume that there exists a linear map (identification map) $J:\widetilde{\mathcal{D}}\to D(t)$ such that for any $\varphi\in\widetilde{D}$ one has:

$$\|J\varphi\|_{\mathcal{H}}=\|\varphi\|_{\widetilde{\mathcal{H}}},\qquad t(J\varphi,J\varphi)\le\widetilde{t}(\varphi,\varphi),$$

then we write

$$T\le\widetilde{T}\text{ via }J.$$

In the subsequent chapter we will see several types of very specific identification maps, but their main application is as follows:

Corollary 5.9 (Min-max inequality via an identification operator). *Let T and $\widetilde{T}$ be lower semibounded self-adjoint operators in Hilbert spaces $\mathcal{H}$ and $\widetilde{\mathcal{H}}$ such that $T\le\widetilde{T}$ via some identification map J. Then $\Lambda_n(T)\le\Lambda_n(\widetilde{T})$ for all $n\in\mathbb{N}$ (in particular, all assertions of Corollary 5.7 hold true).*

Proof. Let t and $\widetilde{t}$ be the sesquilinear forms for t and $\widetilde{t}$ Let $\widetilde{\mathcal{D}} \subset D(\widetilde{t}\,)$ be a dense subsset and $J : \widetilde{\mathcal{D}} \to D(t)$ be an identification map. Then J is injective, and $\dim J(V) = \dim V$ for any subspace $V \subset \widetilde{\mathcal{D}}$. Then one has

$$\begin{aligned}
\Lambda_n(T) &= \inf_{\substack{V\subset D(t)\\ \dim V=n}} \sup_{\substack{\varphi\in V\\ \varphi\neq 0}} \frac{t(\varphi,\varphi)}{\langle\varphi,\varphi\rangle_{\mathcal{H}}} \le \inf_{\substack{V=J(\widetilde{V})\\ \widetilde{V}\subset\widetilde{D}\\ \dim \widetilde{V}=n}} \sup_{\substack{\varphi\in V\\ \varphi\neq 0}} \frac{t(\varphi,\varphi)}{\langle\varphi,\varphi\rangle_{\mathcal{H}}}\\
&= \inf_{\substack{\widetilde{V}\subset\widetilde{D}\\ \dim \widetilde{V}=n}} \sup_{\substack{\phi\in\widetilde{V}\\ \phi\neq 0}} \frac{t(J\phi,J\phi)}{\langle J\phi,J\phi\rangle_{\mathcal{H}}} \le \inf_{\substack{\widetilde{V}\subset\widetilde{D}\\ \dim \widetilde{V}=n}} \sup_{\substack{\phi\in\widetilde{V}\\ \phi\neq 0}} \frac{\widetilde{t}(\phi,\phi)}{\langle\phi,\phi\rangle_{\widetilde{\mathcal{H}}}} = \Lambda_n(\widetilde{T}). \qquad \square
\end{aligned}$$

Definition 5.10 (Comparison of operators). Let T and $\widetilde{T}$ be lower semibounded self-adjoint operators in a Hilbert space $\mathcal{H}$. We write

$$T \le \widetilde{T}$$

if $T \le \widetilde{T}$ via the identity map, which means more precisely that the closed sesquilinear forms t and $\widetilde{t}$ of T and $\widetilde{T}$ satisfy

$$D(\widetilde{t}) \subset D(t), \qquad t(\varphi,\varphi) \le \widetilde{t}(\varphi,\varphi) \text{ for all } \varphi \in D(\widetilde{t}).$$

By applying Corollary 5.9 one arrives at:

Corollary 5.11. *Let T and $\widetilde{T}$ be lower semibounded self-adjoint operators in $\mathcal{H}$ with $T \le \widetilde{T}$. Then $\Lambda_n(T) \le \Lambda_n(\widetilde{T})$ for all $n \in \mathbb{N}$ (and all assertions of Corollary 5.7 hold true).*

Remark 5.12 (Spectral properties of direct sums). The preceding comparison results will be often used for the case when one of the operators is a direct sum. So let us look at this case more attentively.

(A) Let Q and Q' be self-adjoint operators in Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$. Recall that their direct sum $Q \oplus Q'$ is the self-adjoint operator in $\mathcal{H} \times \mathcal{H}'$ given by

$$D(Q\oplus Q') = D(Q) \times D(Q'), \quad (Q\oplus Q')(u,u') = (Qu, Q'u').$$

As discussed in the exercises, one has

$$\operatorname{spec}(Q\oplus Q') = \operatorname{spec} Q \cup \operatorname{spec} Q', \quad \operatorname{spec}_{\mathrm{p}}(Q\oplus Q') = \operatorname{spec}_{\mathrm{p}} Q \cup \operatorname{spec}_{\mathrm{p}} Q'.$$

(B) Remark that a number λ belong to $\operatorname{spec}_{\mathrm{disc}}(Q\oplus Q')$ if and only if

- it is an isolated point of $\operatorname{spec}(Q\oplus Q') \equiv \operatorname{spec} Q \cup \operatorname{spec} Q'$,
- and $\dim\ker\big(Q\oplus Q' - \lambda\big) < \infty$.

Due to

$$\begin{aligned}
\dim\ker\big((Q\oplus Q' - \lambda\big) &= \dim\big(\ker(Q-\lambda) \times \ker(Q'-\lambda)\big)\\
&= \dim\ker(Q-\lambda) + \dim\ker(Q'-\lambda)
\end{aligned}$$

we conclude that $\lambda \in \operatorname{spec}_{\mathrm{disc}}(Q \oplus Q')$ if and only if it is in the discrete spectrum of one of Q, Q' and not in the essential spectrum of the other operator, or, equivalently,

$$\operatorname{spec}_{\mathrm{ess}}(Q \oplus Q') = \operatorname{spec}_{\mathrm{ess}} Q \cup \operatorname{spec}_{\mathrm{ess}} Q'.$$

(C) Recall that if both Q and Q' are semibounded from below and q and q' are closed sesquilinear forms for Q and Q', then $Q \oplus Q'$ is also semibounded from below and its closed sesquilinear form $q \oplus q'$ is given by

$$D(q \oplus q') = D(q) \times D(q'), \quad (q \oplus q')\big((u,u'),(u,u')\big) = q(u,u) + q'(u',u').$$

(D) Furthermore, if for some $\mu \in \mathbb{R}$ one has $Q' \geq \mu$, then Q' has no spectrum if $(-\infty,\mu)$, and the spectrum of $Q \oplus Q'$ in $(-\infty,\mu)$ coincides wit the spectrum of Q in $(-\infty,\mu)$ (this also holds for the discrete and the essential spectrum). In particular,

$$\text{if } Q' \geq \mu, \text{ then } \Lambda_n(Q \oplus Q') = \Lambda_n(Q) \text{ for any } n \in \mathbb{N} \text{ with } \Lambda_n(Q) < \mu.$$

5.3 Basic inequalities for Laplacian eigenvalues

In this section we discuss some application of the general spectral theory to the eigenvalues of the Dirichlet and Neumann Laplacians. Let us introduce the precise setting.

Let $\Omega \subset \mathbb{R}^d$ be a open set, then By definition, the Dirichlet Laplacian $T^D \equiv T^D_\Omega$ and the Neumann Laplacian $T^N \equiv T^N_\Omega$ are the self-adjoint operators in $\mathcal{H} := L^2(\Omega)$ associated with the sesqulinear forms $t_D \equiv t^D_\Omega$ and t^N_Ω respectively,

$$\begin{aligned} t^D(u,v) = t^D_\Omega(u,v) &= \int_\Omega \overline{\nabla u} \cdot \nabla v \,\mathrm{d}x, \quad D(t^D) = Q(T^D) = H^1_0(\Omega), \\ t^N(u,v) = t^N_\Omega(u,v) &= \int_\Omega \overline{\nabla u} \cdot \nabla v \,\mathrm{d}x, \quad D(t^N) = Q(T^N) = H^1(\Omega). \end{aligned}$$

If T^Ω_D resp. T^Ω_N have compact resolvent, we consider their eigenvalues Consider their eigenvalues

$$E^D_n(\Omega) := E_n(T^\Omega_D) \quad \text{resp.} E^N_n(\Omega) := E_n(T^\Omega_N), \quad n \in \mathbb{N},$$

enumerated in the non-decreasing order and taking into accound the multiplicities. These eigenvalues are clearly non-negative (as $T^\Omega_{D/N} \geq 0$), and they are usually referred to as the Dirichlet resp. Neumann eigenvalues of Ω (the presence of the Laplace operator is assumed implicitly).

We would like to discuss various inequalities between these eigenvalues their relations to the geometric object Ω. In view of the min-max principle one has

$$E^{D/N}_n(\Omega) = \Lambda_n(T^\Omega_{D/N}) \text{ for any } n \in \mathbb{N}, \tag{5.3}$$

which delivers the principal method of study.

Proposition 5.13 (First Neumann eigenvalue). *If Ω is a open set of finite measure, then* *(a)* 0 *is an eigenvalue of the Neumann Laplacian* T_N^Ω*, and it is the lower edge of spectrum,*

(b) $\ker T_N^\Omega$ *is spanned by the locally constant functions,*

(c) $\dim\ker T_\Omega^N$ *=the number of connected component of* Ω*.*

(Remark that no compact embedding and no assumption on the boundary are required for the above statements.)

Proof. We abbreviate $T := T_N^\Omega$ and $t := t_N^\Omega$. Due to $T \geq 0$ we have $\operatorname{spec} T \subset [0, \infty)$.

Let v be a function which is constant on each connected component of Ω, then $\nabla v = 0$. Due to $|\Omega| < \infty$ such v belongs to $L^2(\Omega)$, and due to $\nabla v = 0$ one has $v \in H^1(\Omega)$. For any $u \in H^1(\Omega)$ we have

$$t(u, v) = \int_\Omega \overline{\nabla u} \cdot \nabla v \,\mathrm{d}x = \int_\Omega \overline{\nabla u} \cdot \underbrace{\nabla v}_{=0} \,\mathrm{d}x = 0 = \langle u, 0\rangle_{L^2},$$

which means that $v \in D(T_N^\Omega)$ with $Tv = 0$. This shows that $\ker T$ contains all locally constant functions (remark the dimension of the space of local constant functions on Ω is exactly the number of connected components of Ω).

On the other hand, let $u \in \ker T$, then $Tu = -\Delta u$ weakly. If $Tu = 0$, then the elliptic regularity (Theorem 1.54 gives $u \in C^\infty(\Omega)$. Furthermore,

$$0 = \langle u, Tu\rangle = t(u, u) = \int_\Omega |\nabla u|^2 \,\mathrm{d}x = 0,$$

i.e. $\partial_j u = 0$ for each $j \in \{1, \dots, d\}$, which shows that u is constant on each connected component of Ω. $\square$

Proposition 5.14 (First Dirichlet eigenvalue). *If the Dirichlet Laplacian on* Ω *has compact resolvent, then its first eigenvalue is strictly positive.*

Proof. Let $T := T_D^\Omega$ and $t := t_D^\Omega$. Let E be the first eigenvalue, then due to $T \geq 0$ we have at least $E \geq 0$. Let $v \in \ker(T - E) \subset D(T) \subset D(t) = H_0^1(\Omega)$ with $\|v\|_{L^2} = 1$, then

$$E = E\|v\|_{L^2}^2 = E\langle v, v\rangle_{L^2} = \langle v, Tv\rangle_{L^2} = t(v, v) = \int_\Omega |\nabla v|^2 \,\mathrm{d}x.$$

Assume that $E = 0$, then $\nabla v = 0$. Let $\widetilde{v}$ be the extension of v by zero to $\mathbb{R}^d$, then (Prop. 1.67) we have $\widetilde{v} \in H^1(\mathbb{R}^d)$ with $\nabla\widetilde{u} = 0$ and $\|\widetilde{v}\|_{L^2(\mathbb{R}^d)} = 1$. Then for any $u \in H^1(\mathbb{R}^d)$ we have

$$\int_{\mathbb{R}^d} \overline{\nabla u} \cdot \underbrace{\nabla\widetilde{v}}_{=0} \,\mathrm{d}x = 0 = \langle u, 0\rangle_{L^2(\mathbb{R}^d)},$$

which shows that $\widetilde{v}$ an eigenfunction of the free Laplacian in $\mathbb{R}^d$ for the zero eigenvalue .This is a contradiction, as the free Laplacian in $\mathbb{R}^d$ has no eigenvalues. $\square$

Proposition 5.15 (Domain monotonicity for Dirichlet). *If $\widetilde{\Omega} \subset \Omega$ and the embedding $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ is compact, then T_D^Ω and $T_D^{\widetilde{\Omega}}$ have compact resolvents, and $E_n^D(\widetilde{\Omega}) \geq E_n^D(\Omega)$ for all $n \in \mathbb{N}$ (i.e. "smaller domains have higher eigenvalues").*

Proof. Let $J : L^2(\widetilde{\Omega}) \to L^2(\Omega)$ be the operator of extension by zero and $\widetilde{J}$: $L^2(\Omega) \to L^2(\widetilde{\Omega})$ be the operator of restriction to $\widetilde{\Omega}$. Then for any $u \in H_0^1(\widetilde{\Omega})$ we have $\|u\|_{L^2(\widetilde{\Omega})} = \|Ju\|_{L^2(\Omega)}$ and $Ju \in H_0^1(\Omega)$ with and $J\partial_j u = \partial_j Ju$, which shows

$$t_D^\Omega(Ju, Ju) = \int_\Omega |\nabla(Ju)|^2 \, \mathrm{d}x = \int_{\widetilde{\Omega}} |\nabla u|^2 \, \mathrm{d}x = t_D^{\widetilde{\Omega}}(u,u).$$

In other words, $J : H_0^1(\widetilde{\Omega}) \to H_0^1(\Omega)$ is an isometry (and then continuous), and the embedding $H_0^1(\widetilde{\Omega}) \hookrightarrow L^2(\widetilde{\Omega})$ can be decomposed as

$$\widetilde{J} \underbrace{(\text{embedding } H_0^1(\Omega) \hookrightarrow L^1(\Omega))}_{\text{compact}} J = \text{compact}.$$

Therefore, one has compact embedding $H_0^1 \hookrightarrow L^2$ for both Ω and $\widetilde{\Omega}$, which gives the resolvent compactness for T_D^Ω and $T_D^{\widetilde{\Omega}}$. Furthermore; the above equalities means that $T_D^\Omega \leq T_D^\Omega$ via J (Definition 5.8), and the min-max argument (Corollary 5.9) shows that $\Lambda_n(T_D^\Omega) \geq \Lambda_n(T_D^{\widetilde{\Omega}})$ for all $n \in \mathbb{N}$, and in our case $E_n = \Lambda_n$ as we deal with operators with compact resolvents. □

Remark 5.16 (No domain monotonicity for Neumann). The inclusion $\widetilde{\Omega} \subset \Omega$ does not imply any inequality between the respective Neumann eigenvalues. [7]

(a) Let $\widetilde{\Omega} = (0,1) \times (0,1)$ and $\Omega = (0,1) \times (0,2)$, then $\widetilde{\Omega} \subset \Omega$ and

$$E_2^N(\widetilde{\Omega}) = \pi^2 > \frac{\pi^2}{4} = E_2^N(\Omega).$$

(b) Let $t \in (0,1)$ and $\widetilde{\Omega}$ be the rectangle with the vertices $(t,0)$, $(0,t)$, $(1-t,t)$ and $(1,1-t)$, then the side lengths of $\widetilde{\Omega}$ are $\sqrt{2}t$ and $\sqrt{2}(1-t)$. To be definite, assume that $t < \frac{1}{2}$. The eigenvalues of the Neumann Laplacian in $\widetilde{\Omega}$ are

$$\frac{\pi^2 m^2}{2t^2} + \frac{\pi^2 n^2}{2(1-t)^2}, \quad m,n \in \mathbb{N}_0, \quad \text{which gives } E_2^N(\widetilde{\Omega}) = \frac{\pi^2}{2(1-t)^2}.$$

On the other hand, $\widetilde{\Omega} \subset (0,1) \times (0,1) =: \Omega$ and $E_2^N(\widetilde{\Omega}) < \pi^2 = E_2^N(\Omega)$ if t is sufficiently close to 0.

[7] In fact, the resolvent compactness of T_N^Ω does not imply the resolvent compactness of $T_N^{\widetilde{\Omega}}$ (will be considered in the exercises).

Nevertheless, we mention at least one important case for which the domain monotonicity for the Neumann case still holds.

Proposition 5.17 (Restricted domain monotonicity for Neumann). *Let $\widetilde{\Omega}, \Omega \subset \mathbb{R}^d$ be open sets such that $\widetilde{\Omega} \subset \Omega$, the embedding $H^1(\widetilde{\Omega}) \hookrightarrow L^2(\widetilde{\Omega})$ is compact* <u>*and*</u> *$|\Omega \setminus \widetilde{\Omega}| = 0$, then:*

(a) the embedding $H^1(\Omega) \hookrightarrow L^2(\Omega)$ is also compact,

(b) for any $n \in \mathbb{N}$ one has $E_n(\widetilde{\Omega}) \leq E_n(\Omega)$.

Proof. For a function u defined on Ω denote by $\widetilde{u}$ its restriction on $\widetilde{\Omega}$. Due to $|\Omega \setminus \widetilde{\Omega}| = 0$ the map $\widetilde{J} : L^2(\Omega) \ni u \mapsto \widetilde{u} \in L^2(\widetilde{\Omega})$ is an isometry, in particular, continuous. Furthermore, for $u \in H^1(\Omega)$ one has $\widetilde{u} \in H^1(\widetilde{\Omega})$ with $\nabla\widetilde{u} = (\nabla u)|_{\widetilde{\Omega}}$, which means that $\widetilde{J}$ also defines an isometry $H^1(\Omega) \to H^1(\widetilde{\Omega})$. Finally, if J denotes the extension by zero from $\widetilde{\Omega}$ to Ω, then J is also an isometry. The embedding $H^1(\Omega) \hookrightarrow L^2(\Omega)$ can be decomposed as

$$J \underbrace{\big(\text{embedding } H^1(\widetilde{\Omega}) \hookrightarrow L^2(\widetilde{\Omega})\big)}_{\text{compact}} \widetilde{J}$$

and, hence, is a compact operator, and this shows (a). For any $u \in H^1(\Omega)$ we have

$$t_N^{\widetilde{\Omega}}(\widetilde{u}, \widetilde{u}) = \int_{\widetilde{\Omega}} |\nabla \widetilde{u}|^2 \,\mathrm{d}x = \int_{\Omega} |\nabla u|^2 \,\mathrm{d}x = t_N^{\Omega}(u, u)$$

which means that $T_N^{\widetilde{\Omega}} \leq T_N^{\Omega}$ via $\widetilde{J}$, and (b) follows by the comparison principle (Corollary 5.9). □

Finally we obtain the following classical result:

Proposition 5.18 (Neumann ≤ Dirichlet). *If the embedding $H^1(\Omega) \hookrightarrow L^2(\Omega)$ is compact, then the Dirichlet and Neumann Laplacians in Ω have compact resolvent, and for any $n \in \mathbb{N}$ one has $E_n^N(\Omega) \leq E_n^D(\Omega)$.*

Proof. The embedding $H_0^1(\Omega) \hookrightarrow H^1(\Omega)$ is always bounded, which implies the compactness of the embedding $H_0^1(\Omega) \subset L^2(\Omega)$ and the resolvent compactness for both T_N^{Ω} and T_D^{Ω}. The sesquilinear form of the Neumann Laplacian extends the sesquilinear form for the Dirichlet Laplacian, so $T_N^{\Omega} \leq T_D^{\Omega}$ (Definition 5.10), and the claim follows by the min-max principle (Corollary 5.11). □

Remark that if $d = 1$ and $\Omega = (0, \ell)$ with $\ell > 0$, then $E_1(T_N^{\Omega}) = 0$, while

$$E_{n+1}(T_N^{\Omega}) = \frac{\pi^2}{n^2 \ell^2} \equiv E_n(T_D^{\Omega}) \text{ for any } n \in \mathbb{N},$$

which is much stronger than the statement of Prop. 5.18. In higher dimensions, there are very few Ω for which the eigenvalues can be computed and compared explicitly, but, in fact, an even stronger inequality can be proved:

Theorem 5.19 (Friedlander's inequality[8]). *Let $d \geq 2$ and $\Omega \subset \mathbb{R}^d$ be an open set such that the embedding $H^1(\Omega) \hookrightarrow L^2(\Omega)$ is compact* <u>*and*</u> *$|\Omega| < \infty$, then*

$$E_{n+1}^N(\Omega) < E_n^D(\Omega) \text{ for any } n \in \mathbb{N}.$$

The proof is accessible but will not be discussed in detail here.

5.4 Weyl asymptotics for Dirichlet Laplacians

In this subsection we will discuss some aspects of the asymptotic behavior of the Laplacian eigenvalues $E_n^D(\Omega)$ as n becomes large.

For $\Omega \subset \mathbb{R}^d$ satisfying the "compactness assumption"

$$\text{the embedding } H^1(\Omega) \hookrightarrow L^2(\Omega) \text{ is compact} \tag{5.4}$$

one introduces the counting functions $\lambda \mapsto \mathcal{N}_{D/N}(\lambda, \Omega)$ as follows: for each $\lambda \in \mathbb{R}$,

$$\mathcal{N}_{D/N}(\lambda, \Omega) = \text{the number of } n \in \mathbb{N} \text{ for which } E_n^{D/N}(\Omega) \in (-\infty, \lambda].$$

Clearly, $\mathcal{N}_{D/N}(\lambda, \Omega)$ is finite for any λ, and the function $\lambda \mapsto \mathcal{N}_{D/N}(\lambda, \Omega)$ in non-decreasing, with values in $\mathbb{N}_0$ (there it is piecewise constant), and has a jump at each eigenvalue of $T_{D/N}^{\Omega}$ (the jump is equal to the multiplicity of the eigenvalue). Our main result is:

Theorem 5.20 (Weyl asymptotics for Dirichlet eigenvalues). *For any bounded open subset $\Omega \subset \mathbb{R}^d$ with $|\partial\Omega| = 0$ we have*

$$\lim_{\lambda \to +\infty} \frac{\mathcal{N}_D(\lambda, \Omega)}{\lambda^{\frac{d}{2}}} = \frac{\omega_d}{(2\pi)^d} |\Omega|,$$

where ω_d denotes the volume of the unit ball in $\mathbb{R}^d$.

To keep simple notation we proceed with the proof for the case $d = 2$ only. Due to $\omega_2 = \pi$ we are reduced to prove

$$\lim_{\lambda \to +\infty} \frac{\mathcal{N}_D(\lambda, \Omega)}{\lambda} = \frac{|\Omega|}{4\pi}. \tag{5.5}$$

The proof consists of several steps.

Lemma 5.21. *If Ω is a rectangle, then*

$$\lim_{\lambda \to +\infty} \frac{\mathcal{N}_{D/N}(\lambda, \Omega)}{\lambda} = \frac{|\Omega|}{4\pi}.$$

[8] The result was first proved in the paper L. Friedlander, Some inequalities between Dirichlet and Neumann eigenvalues. Arch. Rat. Mech. Anal. **116** (1991) 153–160 for domains with smooth boundaries. The proof we present here is from the paper N. Filonov, On an inequality between Dirichlet and Neumann eigenvalues for the Laplace operator. St. Petersburg Math. J. **16** (2005) 413–416.

Proof. As the spectra of $T^{\Omega}_{D/N}$ are invariant under isometries of Ω, without loss of generality consider $\Omega = (0, a) \times (0, b)$ with $a, b > 0$. As discussed in the exercises, the Neumann eigenvalues of Ω are the numbers

$$\lambda(m, n) := \Big(\frac{\pi m}{a}\Big)^2 + \Big(\frac{\pi n}{b}\Big)^2$$

with $m, n \in \mathbb{N}_0$, and the Dirichlet spectrum consists of the eigenvalues $\lambda(m, n)$ with $m, n \in \mathbb{N}$. Denote

$$S_\lambda := \Big\{(x, y) \in \mathbb{R}^2 : \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} \le \frac{\lambda}{\pi^2},\ x \ge 0,\ y \ge 0\Big\},$$

then $\mathcal{N}_D(\lambda, \Omega) = \# S_\lambda \cap \big(\mathbb{N} \times \mathbb{N}\big)$ and $\mathcal{N}_N(\lambda, \Omega) = \# S_\lambda \cap \big(\mathbb{N}_0 \times \mathbb{N}_0\big)$.

First, counting the points $(n, 0)$ and $(0, n)$ with $n \in \mathbb{N}_0$ that lie in S_λ we obtain the upper bound

$$\mathcal{N}_N(\lambda, \Omega) - \mathcal{N}_D(\lambda, \Omega) \le \frac{a + b}{\pi} \sqrt{\lambda} + 2, \quad \lambda > 0.$$

At the same time, S_λ contains the union of the unit squares $[m - 1, m] \times [n - 1, n]$ with $(m, n) \in S_\lambda \cap \big(\mathbb{N} \times \mathbb{N}\big)$. As there are exactly $\mathcal{N}_D(\lambda, \Omega)$ such squares, we have

$$\mathcal{N}_D(\lambda, \Omega) \le |S_\lambda| = \frac{\lambda ab}{4\pi}.$$

We also observe that S_λ is contained in the union of the unit squares $[m, m + 1] \times [n, n + 1]$ with $(m, n) \in S_\lambda \cap \big(\mathbb{N}_0 \times \mathbb{N}_0\big)$. As the number of such squares is exactly $\mathcal{N}_N(\lambda, \Omega)$, this gives

$$\mathcal{N}_N(\lambda, \Omega) \ge |S_\lambda| = \frac{\lambda ab}{4\pi}.$$

Putting all together we arrive at

$$\frac{\lambda ab}{4\pi} \le \mathcal{N}_N(\lambda, \Omega) \le \mathcal{N}_D(\lambda, \Omega) + \frac{a + b}{\pi} \sqrt{\lambda} + 2 \le \frac{\lambda ab}{4\pi} + \frac{a + b}{\pi} \sqrt{\lambda} + 2.$$

Then

$$\begin{aligned} \frac{ab}{4\pi} \le \frac{\mathcal{N}_N(\lambda, \Omega)}{\lambda} &\le \frac{ab}{4\pi} + \frac{a + b}{\pi} \frac{1}{\sqrt{\lambda}} + \frac{2}{\lambda}, \\ \frac{ab}{4\pi} + \frac{a + b}{\pi} \frac{1}{\sqrt{\lambda}} + \frac{2}{\lambda} \le \frac{\mathcal{N}_D(\lambda, \Omega)}{\lambda} &\le \frac{ab}{4\pi}, \end{aligned}$$

and it remains to recall that $|\Omega| = ab$. □

Remark 5.22 (Dirichlet-Neumann bracketing). Now let us discuss in greater details the behavior of $\mathcal{N}(\lambda, \Omega)$ with respect to Ω.

(A) By Proposition 5.18 one has $\mathcal{N}_D(\lambda, \Omega) \le \mathcal{N}_N(\lambda, \Omega)$.

(B) The domain monotonicity shows that $N_D(\lambda, \widetilde{\Omega}) \le \mathcal{N}_D(\lambda, \Omega)$ for $\widetilde{\Omega} \subset \Omega$.

(C) Let Ω_1, Ω_2 be two <u>disjoint</u> open sets, then

- $T_N^{\Omega_1\cup\Omega_2}$ is unitarily equivalent to $T_N^{\Omega_1}\oplus T_N^{\Omega_2}$,
- $T_D^{\Omega_1\cup\Omega_2}$ is unitarily equivalent to $T_D^{\Omega_1}\oplus T_D^{\Omega_2}$.

In fact, if one introduces the unitary map

$$\Theta : L^2(\Omega)\to L^2(\Omega_1)\times L^2(\Omega_2),\quad \Theta u := (u_1,u_2) \text{ with } u_j := u|_{\Omega_j},$$

then

$$\begin{aligned}
\Theta D(t_N^{\Omega_1\cup\Omega_2}) &= H^1(\Omega_1\cup\Omega_2) = H^1(\Omega_1)\times H^1(\Omega_2) = D(t_N^{\Omega_1}\oplus t_N^{\Omega_2}),\\
t_N^{\Omega_1\cup\Omega_2}(u,u) &= \int_{\Omega_1\cup\Omega_2} |\nabla u|^2\,\mathrm{d}x = \int_{\Omega_1} |\nabla u|^2\,\mathrm{d}x + \int_{\Omega_2} |\nabla u|^2\,\mathrm{d}x\\
&= \int_{\Omega_1} |\nabla u_1|^2\,\mathrm{d}x + \int_{\Omega_2} |\nabla u_2|^2\,\mathrm{d}x = t_N^{\Omega_1}(u_1,u_1)+t_N^{\Omega_2}(u_2,u_2)\\
&= (t_N^{\Omega_1}\oplus t_N^{\Omega_2})\big((u_1,u_2),(u_1,u_2)\big) = (t_N^{\Omega_1}\oplus t_N^{\Omega_2})(\Theta u,\Theta u),
\end{aligned}$$

which shows the unitary equivalent for the Neumann case, and the Dirichlet case is considered in the same way. In particular, if for both Ω_1 and Ω_2 satisfy the compactness assumption (5.4), then the same holds for $\Omega_1\cup\Omega_2$ too, and

$$\mathcal{N}_D(\lambda,\Omega_1\cup\Omega_2) = \mathcal{N}_D(\lambda,\Omega_1)+\mathcal{N}_D(\lambda,\Omega_2),\quad \mathcal{N}_N(\lambda,\Omega_1\cup\Omega_2) = \mathcal{N}_N(\lambda,\Omega_1)+\mathcal{N}_N(\lambda,\Omega_2).$$

(D) Now assume that Ω_1,Ω_2 are disjoint open subsets of Ω, both satisfying (5.4), such that $|\Omega\setminus(\Omega_1\cup\Omega_2)|=0$. Then using the above considerations and the restricted domain monotonicity for the Neumann eigenvalues (Prop. 5.17) we arrive at the so-called *Dirichlet-Neumann bracketing*

$$\begin{aligned}
\mathcal{N}_D(\lambda,\Omega_1)+\mathcal{N}_D(\lambda,\Omega_2) = \mathcal{N}_D(\lambda,\Omega_1\cup\Omega_2) &\le \mathcal{N}_D(\lambda,\Omega)\le \mathcal{N}_N(\lambda,\Omega)\\
&\le \mathcal{N}_N(\lambda,\Omega_1\cup\Omega_2) = \mathcal{N}_N(\lambda,\Omega_1)+\mathcal{N}_N(\lambda,\Omega_2),
\end{aligned} \tag{5.6}$$

which will be our main argument below. By iterations this extends to finitely many disjoint unions.

Definition 5.23 (Open sets composed from rectangles). We say that an open set Ω is *composed from rectangles* if there exists a finite family of disjoint open rectangles $\Omega_j=(a_j,b_j)\times(c_j,b_j)\subset\Omega$, $j=1,\dots,k$, such that $\Big|\Omega\setminus\bigcup_{j=1}^k\Omega_j\Big|=0$.

We remark that the unions and intersections of finitely many open sets composed from rectangles are again open sets composed from rectangles. Furthermore, the discussion in Proposition 5.17 and Remark 5.22 shows that any open composed from rectangles satsfies the compactness assumption (5.4).

Lemma 5.24. *The Weyl asymptotics* (5.5) *holds for* Ω *composed from rectangles.*

Proof. Let Ω_j, $j=1,\dots,k$ be rectangles as in Definition 5.23, then

$$\Big|\Omega\setminus\bigcup_{j=1}^k\Omega_j\Big|=0,\qquad |\Omega_1|+\dots+|\Omega_k|=|\Omega|.$$

Using the Dirichlet-Neumann bracketing (5.6) we obtain the chain

$$\frac{\mathcal{N}_D(\lambda,\Omega_1)+\dots+\mathcal{N}_N(\lambda,\Omega_k)}{\lambda}=\frac{\mathcal{N}_D(\lambda,\Omega_1\cup\dots\cup\Omega_k)}{\lambda}\leq\frac{\mathcal{N}_D(\lambda,\Omega)}{\lambda}$$
$$\leq\frac{\mathcal{N}_N(\lambda,\Omega)}{\lambda}\leq\frac{\mathcal{N}_N(\lambda,\Omega_1\cup\dots\cup\Omega_k)}{\lambda}=\frac{\mathcal{N}_N(\lambda,\Omega_1)+\dots+\mathcal{N}_D(\lambda,\Omega_k)}{\lambda},$$

and it remains to use

$$\frac{\mathcal{N}_{D/N}(\lambda,\Omega_j)}{\lambda}\xrightarrow{\lambda\to+\infty}\frac{|\Omega_j|}{4\pi}\text{ by Lemma 5.21.}\qquad\square$$

Lemma 5.25 (Approximation by rectangles). *Let Ω be bounded with $|\partial\Omega|=0$. Then for any $\varepsilon>0$ one can find two bounded open sets $\widetilde{\Omega}_\varepsilon\subset\Omega\subset\Omega_\varepsilon$, both composed from rectangles, such that $|\Omega_\varepsilon\setminus\widetilde{\Omega}_\varepsilon|<\varepsilon$.*

Proof. Let $\varepsilon>0$, then the condition $|\partial\Omega|=0$ means that one can cover $\partial\Omega$ by a family of rectangles $\omega_j=(a_j,b_j)\times(c_j,b_j)$ with $\sum_{j=1}^\infty|\omega_j|<\varepsilon$. As $\partial\Omega$ is compact (because Ω is bounded), there is a finite subcovering, so let $n\in\mathbb{N}$ with

$$\partial\Omega\subset\bigcup_{j=1}^n\omega_j=:\omega,\quad\sum_{j=1}^n|\omega_j|<\varepsilon,\quad\text{then }|\omega|<\varepsilon.$$

Pick $R>0$ such that the square $S:=(-R,R)\times(-R,R)$ contains Ω and ω. The open set $W:=S\setminus\overline{\omega}$ is composed from rectangles, so let $W_1,\dots,W_N\subset W$ be mutually disjoint rectangles with $|W\setminus\bigcup_{k=1}^N W_k|=0$. Note that none of W_k intersects $\partial\Omega$ and, moreover, each W_k is contained either in Ω or in $\overline{\Omega}^{\mathsf{C}}$ (if the intersection of W_k with each of these two sets is non-empty, then W_k becomes a disjoint union of two non-empty open sets, which is impossible as W_k is connected). So let

$$K:=\big\{k\in\{1,\dots,N\}:\ W_k\subset\Omega\big\},\quad K':=\{1,\dots,N\}\setminus K,$$

then

$$\widetilde{\Omega}_\varepsilon:=\bigcup_{k\in K}W_j\subset\Omega\subset S\setminus\overline{\bigcup_{k\in K'}W_k}=:\Omega_\varepsilon.$$

The three mutually disjoint open sets

$$\bigcup_{k\in K}W_j,\quad\omega,\quad\bigcup_{k\in K'}W_k$$

exhausts S up to zero measure sets (recall that the boundaries of sets composed from rectangles have zero measure), which means that $|\Omega_\varepsilon\setminus\widetilde{\Omega}_\varepsilon|=|\omega|<\varepsilon$. $\square$

Proof of the Weyl asymptotics (Theorem 5.20). Let $\varepsilon>0$. By Lemma 5.25 one can find two bounded open sets $\widetilde{\Omega}_\varepsilon\subset\Omega\subset\Omega_\varepsilon$, both composed from rectangles, such that $|\Omega_\varepsilon\setminus\widetilde{\Omega}_\varepsilon|<\varepsilon$, then $|\Omega|-\varepsilon\leq|\widetilde{\Omega}_\varepsilon|\leq|\Omega|\leq|\widetilde{\Omega}_\varepsilon|\leq|\Omega|+\varepsilon$. Due to the domain monotonicity one has, for any $\lambda>0$,

$$\frac{\mathcal{N}_D(\lambda,\widetilde{\Omega}_\varepsilon)}{\lambda}\leq\frac{\mathcal{N}_D(\lambda,\Omega)}{\lambda}\leq\frac{\mathcal{N}_D(\lambda,\Omega_\varepsilon)}{\lambda}.$$

The terms with $\widetilde{\Omega}_\varepsilon$ and Ω_ε are covered by Lemma 5.24 and one obtains

$$\frac{|\Omega|-\varepsilon}{4\pi} \le \frac{|\widetilde{\Omega}_\varepsilon|}{4\pi} \le \liminf_{\lambda\to+\infty} \frac{\mathcal{N}_D(\lambda,\Omega)}{\lambda} \le \limsup_{\lambda\to+\infty} \frac{\mathcal{N}_D(\lambda,\Omega)}{\lambda} \le \frac{|\Omega_\varepsilon|}{4\pi} \le \frac{|\Omega|+\varepsilon}{4\pi}.$$

As ε was arbitrary, the theorem is proved. □

We note that the Weyl asymptotics also holds for the Neumann Laplacian if the domain is sufficiently smooth, which can be proved using suitable extension theorem for Sobolev spaces. The Weyl asymptotics is one of the basic results on the relations between the Dirichlet/Neumann eigenvalues and the geometric properties of the domain. It states, in particular, that the Dirichlet eigenvalues contain the information on the dimension and the volume of the domain. There are various refinements involving lower order terms with respect to λ, and the respective coefficients contain some information on the topology of the domain, on its boundary etc.

5.5 Discrete spectrum of Schrödinger operators

As seen above in Theorem 4.26, if V is a Kato class potential on $\mathbb{R}^d$, then the associated Schrödinger operator $T = -\Delta + V$ in $L^2(\mathbb{R}^d)$ acting is semibounded below with $\operatorname{spec}_{\text{ess}} T = [0, +\infty)$ and $\Sigma = 0$. It means for the spectrum of T in $(-\infty, 0)$ (which is usually called *negative spectrum*) there are only three possibilities:

- the negative spectrum is empty,
- the negative spectrum spectrum consists of finitely many eigenvalues of finite multiplicites,
- the negative spectrum consists of infinitely many eigenvalues of finite multilplicities, and these eigenvalues form a sequence converging to 0 (no other accumulation point is possible, as any accumulation point of the spectrum belongs to the essential spectrum).

In this section we discuss some conditions on V allowing one to understand which of these options is realized. We remark first that the condition $V \geq 0$ (for V from a Kato class) excludes the existence of negative eigenvalues: in this case one has $T \geq 0$, which finally gives $\operatorname{spec} T = [0, \infty)$, so the above questions only make sense if V takes negative eigenvalues.

For the existence of negative eigenvalues we make first a very simple observation:

Lemma 5.26. *Let T be a lower semibounded self-adjoint operator generated by a closed sesquilinear form t. If there exists $0 \neq u \in D(t)$ with $t(u, u) < \Sigma \|u\|^2$, then T has at least one eigenvalue in $(-\infty, \Sigma)$.*

Proof. One has

$$\Lambda_1(T) \leq \frac{t(u, u)}{\|u\|^2} < \Sigma,$$

which shows that $\Lambda_1(T)$ is the first eigenvalue of T. □

We have a simple sufficient condition for the one- and two-dimensional cases.

Theorem 5.27 (Existence of negative eigenvalues in 1D and 2D). *Let $d \in \{1, 2\}$ and $V \in L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ be real-valued such that*

$$\int_{\mathbb{R}^d} V(x)\, \mathrm{d}x < 0, \tag{5.7}$$

then the Schrödinger operator $T = -\Delta + V$ has at least one negative eigenvalue.

Proof. The potential V is in the Kato class, which shows that the operator T is semibounded from below and its essential spectrum is $[0, \infty)$, see Theorems 4.13 and 4.26, and $\Sigma(T) = 0$. In view of Lemma 5.26 it is sufficient to show that there exists $u \in D(T)$ with $t(u, u) < 0$.

Case $d = 1$. Let $\varphi \in C_c^\infty(\mathbb{R})$ with $\varphi(0) = 1$ and

$$u_\varepsilon : \mathbb{R} \ni x \mapsto \varphi(\varepsilon x).$$

Then $u_\varepsilon \in C_c^\infty(\mathbb{R}) \subset D(T)$, and due to $|V||u_\varepsilon|^2 \le |V|\|\varphi\|_\infty^2 \in L^1(\mathbb{R})$ the dominated convergence gives

$$\int_{\mathbb{R}} V(x)\big|u_\varepsilon(x)\big|^2 \,\mathrm{d}x = \int_{\mathbb{R}} V(x)\big|\varphi(\varepsilon x)\big|^2 \,\mathrm{d}x \overset{\varepsilon\to 0^+}{\longrightarrow} \int_{\mathbb{R}} V(x)\,\mathrm{d}x < 0,$$

while

$$\begin{aligned}\int_{\mathbb{R}} \big|u_\varepsilon'(x)\big|^2 \,\mathrm{d}x &= \varepsilon^2 \int_{\mathbb{R}} \big|\varphi'(\varepsilon x)\big|^2 \,\mathrm{d}x \quad \text{Substitution } y = \varepsilon x\\ &= \varepsilon^2 \cdot \frac{1}{\varepsilon}\int \mathbb{R}\big|\varphi'(y)\big|^2 \,\mathrm{d}y = \varepsilon\,\|\varphi'\|^2_{L^2(\mathbb{R})} = O(\varepsilon),\end{aligned}$$

which shows that

$$t(u_\varepsilon,u_\varepsilon) = \int_{\mathbb{R}} |u_\varepsilon'(x)|^2 \,\mathrm{d}x + \int_{\mathbb{R}} V(x)|u_\varepsilon(x)|^2 \,\mathrm{d}x \overset{\varepsilon\to 0^+}{\longrightarrow} V_0 < 0,$$

and $t(u_\varepsilon,u_\varepsilon) < 0$ if $\varepsilon > 0$ is sufficiently small.

Case $d = 2$. Pick $\varphi \in C^\infty(\mathbb{R})$ such that $\varphi(t) = 1$ for all $t \le 1$ and $\varphi(t) = 0$ for $|t| > 2$. Then the functions

$$u_\varepsilon : \mathbb{R}^2 \ni x \mapsto \varphi\big(\varepsilon \ln|x|\big)$$

belong to $C_c^\infty(\mathbb{R}^2) \subset D(T)$ with $u_\varepsilon(x) = 1$ for $|x| \le e^{1/\varepsilon}$ and $u_\varepsilon(x) = 0$ for $|x| \ge e^{2/\varepsilon}$. As in the case $d = 1$ one easily obtains

$$\lim_{\varepsilon\to 0^+} \int_{\mathbb{R}^2} V(x)|u_\varepsilon(x)|^2 \,\mathrm{d}x = \int_{\mathbb{R}^2} V(x)\,\mathrm{d}x < 0,$$

Furthermore,

$$\begin{aligned}\partial_j u_\varepsilon(x) &= \frac{\varepsilon x_j}{|x|^2}\,\varphi'\big(\varepsilon\ln|x|\big), \qquad \big|\nabla u_\varepsilon(x)\big|^2 = \frac{\varepsilon^2}{|x|^2}\Big|\varphi'\big(\varepsilon\ln|x|\big)\Big|^2,\\ \int_{\mathbb{R}^2} |\nabla u_\varepsilon(x)|^2 \,\mathrm{d}x &= \varepsilon^2 \int_{\mathbb{R}^2} \frac{1}{|x|^2}\Big|\varphi'\big(\varepsilon\ln|x|\big)\Big|^2 \,\mathrm{d}x\\ &= \varepsilon^2 \int_{e^{1/\varepsilon}\le|x|\le e^{2/\varepsilon}} \frac{1}{|x|^2}\Big|\varphi'\big(\varepsilon\ln|x|\big)\Big|^2 \,\mathrm{d}x\\ \text{(use polar coordinates)} \ &= 2\pi\varepsilon^2 \int_{e^{1/\varepsilon}}^{e^{2/\varepsilon}} \frac{|\varphi'(\varepsilon\ln r)|^2}{r}\,\mathrm{d}r \le 2\pi\|\varphi'\|_\infty^2\varepsilon^2 \int_{e^{1/\varepsilon}}^{e^{2/\varepsilon}} \frac{\mathrm{d}r}{r}.\end{aligned}$$

The last integral can be directly computed:

$$\int_{e^{1/\varepsilon}}^{e^{2/\varepsilon}} \frac{\mathrm{d}r}{r} = \big[\ln r\big]_{r=e^{1/\varepsilon}}^{r=e^{2/\varepsilon}} = \frac{2}{\varepsilon} - \frac{1}{\varepsilon} = \frac{1}{\varepsilon},$$

which gives

$$\int_{\mathbb{R}^2} |\nabla u_\varepsilon(x)|^2 \,\mathrm{d}x = O(\varepsilon).$$

It follows that

$$t(u_\varepsilon,u_\varepsilon)=\int_{\mathbb{R}^2}|\nabla u_\varepsilon(x)|^2\,\mathrm{d}x+\int_{\mathbb{R}^2}V(x)|u_\varepsilon(x)|^2\,\mathrm{d}x \xrightarrow{\varepsilon\to 0^+} \int_{\mathbb{R}^2}V(x)\,\mathrm{d}x << 0,$$

and $t(u_\varepsilon,u_\varepsilon)<0$ is $\varepsilon>0$ is chosen sufficiently small. $\square$

Remark 5.28. The assumption (5.7) can be satisfied for very small V. For example, if (5.7) holds for some V, then it also holds for λV with any $\lambda>0$. It follows that for any $\lambda>0$ the Schrödinger operator $-\Delta+\lambda V$ has the essential spectrum $[0,\infty)$ and at least one negative eigenvalue.

It is remarkable that Theorem 5.27 cannot be extended to higher dimensions:

Proposition 5.29 (Absense of eigenvalues for small potentials in higher dimensions). *Let $d\geq 3$ and $V:\mathbb{R}^d\to\mathbb{R}$ be bounded and such that*

$$V(x)=O\Big(\frac{1}{|x|^2}\Big) \text{ as } |x|\to\infty.$$

For $\lambda\in\mathbb{R}$ consider $T_\lambda:=-\Delta+\lambda V$ in $L^2(\mathbb{R}^d)$. Then there is $\lambda_0>0$ with

$$\operatorname{spec}T_\lambda=[0,+\infty) \text{ for all } \lambda\in(-\lambda_0,\lambda_0).$$

Proof. Due to the assumptions on V one can find a constant $C>0$ such that

$$\big|V(x)\big|\leq\frac{C}{|x|^2} \text{ for all } x\in\mathbb{R}^d.$$

Remark first that T_λ is essentially self-adjoint on $C_c^\infty(\mathbb{R}^d)$. For $R>0$ one has the representation $V=1_{|x|\leq R}V+1_{|x|>R}V$ in which the first summand is in $L^2(\mathbb{R}^d)$ and the sup-norm of the second summand is $\leq C/R^2$ and can be made arbitrarily small by choosing a sufficiently large $R>0$. This shows that λV is in Kato class and $\operatorname{spec}_{\mathrm{ess}}T_\lambda=[0,\infty)$ for any $\lambda\in\mathbb{R}$ (Theorem 4.26). It remains to show the inclusion $\operatorname{spec}T_\lambda\subset[0,\infty)$ if λ is sufficiently small.

Recall the Hardy inequality (Proposition 1.81): for any $u\in C_c^\infty(\mathbb{R}^d)$ we have

$$\frac{(d-2)^2}{4}\int_{\mathbb{R}^d}\frac{|u(x)|^2}{|x|^2}\,\mathrm{d}x\leq\int_{\mathbb{R}^d}|\nabla u(x)|^2\,\mathrm{d}x.$$

Let $\lambda_0:=(d-2)^2/(4C)$ and $\lambda\in(-\lambda_0,\lambda_0)$. For any $u\in C_c^\infty(\mathbb{R}^d)$ we have

$$\begin{aligned}\langle u,T_\lambda u\rangle&=\int_{\mathbb{R}^d}\big|\nabla u(x)\big|^2\,\mathrm{d}x+\lambda\int_{\mathbb{R}^d}V(x)\big|u(x)\big|^2dx\\&\geq\int_{\mathbb{R}^d}\big|\nabla u(x)\big|^2dx-|\lambda|\int_{\mathbb{R}^d}\big|V(x)\big|\,\big|u(x)\big|^2\,\mathrm{d}x\\&\geq\int_{\mathbb{R}^d}\big|\nabla u(x)\big|^2dx-\underbrace{|\lambda|C}_{\leq\frac{(d-2)^2}{4}}\int_{\mathbb{R}^d}\frac{\big|u(x)\big|^2}{|x^2|}\,\mathrm{d}x\\&\geq\int_{\mathbb{R}^d}\big|\nabla u(x)\big|^2dx-\frac{(d-2)^2}{4}\int_{\mathbb{R}^d}\frac{|u(x)|^2}{|x|^2}\,\mathrm{d}x\geq 0.\end{aligned}$$

the inequality extends to all $u\in D(T_\lambda)$. Hence, $T_\lambda\geq 0$ and $\operatorname{spec}T_\lambda\subset[0,+\infty)$. $\square$

Theorem 5.30 (Discrete spectrum for potential wells). *Let $V \in L^\infty_{\rm loc}(\mathbb{R}^d)$ be semibounded from below and let $E \in \mathbb{R}$ be such that the set $S = \{x : V(x) < E\}$ is bounded.*[9] *Then the spectrum of $T = -\Delta + V$ in $(-\infty, E)$ is purely discrete and consists of (at most) finitely many eigenvalues.*

Proof. The operator T is essentially self-adjoint on $C_c^\infty(\mathbb{R}^d)$, see Theorem 4.16 and, at the same time, it is generated by its closed sesquilinear form t defined on $H^1_V(\mathbb{R}^d)$:

$$t(u,u) = \int_{\mathbb{R}^d} \big(|\nabla u|^2 + V|u|^2\big)\,\mathrm{d}x.$$

Let B be an open ball containing the above set S. The idea is to "decouple" the two sides of ∂B and to compare T with the direct sum of two operators acting in the Hilbert spaces $\mathcal{G} := L^2(B)$ and $\mathcal{G}' := L^2(\mathbb{R}^d \setminus \overline{B})$.

Consider the following closed sesquilinear form $\widetilde{t}$ extending t:

$$\begin{aligned}\widetilde{t}(u,u) &= \int_{\mathbb{R}^d\setminus\partial B} \big(|\nabla u|^2 + V|u|^2\big)\,\mathrm{d}x,\\ D(\widetilde{t}) &= \big\{u \in H^1(\mathbb{R}^d\setminus\partial B) : \int_{\mathbb{R}^d} V|u|^2\,dx < \infty\big\},\end{aligned}$$

and let $\widetilde{T}$ be the self-adjoint operator generated by $\widetilde{t}$. Then $\widetilde{T} \le T$ (see Definition 5.10) and $\Lambda_n(\widetilde{T}) \le \Lambda_n(T)$ for any $n \in \mathbb{N}$.

Remark that $\mathbb{R}^d \setminus \partial B$ consists of two connected components B and $\mathbb{R}^d \setminus \overline{B}$. Consider the unitary transform

$$\Theta : L^2(\mathbb{R}^d) \mapsto \mathcal{G}\times\mathcal{G}', \quad \Theta u = (u|_B, u|_{\mathbb{R}^d\setminus\overline{B}}),$$

then one easily sees (Remark 5.22) that $\Theta\widetilde{T}\Theta^{-1} = Q \oplus Q'$, where:

- Q is the self-adjoint operator in $\mathcal{G}$ given by its sesquilinear form

  $$q(u,u) = \int_B \big(|\nabla u|^2 + V|u|^2\big)\,\mathrm{d}x, \quad D(q) = H^1(B);$$

  we recall that V is bounded on B,

- Q' is the self-adjoint operator in $\mathcal{G}'$ given by its sesquilinear form

  $$\begin{aligned}q'(u,u) &= \int_{\mathbb{R}^d\setminus\overline{B}} \big(|\nabla u|^2 + V|u|^2\big)\,\mathrm{d}x,\\ D(q') &= \Big\{u \in H^1(\mathbb{R}^d\setminus\overline{B}) : \int_{\mathbb{R}^d\setminus\overline{B}} V|u|^2\,\mathrm{d}x < \infty\Big\},\end{aligned}$$

and then $\Lambda_n(\widetilde{T}_n) = \Lambda_n(Q \oplus Q')$. So we have proved that

$$\Lambda_n(T) \ge \Lambda_n(Q\oplus Q') \text{ for any } n \in \mathbb{N}. \tag{5.8}$$

Now remark that:

[9] In quantum mechanics one says that V is a finite potential well below E.

- the operator Q has compact resolvent, as $D(q)$ is compactly embedded in $\mathcal{G}$. Then $\Lambda_n(Q)$ is the nth eigenvalue of Q, and the eigenvalues converge to $+\infty$, and the number N of eigenvalues of Q in $(-\infty, E)$ is finite.
- in $\mathbb{R}^d \setminus \overline{B}$ we have $V \geq E$, which gives $Q' \geq E$, and Q' has no spectrum in $(-\infty, E)$.

It follows that the spectrum of $Q \oplus Q'$ in $(-\infty, E)$ consists of exactly N eigenvalues (=the first N eigenvalues of Q), which gives $\Lambda_{N+1}(Q \oplus Q') \geq E$. By (5.8) one has $\Lambda_{N+1}(T) \geq E$, and the spectrum of T in $(-\infty, E)$ consists of at most N eigenvalues. □

Corollary 5.31 (Compactly supported potentials). *Let $V \in L^\infty(\mathbb{R}^d)$ be real-valued with compact support and $T = -\Delta + V$ in $L^2(\mathbb{R}^d)$, then* $\operatorname{spec}_{\text{ess}} T = [0, \infty)$, *and T has at most finitely many negative eigenvalues.*

Proof. The potential V is in Kato class, hence, $\operatorname{spec}_{\text{ess}} H = [0, +\infty)$ (see Subsection 4.5), and the finiteness of the discrete spectrum follows by Theorem 5.30 with $E = 0$. □

In fact, one can show the finiteness of the negative discrete spectrum under a weaker assumption that the potentials decay "rapidly" at infinity (the condition to have a compact support in an "extreme" version of such a decay): this will considered in the exercises.

Theorem 5.32 (Potentials producing infinite discrete spectrum). *Let V be a Kato class potential in $\mathbb{R}^d$ and $T = -\Delta + V$ in $L^2(\mathbb{R}^d)$. Assume that for some $R > 0$, $c > 0$ and $p \in (0, 2)$ one has*

$$V(x) \leq -\frac{c}{|x|^p} \text{ for all } x \in \mathbb{R}^d \text{ with } |x| \geq R.$$

Then $\operatorname{spec}_{\text{ess}} T = [0, +\infty)$ *and T has* infinitely many *negative eigenvalues.*

Proof. The equality for the essential spectrum is already proved (Theorem 4.26), and one simply needs to show that $\Lambda_n(T) < 0 \equiv \Sigma(T)$ for any $n \in \mathbb{N}$.

Pick any $\varphi \in C_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp} \varphi \subset \{x \in \mathbb{R}^d : R < |x| < 2R\}$ and $\|\varphi\|_{L^2(\mathbb{R}^d)} = 1$. For $t > 1$ consider the functions $\varphi_t(x) = t^{-d/2}\varphi(x/t)$, then $\varphi_t \in C_c^\infty(\mathbb{R}^d)$ with $\|\varphi_t\|_{L^2(\mathbb{R}^d)} = 1$ and $\operatorname{supp} \varphi_t \subset \{x \in \mathbb{R}^d : tR < |x| < 2tR\}$. We compute

$$\begin{aligned}
\int_{\mathbb{R}^d} |\nabla \varphi_t|^2 \,\mathrm{d}x &= \int_{\mathbb{R}^d} \frac{1}{t^{2+d}} \Big|(\nabla\varphi)\Big(\frac{x}{t}\Big)\Big|^2 \,\mathrm{d}x = (\text{Substitute } x = ty) = \\
&= \frac{1}{t^2} \underbrace{\int_{\mathbb{R}^d} |\nabla \varphi(y)|^2 \,\mathrm{d}y}_{=:a>0} \equiv \frac{a}{t^2}, \\
\int_{\mathbb{R}^d} V|\varphi_t|^2 \,\mathrm{d}x &= \frac{1}{t^d} \int\limits_{tR<|x|<2tR} V(x) \Big|\varphi\Big(\frac{x}{t}\Big)\Big|^2 \,\mathrm{d}x
\end{aligned}$$

$$\leq \frac{1}{t^d} \int\limits_{tR<|x|<2tR} \Big(-\frac{c}{|x|^p}\Big) \Big| \varphi\Big(\frac{x}{t}\Big)\Big|^2 \,\mathrm{d}x = (\text{Substitute } x = ty) =$$

$$= -\frac{c}{t^p} \underbrace{\int\limits_{R<|y|<2R} \frac{1}{|y|^p} \big|\varphi(y)\big|^2 \,\mathrm{d}y}_{=:b>0} = -\frac{bc}{t^p}.$$

As $p < 2$, one can choose $s > 1$ sufficiently large to have

$$\begin{aligned} \langle \varphi_t, T\varphi_t\rangle &= \int_{\mathbb{R}^d} \big(|\nabla \varphi_t|^2 + V|\varphi_t|^2\big) \,\mathrm{d}x \\ &= \frac{a}{t^2} - \frac{bc}{t^p} = \frac{a - bct^{2-p}}{t^2} < 0 \text{ for all } t \geq s. \end{aligned}$$

Now for $n \in \mathbb{N}$ put $\psi_n := \varphi_{2^n s}$, then ψ_n have mutually disjoint supports and, therefore, form an orthonormal family, and

$$\langle \psi_m, T\psi_n\rangle = 0 \text{ for } m \neq n, \quad \lambda_n := \langle \psi_n, T\psi_n\rangle < 0.$$

Let $N \in \mathbb{N}$ and consider $F := \operatorname{span}\{\psi_1, \ldots, \psi_N\}$, then $\dim F = N$. If

$$\psi \in F, \quad \psi = \sum_{n=1}^{N} \xi_n \psi_n, \quad \xi = (\xi_1, \ldots, \xi_N) \in \mathbb{C}^N,$$

then $\|\psi\|^2 = \sum_{n=1}^N |\xi_n|^2$ and

$$\langle \psi, T\psi\rangle = \sum_{m,n=1}^{N} \overline{\xi_m}\xi_n \langle \psi_m, T\psi_n\rangle = \sum_{n=1}^{N} \lambda_n |\xi_n^2| \leq \underbrace{\max\{\lambda_1, \ldots, \lambda_N\}}_{=:\mu_N<0} \sum_{n=1}^{N} |\xi_n^2| \equiv \mu_N \|\psi\|^2.$$

Therefore,

$$\Lambda_N(T) \leq \sup_{\psi\in F,\, \psi\neq 0} \frac{\langle \psi, T\psi\rangle}{\langle \psi, \psi\rangle} \leq \mu_N < 0,$$

while $N \in \mathbb{N}$ was arbitrary. □

6 Some asymptotic aspects

6.1 Definitions and preliminary observations

In the present section we proceed with a more detailed study using the mix-max principle. Let us introduce a general framework, which will cover a variety of situations.

Throughout the whole chapter, let $\Omega \subset \mathbb{R}^d$ be a non-empty open subset. Furthermore, let $V \in L^2_{\mathrm{loc}}(\Omega)$ be a real-valued potential. We consider the negative and positive parts of V,

$$V_- := \max\{-V, 0\}, \quad V_+ := \max\{V, 0\},$$

then $V = V_+ - V_-$ and $|V| = V_+ + V_-$. We will make the rather general "smallness" assumption for V_-

for any $a > 0$ there is $b > 0$ such that

$$\int_\Omega V_-|u|^2 \,\mathrm{d}x \le a \int_\Omega |\nabla u|^2 \,\mathrm{d}x + b \int_\Omega |u|^2 \,\mathrm{d}x \tag{6.1}$$

for all $u \in C_c^\infty(\Omega)$.

Remark that this assumption holds if V is semibounded from below: then $V_- \in L^\infty$ and one takes $b := \|V\|_\infty$ and any $a > 0$), which covers a large class of reasonable situation. But some unbounded V_- can be included as well: if $d = 3$, the computation of Example 1.83 show that (6.1) holds if $V(x) \ge -q/|x|$ for some $q > 0$. Many further examples are possible.

Now consider the operator

$$\widetilde{T} : C_c^\infty(\Omega) \ni u \mapsto -\Delta u + Vu \in L^2(\Omega) \tag{6.2}$$

and remark that (6.1) guarantees that $\widetilde{T}$ is semibounded from below in $L^2(\Omega)$: if one chooses $b > \frac{1}{2}$ such that

$$\int_\Omega V_-|u|^2 \,\mathrm{d}x \le \frac{1}{2} \int_\Omega |\nabla u|^2 \,\mathrm{d}x + b \int_\Omega |u|^2 \,\mathrm{d}x,$$

for all $u \in C_c^\infty(\Omega)$, then for the same u one has

$$\begin{aligned}
\langle u, \widetilde{T}u\rangle &= \int_\Omega \overline{u}\big(-\Delta u + Vu\big) \,\mathrm{d}x = \int_\Omega |\nabla u|^2 \,\mathrm{d}x + \underbrace{\int_\Omega V_+|u|^2 \,\mathrm{d}x}_{\ge 0} - \int_\Omega V_-|u|^2 \,\mathrm{d}x \\
&\ge \int_\Omega |\nabla u|^2 \,\mathrm{d}x - \Big(\frac{1}{2} \int_\Omega |\nabla u|^2 \,\mathrm{d}x + b \int_\Omega |u|^2 \,\mathrm{d}x\Big) \\
&= \frac{1}{2} \int_\Omega |\nabla u|^2 \,\mathrm{d}x - b \int_\Omega |u|^2 \,\mathrm{d}x,
\end{aligned}$$

and one obtains

$$\begin{aligned}\langle u, \widetilde{T}u\rangle + 2b\|u\|^2 &\geq \frac{1}{2}\int_\Omega |\nabla u|^2\,\mathrm{d}x + b\int_\Omega |u|^2\,\mathrm{d}x\\ &\geq \frac{1}{2}\Big(\int_\Omega |\nabla u|^2\,\mathrm{d}x + \int_\Omega |u|^2\,\mathrm{d}x\Big) \equiv \frac{1}{2}\,\|u\|^2_{H^1(\Omega)} \geq 0.\end{aligned} \tag{6.3}$$

Now we denote

$$T(\Omega, V) := \text{the Friedrichs extension in } L^2(\Omega) \text{ of the operator (6.2)}$$

and let t_V^Ω be the sesquilinear form for $T(\Omega, V)$. By construction we have

$$t_V^\Omega(u,u) \equiv \langle u, \widetilde{T}u\rangle = \int_\Omega \big(|\nabla u|^2 + V|u|^2\big)\,\mathrm{d}x \text{ for all } u \in C_c^\infty(\Omega),$$

and the inequality (6.1) extends by density to all $u \in D(t_V^\Omega)$. The inequality (6.3) shows that $D(t_V^\Omega) \subset H_0^1(\Omega)$ and, moreover, the embedding $D(t_V^\Omega) \hookrightarrow H_0^1(\Omega)$ is continuous. As the embedding $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ is compact for bounded Ω (Prop. 2.36) one concludes by Theorem 2.33 that

$$\text{if } \Omega \text{ is bounded, then } T(\Omega, V) \text{ has compact resolvent.} \tag{6.4}$$

Furthermore, if $u \in D(t_V^\Omega)$ and $h \in C^\infty(\Omega)$ such that h and ∇h are bounded, then $hu \in D(t_V^\Omega)$: if $u_n \in C_c^\infty(\Omega)$ converge to u in $D(t_V^\Omega)$, then a simple computation shows that hu_n is a Cauchy sequence in $D(t_V^\Omega)$ and hu_n converge to hu in $L^2(\Omega)$, which gives the result (see Prop. 1.73).

In fact, the most important cases for us are:

- $\Omega = \mathbb{R}^d$, then $T(\Omega, V)$ is the usual Schrödinger operator $-\Delta + V$ in $L^2(\mathbb{R}^d)$,
- $V \equiv 0$, then $T(\Omega, V)$ is the Dirichlet Laplacian in Ω,

but many further situations are possible.

Remark that the precise description of the domains of the operator $T(\Omega, V)$ and the form t_V^Ω will not be very important: as $C_c^\infty(\Omega)$ is dense in $D(t_V^\Omega)$ due to the construction of the Friedrichs extension, due to the form version of the min-max principle (Theorem 5.2) one has

$$\Lambda_n\big(T(\Omega, V)\big) = \inf_{\substack{F \subset C_c^\infty(\Omega)\\ \dim F = n}} \sup_{u \in F,\, u \neq 0} \frac{t_V^\Omega(u,u)}{\|u\|^2_{L^2(\Omega)}} \text{ for any } n \in \mathbb{N}. \tag{6.5}$$

We also remark that if the main assumption (6.1) holds for a pair (Ω, V), then it also holds for $(\widetilde{\Omega}, \lambda V)$ for any open subset $\widetilde{\Omega} \subset T$ and any $\lambda \geq 0$, which gives rise to the associated operators $T(\widetilde{\Omega}, \lambda V)$. The inclusion $C_c^\infty(\widetilde{\Omega}) \subset C_c^\infty(\Omega)$ immediately gives

$$\text{the domain monotonicity: } \widetilde{\Omega} \subset \Omega \quad \Rightarrow \quad T(\Omega, V) \leq T(\widetilde{\Omega}, V), \tag{6.6}$$

which holds for any admissible V.

6.2 Truncated operators and IMS partitions

The open set Ω and the potential V will be fixed this section. For $R > 0$ we denote

$$\begin{gathered}\Omega_R := \Omega \cap \{x \in \mathbb{R}^d : |x| < R\}, \qquad \Omega_R^{\complement} := \Omega \cap \{x \in \mathbb{R}^d : |x| > R\},\\ T := T(\Omega, V), \qquad T_R := T(\Omega_R, V), \qquad T_R^{\complement} := T(\Omega_R^{\complement}, V),\end{gathered} \tag{6.7}$$

and denote by t, t_R, $t_R^{\complement}$ the sesquilinear forms for T, T_R, $T_R^{\complement}$.

Remark that for any $R > 0$ one has $\Omega_R \subset \Omega$ and $\Omega_R^{\complement} \subset \Omega$, and the domain monotonicity shows that

$$\Lambda_n(T) \le \Lambda_n(T_R) \text{ and } \Lambda_n(T) \le \Lambda_n(T_R^{\complement}) \text{ for any } n \in \mathbb{N} \text{ and } R > 0. \tag{6.8}$$

Furthermore, if R becomes larger, then Ω_R becomes larger but $\Omega_R^{\complement}$ becomes smaller, and the domain monotonicity shows that

$$\begin{gathered}R \mapsto \Lambda_n(T_R) \text{ is non-increasing for any fixed } n \in \mathbb{N},\\ R \mapsto \Lambda_n(T_R^{\complement}) \text{ is non-decreasing for any fixed } n \in \mathbb{N},\end{gathered} \tag{6.9}$$

which holds for all R such that Ω_R resp. $\Omega_R^{\complement}$ are non-empty.

We are going to compare our operator T with the direct sum of two other operators using the following identity:

Lemma 6.1 (IMS formula[10]). *Let $u \in H^1(\Omega)$ and $\chi, \widetilde{\chi} \in C^\infty(\Omega)$ be real-valued functions such that:*

- $\chi, \widetilde{\chi}, \nabla\chi, \nabla\widetilde{\chi}$ *are bounded,*
- $\chi^2 + \widetilde{\chi}^2 = 1$,

then

$$\int_\Omega |\nabla u|^2 \,\mathrm{d}x = \int_\Omega |\nabla(\chi u)|^2 \,\mathrm{d}x + \int_\Omega |\nabla(\widetilde{\chi} u)|^2 \,\mathrm{d}x - \int_\Omega \big(|\nabla\chi|^2 + |\nabla\widetilde{\chi}|^2\big)\, |u|^2 \,\mathrm{d}x.$$

Proof. We have

$$\begin{aligned}\big|\nabla(\chi u)\big|^2 = |u\nabla\chi + \chi\nabla u|^2 &= \big|(\nabla\chi)u\big|^2 + 2\Re\big(\overline{u}\,\chi\nabla\chi\cdot\nabla u\big) + \chi^2|\nabla u|^2\\ &= \big|(\nabla\chi)u\big|^2 + \Re\Big[\overline{u}\,\nabla(\chi^2)\cdot\nabla u\Big] + \chi^2|\nabla u|^2\end{aligned}$$

and similarly

$$\big|\nabla(\widetilde{\chi} u)\big|^2 = \big|(\nabla\widetilde{\chi})u\big|^2 + \Re\Big[\overline{u}\,\nabla(\widetilde{\chi}^{\,2})\cdot\nabla u\Big] + \widetilde{\chi}^2|\nabla u|^2.$$

It follows that

$$\begin{aligned}&\big|\nabla(\chi u)\big|^2 + \big|\nabla(\widetilde{\chi} u)\big|^2\\ &\qquad= \big(|\nabla\chi|^2 + |\nabla\widetilde{\chi}|^2\big)\,|u|^2 + \Re\Big[\overline{u}\,\nabla(\underbrace{\chi^2 + \widetilde{\chi}^2}_{\equiv 1})\cdot\nabla u\Big] + \big(\underbrace{\chi^2 + \widetilde{\chi}^2}_{\equiv 1}\big)|\nabla u|^2\\ &\qquad= \big(|\nabla\chi|^2 + |\nabla\widetilde{\chi}|^2\big)\,|u|^2 + |\nabla u|^2,\end{aligned}$$

and the integration over Ω gives the result. $\square$

[10] The IMS formula is usually attributed to Ismagilov, Morgan, Simon, Sigal.

Now we make some special choice for χ and $\widetilde{\chi}$ (it will be used several times):

Remark 6.2 (IMS partition of unity). Let $\chi, \widetilde{\chi} : \mathbb{R}^d \to \mathbb{R}$ be C^∞-functions such that $\chi^2 + \widetilde{\chi}^2 = 1$ and

$$\chi(x) = 1 \text{ for } |x| \le \frac{5}{4}, \quad \widetilde{\chi}(x) = 1 \text{ for } |x| \ge \frac{7}{4}.$$

For any $R > 0$ consider the C^∞-functions $\chi_R, \widetilde{\chi}_R : \mathbb{R}^d \to \mathbb{R}$,

$$\chi_R : \ x \mapsto \chi\Big(\frac{x}{R}\Big), \qquad \widetilde{\chi}_R : \ x \mapsto \widetilde{\chi}\Big(\frac{x}{R}\Big).$$

For $R > 0$ we have $\chi_R^2 + \widetilde{\chi}_R^2 = 1$, and if $u \in C_c^\infty(\Omega)$, then

$$\begin{aligned} \chi_R u \in C_c^\infty(\Omega_{2R}), \quad \widetilde{\chi}_R u \in C_c^\infty(\Omega_R^\complement), \\ \|u\|^2_{L^2(\Omega)} = \|\chi_R u\|^2_{L^2(\Omega_{2R})} + \|\widetilde{\chi}_R u\|^2_{L^2(\Omega_R^\complement)}. \end{aligned} \tag{6.10}$$

In addition, the function $|\nabla\chi|^2 + |\nabla\widetilde{\chi}|^2$ is smooth and supported in the bounded set $\big\{x \in \mathbb{R}^d : 1 < |x| < 2\big\}$. Therefore, it is bounded, so we denote

$$B := \big\| \, |\nabla\chi|^2 + |\nabla\widetilde{\chi}|^2 \big\|_\infty < \infty.$$

For any $x \in \mathbb{R}^d$ there holds then

$$\big(|\nabla\chi_R|^2 + |\nabla\widetilde{\chi}_R|^2\big)(x) = \frac{1}{R^2}\,\big(|\nabla\chi|^2 + |\nabla\widetilde{\chi}|^2\big)\Big(\frac{x}{R}\Big) \le \frac{B}{R^2}.$$

Lemma 6.3 (IMS decoupling). *Let $\chi_R, \widetilde{\chi}_R$ and B be as in Remark 6.2, then for any $u \in D(t)$ and any $R > 0$ such that both Ω_{2R} and $\Omega_R^\complement$ are non-empty one has*

$$t(u,u) \ge t_{2R}(\chi_R u, \chi_R u) + t_R^\complement(\widetilde{\chi}_R u, \widetilde{\chi}_R u) - BR^{-2}\|u\|^2_{L^2(\Omega)}.$$

Proof. One has

$$\begin{aligned} t(u,u) &= \int_\Omega |\nabla u|^2 \,\mathrm{d}x + \int_\Omega V|u|^2 \,\mathrm{d}x \\ \text{(Lemma 6.1)} &= \int_\Omega |\nabla(\chi_R u)|^2 \,\mathrm{d}x + \int_\Omega |\nabla(\widetilde{\chi}_R u)|^2 \,\mathrm{d}x - \int_\Omega \big(|\nabla\chi_R|^2 + |\nabla\widetilde{\chi}_R|^2\big)\,|u|^2 \,\mathrm{d}x \\ &\quad + \int_\Omega V|\chi_R u|^2 \,\mathrm{d}x + \int_\Omega V|\widetilde{\chi}_R u|^2 \,\mathrm{d}x \\ &= \int_{\Omega_{2R}} \big(|\nabla(\chi_R u)|^2 + V|\chi_R u|^2\big)\,\mathrm{d}x + \int_{\Omega_R^\complement} \big(|\nabla(\widetilde{\chi}_R u)|^2 + V|\widetilde{\chi}_R u|^2\big)\,\mathrm{d}x \\ &\quad - \int_\Omega \big(|\nabla\chi_R|^2 + |\nabla\widetilde{\chi}_R|^2\big)\,|u|^2 \,\mathrm{d}x \\ &\ge t_{2R}(\chi_R u, \chi_R u) + t_R^\complement(\widetilde{\chi}_R u, \widetilde{\chi}_R u) - BR^{-2}\|u\|^2_{L^2(\Omega)}. \end{aligned}$$

□

Corollary 6.4. *For any $R > 0$ such that both Ω_{2R} and $\Omega_R^\complement$ are non-empty and any $n \in \mathbb{N}$ one has*

$$\Lambda_n(T + BR^{-2}) \ge \Lambda_n(T_{2R} \oplus T_R^\complement)$$

with $B > 0$ independent of Ω, V and R.

Proof. In view of (6.10) and Lemma 6.3 the map

$$J : C_c^\infty(\Omega) \ni u \mapsto (\chi_R u, \widetilde{\chi}_R u) \in D(t_{2R}) \times D(t_R^\complement) \equiv D(t_{2R} \oplus t_R^\complement)$$

can be viewed as an identification map: one has

$$\begin{aligned} t_{2R}(\chi_R u, \chi_R u) + t_R^\complement(\widetilde{\chi}_R u, \widetilde{\chi}_R u) &= (t_{2R} \oplus t_R^\complement)\big((\chi_R u, \widetilde{\chi}_R u), (\chi_R u, \widetilde{\chi}_R u)\big) \\ &\equiv (t_{2R} \oplus t_R^\complement)(Ju, Ju), \end{aligned}$$

and the result of Lemma 6.3 means that

$$(t_{2R} \oplus t_R^\complement)(Ju, Ju) \le t(u,u) + BR^{-2}\|u\|^2_{L^2(\Omega)} \text{ for all } u \in C_c^\infty(\Omega).$$

In the language of Definition 5.8 one has $T_{2R} \oplus T_R^\complement \le T + BR^{-2}$ using J, and the min-max principle (Corollary 5.9) shows that for any $n \in \mathbb{N}$ one has

$$\Lambda_n(T_{2R} \oplus T_R^\complement) \le \Lambda_n(T + BR^{-2}).$$

It remains to remark that the choice of χ and $\widetilde{\chi}$ (and then the value of B) are independent of Ω, V and R by construction. □

6.3 Persson theorem for the essential spectrum

As the first application we prove the following result:

Theorem 6.5 (Persson theorem for the bottom of the essential spectrum). *If Ω is unbounded, then*

$$\inf \operatorname{spec}_{\text{ess}} T = \lim_{R \to +\infty} \inf \operatorname{spec} T_R^\complement. \tag{6.11}$$

The formula (6.11) has at least two curious aspects:

- the essential spectrum appears explicitly only on the left-hand side,
- there is no "limit" of $\Omega_R^\complement$ as $R \to +\infty$ (the set $\Omega_R^\complement$ escapes to infinity).

Proof. Remark first that $\inf \operatorname{spec} T_R^\complement = \Lambda_1(T_R^\complement)$ for any $R > 0$. The monotonicity of $R \mapsto \Lambda_1(T_R^\complement)$, see (6.9), shows that the limit

$$\Lambda := \lim_{R \to +\infty} \inf \operatorname{spec} T_R^\complement \equiv \lim_{R \to +\infty} \Lambda_1(T_R^\complement)$$

always exists (can be equal to $+\infty$).

(a) Let us show first that $\inf \operatorname{spec}_{\text{ess}} T \le \Lambda$. If $\Lambda = +\infty$, then the inequality holds. Now let $\Lambda < \infty$. Our strategy is as follows: we take $\varepsilon > 0$ and show that

$$\Lambda_n(T) \le \Lambda + \varepsilon \text{ for any } n \in \mathbb{N}. \tag{6.12}$$

Then $\inf \operatorname{spec}_{\text{ess}} T = \lim_{n \to \infty} \Lambda_n(T) \le \Lambda + \varepsilon$ for any $\varepsilon > 0$, which gives the result.

It remains to show (6.12). As $R \mapsto \Lambda_1(T_R^{\complement})$ is non-decreasing, one has $\Lambda_1(T_R^{\complement}) \le \Lambda$ for any $R > 0$. The definition of Λ_1 shows that for any $R > 0$ we have

$$\Lambda_1(T_R^{\complement}) = \inf_{u \in C_c^\infty(\Omega_R^{\complement}),\, u \ne 0} \frac{t_R^{\complement}(u,u)}{\|u\|^2_{L^2(\Omega_R^{\complement})}} \equiv \inf_{u \in C_c^\infty(\Omega_R^{\complement}),\, u \ne 0} \frac{t(u,u)}{\|u\|^2_{L^2(\Omega)}} \le \Lambda.$$

Pick any $R_1 > 0$, then one can find $u_1 \in C_c^\infty(\Omega_{R_1}^{\complement})$ with $\|u_1\|^2_{L^2(\Omega)} = 1$ such that $t(u_1, u_1) < \Lambda + \varepsilon$. Let $R_2 > 0$ with $\operatorname{supp} u_1 \subset B_{R_2}(0)$, then one can find a function $u_2 \in C_c^\infty(\Omega_{R_2}^{\complement})$ with $\|u_2\|^2_{L^2(\Omega)} = 1$ such that $t(u_2, u_2) \le \Lambda + \varepsilon$. Take $R_3 > 0$ with $\operatorname{supp} u_2 \in B_{R_3}(0)$ and continue similarly. We obtain an infinite sequence of functions $u_j \in C_c^\infty(\Omega)$ with mutually disjoint supports such that $\|u_j\|_{L^2(\Omega)=1}$ and $t(u_j, u_j) < \Lambda + \varepsilon$ for all $j \in \mathbb{N}$. In particular, $t(u_j, u_k) = 0$ and $\langle u_j, u_k \rangle_{L^2(\Omega)} = 0$ for all $j \ne k$.

Let $n \in \mathbb{N}$ and $F := \operatorname{span}\{u_1, \dots, u_n\}$, then $F \subset C_c^\infty(\Omega)$ with $\dim F = n$. If $u \in F$, then one has the unique representation $u_n = \xi_1 u_1 + \dots + \xi_n u_n$ with some $\xi_j \in \mathbb{C}$, and then

$$\|u\|^2 = \sum_{j=1}^n |\xi_j|^2, \qquad t(u,u) = \sum_{j,k=1} \overline{\xi_j}\, \xi_k\, t(u_j, u_k)$$
$$= \sum_{j=1}^n |\xi_j|^2 t(u_j, u_j) \le (\Lambda + \varepsilon) \sum_{j=1}^n |\xi_j|^2 = (\Lambda + \varepsilon)\|u\|^2,$$

which gives (6.12) and concludes the part (a).

(b) Let us show the reverse inequality $\inf \operatorname{spec}_{\mathrm{ess}} T \ge \Lambda$. By Corollary 6.4 we have $\Lambda_n(T) + BR^{-2} \ge \Lambda_n(T_{2R} \oplus T_R^{\complement})$, and for $n \to \infty$ we obtain

$$\inf \operatorname{spec}_{\mathrm{ess}} T + BR^{-2} \equiv \inf \operatorname{spec}_{\mathrm{ess}}(T + BR^{-2}) \ge \inf \operatorname{spec}_{\mathrm{ess}}(T_{2R} \oplus T_R^{\complement}).$$

The operator T_{2R} has compact resolvent (as Ω_{2R} is bounded), so $\operatorname{spec}_{\mathrm{ess}} T_{2R} = \emptyset$ and (Remark 5.12)

$$\inf \operatorname{spec}_{\mathrm{ess}}(T_{2R} \oplus T_R^{\complement}) \equiv \inf(\underbrace{\operatorname{spec}_{\mathrm{ess}} T_{2R}}_{=\emptyset} \cup \operatorname{spec}_{\mathrm{ess}} T_R^{\complement}) = \inf \operatorname{spec}_{\mathrm{ess}} T_R^{\complement} \ge \inf \operatorname{spec} T_R^{\complement},$$

so we arrive at $\inf \operatorname{spec}_{\mathrm{ess}} T + BR^{-2} \ge \inf \operatorname{spec} T_R^{\complement}$, and one arrives at the sought conclusion by taking the limit $R \to \infty$. □

Corollary 6.6 (Eigenvalues in truncated domains). *Let Ω be unbounded and $N \in \mathbb{N}$.*

(a) if $E_N(T_R) < \inf \operatorname{spec}_{\mathrm{ess}} T$ for some $R > 0$, then T has at least N eigenvalues below $\inf \operatorname{spec}_{\mathrm{ess}} T$.

(b) if T has at least N eigenvalues in $(-\infty, \inf \operatorname{spec}_{\mathrm{ess}} T)$, then

$$E_N(T_R) = E_N(T) + O\Big(\frac{1}{R^2}\Big) \text{ as } R \to +\infty,$$

in other words, the eigenvalues of the "finite part" T_R of T below $\inf \operatorname{spec}_{\mathrm{ess}} T$ approximate the respective eigenvalues of T as $R \to \infty$.

Proof. (a) Use the domain monotonicity:

$$\Lambda_N(T) \leq \Lambda_N(T_R) \equiv E_N(T_R) < \inf \operatorname{spec}_{\mathrm{ess}} T,$$

and this shows that $\Lambda_N(T) = E_N(T)$.

(b) Recall that for any $n \in \mathbb{N}$ we have $E_n(T_R) = \Lambda_n(T_R)$ and

$$\Lambda_n(T) \leq \Lambda_n(T_{2R}), \quad \Lambda_n(T_{2R} \oplus T_R^{\complement}) \leq \Lambda_n(T) + BR^{-2}. \tag{6.13}$$

By assumption one has $\Lambda_N(T) = E_N(T) \leq \inf \operatorname{spec}_{\mathrm{ess}} T$.

Take any E with $\Lambda_N(T) < E < \inf \operatorname{spec}_{\mathrm{ess}} T$. Due to Persson theorem, for all sufficiently large R one has $\Lambda_N(T)+BR^{-2} < E < \Lambda_1(T_R^{\complement})$, and the second inequality in (6.13) gives $\Lambda_N(T_{2R} \oplus T_R^{\complement}) < E$.

We claim that $\Lambda_N(T_{2R}) < E$. In fact, if one assumes that $\Lambda_N(T_{2R}) \geq E$, then using $T_R^{\complement} > E$ one obtains that the spectrum of $T_{2R} \oplus T_R^{\complement}$ in $(-\infty, E)$ consists of at most $N-1$ eigenvalues, which then shows $\Lambda_N(T_{2R} \oplus T_R^{\complement}) \geq E$ and gives a contradiction.

Due to $\Lambda_N(T_{2R}) < E$ and $T_R^{\complement} > E$ we obtain $\Lambda_N(T_{2R} \oplus T_R^{\complement}) = \Lambda_N(T_{2R})$, and using the both inequalities (6.13) one arrives at

$$\Lambda_N(T_{2R}) - BR^{-2} \leq \Lambda_N(T) \leq \Lambda_N(T_{2R})$$

for all sufficiently large R, and one recalls that both Λ_N are actually E_N. $\square$

6.4 Strong coupling asymptotics

We will continue to work with the truncated domains Ω_R and $\Omega_R^{\complement}$ defined in (6.7) but prefer to use the "full notation" $T(\Omega, V)$ as both Ω and V will be varying.

We are interested in the behavior of the spectrum of $T(\Omega, \lambda V)$ when Ω and V are fixed but $\lambda \to +\infty$. The constant λ (measuring the "strength" of the potential term) is usually referred to as the coupling constant, and the case $\lambda \to +\infty$ is referred to as the strong coupling.

Theorem 6.7 (Strong coupling asymptotics at first order). *Assume that V is semibounded from below and denote $V_{\min} := \operatorname{ess\,inf} V$. Then for any fixed $n \in \mathbb{N}$ there holds*

$$\Lambda_n\big(T(\Omega, \lambda V)\big) = V_{\min}\lambda + o(\lambda) \text{ as } \lambda \to +\infty.$$

Proof. During the proof set $T_\lambda := T(\Omega, \lambda V)$. Without loss of generality assume that $V_{\min} = 0$. Then $V \geq 0$ a.e. and $T_\lambda \geq 0$ for all $\lambda > 0$. This gives $\Lambda_n(T_\lambda) \geq 0$ for all $n \in \mathbb{N}$ and $\lambda > 0$.

For the upper bound remark that the spectrum of M_V (the operator of multiplication by V) is purely essential, so $0 = \operatorname{ess\,inf} V = \inf \operatorname{spec}_{\mathrm{ess}} M_V$ and $\Lambda_n(M_V) = 0$ by the min-max principle. One easily shows that it is essentially self-adjoint on $C_c^\infty(\Omega)$, and for any $n \in \mathbb{N}$ one can find an n-dimensional subspace U of $C_c^\infty(\Omega)$ such that

$$\langle \varphi, M_V \varphi \rangle = \int_\Omega V|\varphi|^2 \,\mathrm{d}x \leq \varepsilon \|\varphi\|^2_{L^2(\Omega)} \text{ for all } \varphi \in U.$$

As U is finite-dimensional, there exists $C > 0$ such that

$$\frac{\displaystyle\int_\Omega |\nabla\varphi|^2 \,\mathrm{d}x}{\|\varphi\|^2_{L^2(\Omega)}} \leq C \text{ for all } \varphi \in U \setminus \{0\}.$$

Then

$$\Lambda_n(T_\lambda) \leq \sup_{\varphi\in U,\, \varphi\neq 0} \frac{\displaystyle\int_\Omega |\nabla\varphi|^2 \,\mathrm{d}x + \lambda \int_\Omega V|\varphi|^2 \,\mathrm{d}x}{\|\varphi\|^2_{L^2(\Omega)}} \leq C + 2\varepsilon\lambda.$$

As $\varepsilon > 0$ is arbitrary, this shows that $\Lambda_n(T_\lambda) = o(\lambda)$ for large λ. □

Remark 6.8. If Ω is bounded, then in Theorem 6.7 one can replace Λ_n by E_n, as $T(\Omega, V)$ has compact resolvent. For unbounded Ω, it may happen that all Λ_n are the same, and no eigenvalues are present (e.g. for $\Omega = \mathbb{R}^d$ and $V = 0$). Nevertheless, one can guarantee the existence of eigenvalues by an additional assumption on V.

Corollary 6.9 (Existence of eigenvalues in the strong coupling regime). *Assume that Ω is unbounded and that V is semibounded from below and denote*

$$V_{\min} := \operatorname{ess\,inf} V, \qquad V_\infty := \liminf_{|x|\to+\infty} V(x).$$

If $V_{\min} < V_\infty$, then for any $n \in \mathbb{N}$:

- *there exists $\lambda_n > 0$ such that $T(\Omega, \lambda V)$ has at least n eigenvalues below the bottom of the essential spectrum for all $\lambda > \lambda_n$,*
- *there holds $E_n\big(T(\Omega, \lambda V)\big) = V_{\min}\,\lambda + o(\lambda)$ as $\lambda \to +\infty$.*

Proof. In view of Theorem 6.7 one simply needs to show that Λ_n is the nth eigenvalue, i.e. that it lies strictly below the essential spectrum (Corollary 5.4).

Take any c with $V_{\min} < c < V_\infty$. By assumption there is $R > 0$ such that $V(x) \geq c$ for all $x \in \Omega_R^\complement$, then $T(\Omega_r^\complement, \lambda V) \geq c\lambda$ for all $r \geq R$, and Persson theorem shows that

$$\inf \operatorname{spec}_{\mathrm{ess}} T(\Omega, \lambda V) = \lim_{r\to\infty} T(\Omega_r^\complement, \lambda V) \geq c\lambda.$$

For large λ one has $\Lambda_n\big(T(\Omega, \lambda V)\big) = V_{\min}\lambda + o(\lambda) < c\lambda \leq \inf \operatorname{spec}_{\mathrm{ess}} T(\Omega, \lambda V)$, which gives the result. $\square$

We are now interested in more precise asymptotic expansions for the eigenvalues $E_n\big(T(\Omega, \lambda V)\big)$ for large λ. This problem has no general solution: in fact, the asymptotics depend on the way how V attains its minimum: it can be reached e.g. at a single point, or on a submanifold, or on an open set, and the respective eigenvalue asymptotics are different. We only consider the "generic" case when the minimum is attained at a single point.

Theorem 6.10 (Detailed strong coupling asymptotics). *Assume that:*

- *$0 \in \Omega$ is the unique global minimum of V on Ω,*
- *for any $r > 0$ there holds $\operatorname{ess\,inf}_{\Omega_r^\complement} V > V(0)$*
  (in other words, V does not approach the value $V(0)$ at other places),
- *V is C^3-smooth near 0 and its Hessian matrix $V''(0)$ in 0 is non-degenerate.*

Denote

$$\begin{aligned}\mu_1, \dots, \mu_d &:= \text{the eigenvalues of } V''(0),\\ \mathbb{E} &:= \text{ the disjoint union } \bigsqcup_{(n_1,\dots,n_d)\in\mathbb{N}^d} \Big\{ \sum_{j=1}^d (2n_j - 1)\sqrt{\frac{\mu_j}{2}} \Big\},\\ \varepsilon_n &:= \text{ the } n\text{-th element of } \mathbb{E}.\end{aligned}$$

Let $n \in \mathbb{N}$, then for $\lambda \to +\infty$ the operator $T(\Omega, \lambda V)$ has the n-th eigenvalue, and

$$E_n\big(T(\Omega, \lambda V)\big) = V(0)\lambda + \varepsilon_n\sqrt{\lambda} + O(\lambda^{\frac{2}{5}}).$$

As a preparation for the proof consider an explicit example.

Example 6.11 (Multidimensional harmonic oscillator). Let A_0 be a positive definite $d \times d$ real matrix with eigenvalues $\alpha_1, \dots, \alpha_d > 0$. Consider the potential $V_0 : \mathbb{R}^d \ni x \mapsto x \cdot A_0 x$ and the Schrödinger operator $H_\lambda := -\Delta + \lambda V_0$ in $L^2(\mathbb{R}^d)$

with $\lambda > 0$. Let us show that $\operatorname{spec} H_\lambda$ can be computed explicitly (remark that V_0 is in the class covered by Theorem 6.10).

There exists an orthogonal matrix θ with $\theta^{-1}A_0\theta = \operatorname{diag}(\alpha_1, \dots, \alpha_d) =: A_1$, and for any $x \in \mathbb{R}^d$ one has

$$V_0(\theta x) = \theta x \cdot A_0 \theta x = x \cdot (\theta^{-1} A_0 \theta) x = x \cdot A_1 x \equiv \sum_{j=1}^{d} \alpha_j x_j^2 =: V_1(x).$$

Consider the unitary transform $\Theta : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ defined by $(\Theta u)(x) = u(\theta x)$ and the Schrödinger operator $G_\lambda := -\Delta + \lambda V_1 = -\Delta + \lambda\alpha_1 x_1^2 + \dots + \lambda\alpha_d x_d^2$ in $L^2(\mathbb{R}^d)$. For any $u \in C_c^\infty(\mathbb{R}^d)$ one has

$$\begin{aligned}(G_\lambda \Theta u)(x) &= -\sum_{k=1}^{d} \partial_{x_k}^2 u(\theta x) + \lambda V_1(x) u(\theta x)\\ &= -\sum_{i,j,k=1}^{d} \theta_{ik}\theta_{jk}\partial_{ij}^2 u(\theta x) + \lambda V_0(\theta x) u(\theta x).\end{aligned}$$

We have $\sum_{k=1}^d \theta_{ik}\theta_{jk} = (\theta\theta^t)_{ij} = \delta_{ij}$, which gives

$$(G_\lambda \Theta u)(x) = -\Delta u(\theta x) + V_0(\theta x) u(\theta x) = (\Theta H_\lambda u)(x),$$

i.e. $\Theta^{-1} G_\lambda \Theta = H_\lambda$ on $C_c^\infty(\mathbb{R}^d)$. As both G_λ and H_λ are essentially self-adjoint on $C_c^\infty(\mathbb{R}^d)$, this extends to the whole domain and shows that $\Theta^{-1} G_\lambda \Theta = H_\lambda$, so G_λ and H_λ are unitarily equivalent and have the same eigenvalues.

We know that the eigenvalues of the one-dimensional harmonic oscillator

$$T = -\frac{d}{dx^2} + \omega^2 x^2, \quad \omega > 0,$$

are $(2n-1)\omega$ with $n \in \mathbb{N}$ and the respective normalized eigenfunctions $\psi_{n,\omega}$ form an orthonormal basis in $L^2(\mathbb{R})$ (for any fixed $\omega > 0$). Then the functions

$$\Psi_{(n_1,\dots,n_d)} : (x_1, \dots, x_d) \mapsto \psi_{n_1,\sqrt{\lambda\alpha_1}}(x_1) \cdot \ldots \cdot \psi_{n_d,\sqrt{\lambda\alpha_d}}(x_d), \quad (n_1, \dots, n_d) \in \mathbb{N}^d,$$

form an orthonormal basis in $L^2(\mathbb{R}^d)$, see Lemma 2.29, and

$$G_\lambda \Psi_{(n_1,\dots,n_d)} = \Big(\underbrace{(2n_1 - 1)\sqrt{\lambda\alpha_1} + \ldots + (2n_d - 1)\sqrt{\lambda\alpha_d}}_{=\alpha(n_1,\dots,n_d)} \Big) \Psi_{(n_1,\dots,n_d)},$$

i.e. each $\Psi_{(n_1,\dots,n_d)}$ is an eigenfunction of G_λ with eigenvalue $\alpha(n_1, \dots, n_d)$, and these eigenvalues exhaust the whole spectrum of G_λ and of the unitary equivalent H_λ. So we note that the spectrum of H_λ consists of the eigenvalues

$$\Big((2n_1 - 1)\sqrt{\alpha_1} + \ldots + (2n_d - 1)\sqrt{\alpha_d} \Big) \sqrt{\lambda}, \quad (n_1, \dots, n_d) \in \mathbb{N}^d.$$

Proof of Theorem 6.10. Without loss of generality assume that $V(0) = 0$. Consider the matrix $A_0 := \frac{1}{2}V''(0)$, whose eigenvalues are $\frac{\mu_j}{2}$ and the potential $V_0 : x \mapsto x \cdot A_0 x$ (=the quadratic approximation of V near 0). The operator $T(\mathbb{R}^d, \lambda V_0)$ is exactly the harmonic oscillator H_λ from Example 6.11 and $\varepsilon_n\sqrt{\lambda}$ is its n-th eigenvalue (for any $\lambda > 0$). We are reduced to prove that for any $n \in \mathbb{N}$ there holds

$$\Lambda_n\big(T(\Omega, \lambda V)\big) = \Lambda_n\big(T(\mathbb{R}^d, \lambda V_0)\big) + O(\lambda^{\frac{2}{5}}) \text{ as } \lambda \to +\infty. \tag{6.14}$$

Let $s > 0$ (to be chosen later) and $n \in \mathbb{N}$. Using the domain monotonicity and the IMS constructions (Corollary 6.4 with $R := \lambda^{-s}$) we obtain with some $B > 0$:

$$\begin{aligned}\Lambda_n\Big(T(\Omega_{2\lambda^{-s}}, \lambda V) \oplus T(\Omega^{\complement}_{\lambda^{-s}}, \lambda V)\Big) - B\lambda^{2s}&\\ \le \Lambda_n\big(T(\Omega, \lambda V)\big) \le \Lambda_n\big(T(\Omega_{2\lambda^{-s}}, \lambda V)\big),&\end{aligned} \tag{6.15}$$

Now let us fix $n \in \mathbb{N}$ and look at $\Lambda_n\big(T(\Omega_{2\lambda^{-s}}, V)\big)$. Using the Taylor expansion of V one finds $c > 0$ such that $\big|V(x) - V_0(x)\big| \le c|x|^3$ as $x \to 0$. It follows that for sufficiently large $\lambda > 0$ one has

$$\big|V(x) - V_0(x)\big| \le 8c\lambda^{-3s} \text{ for all } x \in \Omega_{2\lambda^{-s}}.$$

Let M_{V-V_0} be the operator of multiplication by $V - V_0$ in $L^2(\Omega_{2\lambda^{-s}})$, then

$$\|M_{V-V_0}\| \le 8c\lambda^{-3s}, \qquad T(\Omega_{2\lambda^{-s}}, \lambda V) = T(\Omega_{2\lambda^{-s}}, \lambda V_0) + \lambda M_{V-V_0},$$

and the min-max principle for perturbations (Corollary 5.5) gives

$$\Lambda_n\big(T(\Omega_{2\lambda^{-s}}, \lambda V)\big) = \Lambda_n\big(T(\Omega_{2\lambda^{-s}}, \lambda V_0)\big) + O(\lambda^{1-3s}). \tag{6.16}$$

Now we apply the IMS estimates for V_0 and $\mathbb{R}^d$. Denote

$$B_r := \big\{x \in \mathbb{R}^d : |x| < r\big\} \equiv (\mathbb{R}^d)_r, \quad B^{\complement}_r := \big\{x \in \mathbb{R}^d : |x| > r\big\} \equiv (\mathbb{R}^d)^{\complement}_r.$$

then (Corollary 6.4 with $R := \lambda^{-s}$)

$$\begin{aligned}\Lambda_n\Big(T(B_{2\lambda^{-s}}, \lambda V_0) \oplus T(B^{\complement}_{\lambda^{-s}}, \lambda V_0)\Big) - B\lambda^{2s}&\\ \le \Lambda_n\big(T(\mathbb{R}^d, \lambda V_0)\big) \le \Lambda_n\big(T(B_{2\lambda^{-s}}, \lambda V_0)\big).&\end{aligned} \tag{6.17}$$

One has $\Lambda_n\big(T(\mathbb{R}^d, \lambda V_0)\big) = \varepsilon_n\lambda^{\frac{1}{2}} = O(\lambda^{\frac{1}{2}})$. Remark that $V_0(x) \ge c_0|x|^2$ for all $x \in \mathbb{R}^d$ (with some fixed $c_0 > 0$), and it follows that for $x \in B^{\complement}_{\lambda^{-s}}$ one has $V_0(x) \ge c_0\lambda^{-2s}$ and then $T(B^{\complement}_{\lambda^{-s}}, \lambda V_0) \ge c_0\lambda^{1-2s}$.

From now on assume that $s < \frac{1}{4}$, then $\frac{1}{2} < 1 - 2s$ and for all sufficiently large λ one has $\Lambda_n\big(T(\mathbb{R}^d, \lambda V_0)\big) \le \inf\operatorname{spec} T(B^{\complement}_{\lambda^{-s}}, \lambda V_0)$. Then (Remark 5.12)

$$\Lambda_n\Big(T(B_{2\lambda^{-s}}, \lambda V_0) \oplus T(B^{\complement}_{\lambda^{-s}}, \lambda V_0)\Big) = \Lambda_n\big(T(B_{2\lambda^{-s}}, \lambda V_0)\big).$$

For large λ one has $B_{2\lambda^{-s}} = \Omega_{2\lambda^{-s}}$, and then the estimate (6.17) shows that

$$\Lambda_n\Big(T(\Omega_{2\lambda^{-s}}, \lambda V_0)\Big) \equiv \Lambda_n\big(T(B_{2\lambda^{-s}}, \lambda V_0)\big) = \Lambda_n\big(T(\mathbb{R}^d, \lambda V_0)\big) + O(\lambda^{2s}).$$

The substitution into (6.16) gives

$$\Lambda_n\big(T(\Omega_{2\lambda^{-s}},\lambda V)\big)=\Lambda_n\big(T(B_{2\lambda^{-s}},\lambda V_0)\big)=\underbrace{\Lambda_n\big(T(\mathbb{R}^d,\lambda V_0)\big)}_{\varepsilon_n\lambda^{\frac12}}+O(\lambda^{1-3s}+\lambda^{2s}).\quad(6.18)$$

Assume additionally that $s>\frac16$, then $1-3s<\frac12$ and

$$\Lambda_n\big(T(\Omega_{2\lambda^{-s}},\lambda V)\big)=\varepsilon_n\lambda^{\frac12}+O(\lambda^{1-3s}+\lambda^{2s})=O(\lambda^{\frac12}).$$

Using the Taylor expansion of V we find $r>0$ and $c_1>0$ such that $V(x)\ge c_1|x|^2$ for all $x\in\Omega_r$. By assumption one can find $a>0$ such that $V(x)\ge a$ for a.e. $x\in\Omega_r^\complement$. For large λ one has

$$\operatorname*{ess\,inf}_{\Omega^\complement_{\lambda^{-s}}} V=\min\Big\{\operatorname*{ess\,inf}_{\Omega^\complement_{\lambda^{-s}}\cap\Omega_r} V,\operatorname*{ess\,inf}_{\Omega_r^\complement} V\Big\},$$

$$V(x)\ge c_1|x|^2\ge c_1\lambda^{-2s}\text{ for }x\in\Omega^\complement_{\lambda^{-s}}\cap\Omega_r,\quad V(x)\ge a>0\text{ for }x\in\Omega_r^\complement,$$

which gives $\operatorname{ess\,inf}_{\Omega^\complement_{\lambda^{-s}}} V\ge c_1\lambda^{-2s}$ and (using again $s<\frac14$)

$$\inf\operatorname{spec}T(\Omega^\complement_{\lambda^{-s}},\lambda V)\ge\lambda\operatorname*{ess\,inf}_{\Omega^\complement_{\lambda^{-s}}} V\ge c_1\lambda^{1-2s}>\Lambda_n\big(T(\Omega_{2\lambda^{-s}},\lambda V)\big)=O(\lambda^{\frac12})$$

and then $\Lambda_n\big(T(\Omega_{2\lambda^{-s}},\lambda V)\oplus T(\Omega^\complement_{\lambda^{-s}},\lambda V)\big)=\Lambda_n\big(T(\Omega_{2\lambda^{-s}},\lambda V)\big)$. Using this equality in (6.15) one obtains

$$\Lambda_n\big(T(\Omega,\lambda V)\big)=\Lambda_n\big(T(\Omega_{2\lambda^{-s}},\lambda V)\big)+O(\lambda^{2s})\overset{(6.18)}{=}\Lambda_n\big(T(\mathbb{R}^d,\lambda V_0)\big)+O(\lambda^{1-3s}+\lambda^{2s}).$$

Recall that this estimate holds with arbitrary $s\in(\frac16,\frac14)$. We optimize the remainder by taking $s=\frac15$ and arrive at the sought estimate (6.14). □

Remark 6.12 (Semiclassical asymptotics). In the quantum mechanics one often considers the Schrödinger operators in $L^2(\mathbb{R}^d)$ of the form $-h^2\Delta+V$ with $h\to0^+$. This asymptotic regime is usually referred to as the *semiclassical asymptotics*. This case is equivalent to the strong coupling: if one denotes $h:=\lambda^{-\frac12}$, then

$$E_n(-\Delta+\lambda V)=\lambda E_n(-h^2\Delta+V),$$

and under the assumptions of Theorem 6.10, for $h\to0^+$ one obtains

$$E_n(-h^2\Delta+V)=V(0)+\varepsilon_n h+O(h^{\frac65}).$$

For $d=1$ one has

$$\varepsilon_n=(2n-1)\sqrt{\frac{V''(0)}{2}},\quad E_n(-h^2\Delta+V)=V(0)+(2n-1)\sqrt{\frac{V''(0)}{2}}h+O(h^{\frac65}),$$

and the last formula is often referred to the *WKB*[11] *asymptotics* for the eigenvalues. We remark that the remainders in the above asymptotics can be improved with the help of different approaches.

[11]WKB= Wentzel, Kramers, Brillouin