Harmonic Analysis Techniques for Elliptic Operators

27th International Internet Seminar 2023/24

M. Egert, R. Haller, S. Monniaux, P. Tolksdorf

February 7, 2024

Contents

1. Basics on operator theory

Many linear operators, differential operators in particular, are bounded operators between a pair of Banach spaces. For example, the derivation $L := \frac{\mathrm{d}}{\mathrm{d}x}$ maps $\mathrm{C}^1([0,1])$ into $\mathrm{C}([0,1])$. This point of view suffices if we just want to apply the operator. But what if we want to define $L^2 + L$, solve an eigenvalue problem $Lu = \lambda u$ or even the abstract evolution equation $u'(t) + Lu(t) = 0$? All this requires one common ambient space to work in. The way out is a simple change of perspective, leading to unbounded operators. In the example, L would be an unbounded operator in $X := \mathrm{C}([0,1])$ that is only defined on $\mathrm{dom}(L) := \mathrm{C}^1([0,1]) \subseteq X$ called domain, on which the usual norm can now be written as $\|\cdot\|_{\mathrm{C}^1([0,1])} = \|\cdot\|_X + \|L\cdot\|_X$.

In this first lecture, we introduce the necessary tools and notions about unbounded operators in Banach spaces. We will quickly specialize to Hilbert spaces, which will be our main playground for most of the time.

Notation 1.1. All spaces in this lecture series will be over the complex numbers. Throughout this lecture, X and Y will denote Banach spaces and H and K will denote Hilbert spaces. In order to focus on the essential quantities in estimates, we will occasionally write $A \lesssim B$ instead of $A \leq CB$ for an insignificant constant $C > 0$, the dependence of which is clear from the context. Similarly, we will use the symbols $\gtrsim$ and $\simeq$.

1.1. Closed operators

We start with a generalization of the well-known concept of a linear operator.

Definition 1.2. *(a) Every linear subspace $R \subseteq X \times Y$ is called a* linear relation *between X and Y. The subspaces*

$$\begin{aligned} \mathrm{dom}(R) &:= \{u \in X \mid \exists v \in Y : (u,v) \in R\} \subseteq X, \\ \ker(R) &:= \{u \in X \mid (u,0) \in R\} \subseteq X \quad \text{and} \\ \mathrm{ran}(R) &:= \{v \in Y \mid \exists u \in X : (u,v) \in R\} \subseteq Y \end{aligned}$$

are the domain, kernel, *and* range *of R, respectively.*

(b) If R and S are linear relations between X and Y with $R \subseteq S$, then S is called an extension *of R and R is a* restriction *of S.*

(c) *If R is a linear relation between X and Y, we denote by*

$$R^{-1} := \{(v,u) \in Y \times X \mid (u,v) \in R\}$$

the inverse relation *of R.*

(d) *A linear relation R between X and Y is called a* linear operator *from X to Y if* $\ker(R^{-1}) = \{v \in Y \mid (0,v) \in R\} = \{0\}$. *If R is a linear operator from X to X itself, we will say that R is a linear operator in X.*

Remark 1.3. (a) We have $\mathrm{dom}(R^{-1}) = \mathrm{ran}(R)$ and $\mathrm{ran}(R^{-1}) = \mathrm{dom}(R)$.

(b) If R is a linear operator, then for every $u \in \mathrm{dom}(R)$ there is exactly one $v \in Y$ with $(u,v) \in R$. Indeed, if $(u,v),(u,w) \in R$, then, since R is a linear subspace, also $(u,v)-(u,w) = (0, v-w) \in R$ and Definition 1.2 (d) forces $v = w$. Thus, a linear operator R is indeed the graph of a function from $\mathrm{dom}(R)$ to Y and identifying graph and function, for $(u,v) \in R$ we write $v = Ru$, as usual. Please note that $\mathrm{dom}(R) \subsetneq X$ in general.

(c) If R and S are linear operators, then $R \subseteq S$ just means that $\mathrm{dom}(R) \subseteq \mathrm{dom}(S)$ and $Ru = Su$ for all $u \in \mathrm{dom}(R)$.

Definition 1.4. *Let R be a linear relation between X and Y.*

(a) *R is called* densely defined *if* $\mathrm{dom}(R)$ *is dense in X.*

(b) *R is* closed *if R is closed as a subspace of* $X \times Y$.

As usual, we will write $\overline{R}$ for the closure of R in $X \times Y$.

Remark 1.5. Inverting a linear relation just means to flip the entries, which is a homeomorphism between $X \times Y$ and $Y \times X$. So, R is closed if and only if R^{-1} is closed.

Definition 1.6. *Let L be a linear operator from X to Y.*

(a) *The operator L is* bounded *if* $\mathrm{dom}(L) = X$ *and* $\|Lu\|_Y \lesssim \|u\|_X$ *for all* $u \in X$.

We denote by $\mathcal{L}(X,Y)$ *the vector space of all bounded linear operators from X to Y and abbreviate* $\mathcal{L}(X) := \mathcal{L}(X,X)$.

For $L \in \mathcal{L}(X,Y)$ *the* operator norm *is given by*

$$\|L\|_{\mathcal{L}(X,Y)} := \sup_{u \in X, \|u\|_X = 1} \|Lu\|_Y.$$

(b) *The* graph norm *on* $\mathrm{dom}(L)$ *is given by*

$$\|u\|_L := \|u\|_X + \|Lu\|_Y \qquad (u \in \mathrm{dom}(L)).$$

(c) *If L is closed, then a subspace of* $\mathrm{dom}(L)$ *is a* core *for L if it is dense in* $\mathrm{dom}(L)$ *with respect to the graph norm.*

It is easy to see that the graph norm is indeed a norm on $\operatorname{dom}(L)$ that makes both the inclusion $\operatorname{dom}(L) \hookrightarrow X$ and the operator $L \colon \operatorname{dom}(L) \to Y$ bounded linear operators.

If you have seen closed operators before, you might recall a different definition involving sequences. At this point you should convince yourself that for linear operators this is in fact the same thing, as stated in the following lemma. In particular, bounded operators are closed.

Lemma 1.7. *Let L be a linear operator from X to Y. Then the following assertions are equivalent:*

(a) *L is closed.*

(b) *Whenever (u_j) is a sequence in $\operatorname{dom}(L)$ converging in X to some $u \in X$, such that (Lu_j) converges in Y to some $v \in Y$, then $u \in \operatorname{dom}(L)$ and $Lu = v$.*

(c) *$(\operatorname{dom}(L), \|\cdot\|_L)$ is complete.*

Example 1.8 (Multiplication operators in L^2)**.** A particularly important example for linear operators are multiplication operators in $\mathrm{L}^2(\mathbb{R}^n)$. For a measurable function $m \colon \mathbb{R}^n \to \mathbb{C}$, we set

$$\operatorname{dom}(M_m) := \big\{u \in \mathrm{L}^2(\mathbb{R}^n) \mid mu \in \mathrm{L}^2(\mathbb{R}^n)\big\} \quad \text{and} \quad M_m u := mu.$$

For us, 'measurable' will be synonymous with 'Lebesgue measurable' and we write $|E|$ for the Lebesgue measure of $E \subseteq \mathbb{R}^n$ if the context is clear. Let us show that M_m is densely defined and closed.

In order to show that M_m is densely defined, let $v \in \operatorname{dom}(M_m)^\perp$. Since $(1+|m|^2)^{-1}$ and $m(1+|m|^2)^{-1}$ are bounded functions and $v \in \mathrm{L}^2(\mathbb{R}^n)$, we conclude $\frac{1}{1+|m|^2} v \in \operatorname{dom}(M_m)$. This implies

$$0 = \Big\langle v, \frac{v}{1+|m|^2} \Big\rangle_{\mathrm{L}^2(\mathbb{R}^n)} = \int_{\mathbb{R}^n} \frac{|v|^2}{1+|m|^2} \, \mathrm{d}x,$$

so $v = 0$ almost everywhere, and we are done.

For the proof that M_m is closed, let (u_j) be a convergent sequence in $\mathrm{L}^2(\mathbb{R}^n)$ with limit u, such that $(M_m u_j) = (mu_j)$ converges in $\mathrm{L}^2(\mathbb{R}^n)$ to some function v. Then there is a subsequence $(u_{j_k})_{k \in \mathbb{N}}$ of (u_j) that converges pointwise almost everywhere. Thus, the sequence $(mu_{j_k})_{k \in \mathbb{N}}$ also converges in the same sense towards mu. Since by hypothesis the same sequence converges to v in the L^2-sense, we get $mu = v \in \mathrm{L}^2(\mathbb{R}^n)$. This gives $u \in \operatorname{dom}(M_m)$ and $M_m u = mu = v$. Hence, M_m is closed.

It is a natural question, for which functions m the operator M_m is bounded. Here is the answer.

Proposition 1.9. *The multiplication operator M_m from Example 1.8 is bounded if and only if $m \in \mathrm{L}^\infty(\mathbb{R}^n)$ and in this case $\|M_m\|_{\mathcal{L}(\mathrm{L}^2(\mathbb{R}^n))} = \|m\|_{\mathrm{L}^\infty(\mathbb{R}^n)}$.*

Proof. If $m \in \mathrm{L}^\infty(\mathbb{R}^n)$, then we immediately get

$$\|M_m u\|_{\mathrm{L}^2(\mathbb{R}^n)} = \|mu\|_{\mathrm{L}^2(\mathbb{R}^n)} \leq \|m\|_{\mathrm{L}^\infty(\mathbb{R}^n)} \|u\|_{\mathrm{L}^2(\mathbb{R}^n)} \qquad (u \in \mathrm{L}^2(\mathbb{R}^n)),$$

so M_m is bounded. Conversely, assume that M_m is bounded with operator norm C. For any set E of finite measure we have $\mathbf{1}_E \in \mathrm{L}^2(\mathbb{R}^n)$ and therefore

$$\int_E (|m|^2 - C^2) \, \mathrm{d}x = \|M_m \mathbf{1}_E\|^2_{\mathrm{L}^2(\mathbb{R}^n)} - C^2 \|\mathbf{1}_E\|^2_{\mathrm{L}^2(\mathbb{R}^n)} \leq 0.$$

Since the Lebesgue measure is σ-finite, the measurable function $|m|^2 - C^2$ must be non-positive a.e. on $\mathbb{R}^n$ and $\|m\|_{\mathrm{L}^\infty(\mathbb{R}^n)} \leq C$ follows. □

We also recall the following important result from functional analysis.

Proposition 1.10 (Closed graph theorem, [Alt16, Thm. 7.9])**.** *If L is a closed operator from X to Y with* $\mathrm{dom}(L) = X$*, then L is bounded.*

Working with unbounded linear operators, we always have to pay attention to the domains. If you have not made this experience before, be aware: This is nothing to get sloppy about!

Definition 1.11. *Let L_1, L_2 be linear operators from X to Y, let L_3 be a linear operator from Y into another Banach space Z and let $\lambda \in \mathbb{C}$. Then*

$$\begin{aligned} \mathrm{dom}(\lambda L_1) &:= \mathrm{dom}(L_1), & (\lambda L_1)u &:= \lambda(L_1 u), \\ \mathrm{dom}(L_1 + L_2) &:= \mathrm{dom}(L_1) \cap \mathrm{dom}(L_2), & (L_1 + L_2)u &:= L_1 u + L_2 u, \\ \mathrm{dom}(L_3 L_1) &:= \big\{u \in \mathrm{dom}(L_1) \mid L_1 u \in \mathrm{dom}(L_3)\big\}, & (L_3 L_1)u &:= L_3(L_1 u). \end{aligned}$$

In the case of the identity operator id_X *on X, we write $\lambda + L_1$ instead of $\lambda \mathrm{id}_X + L_1$.*

Be aware that we do not talk about the sum or composition of two linear relations here. In particular, do *not* think of the sum of two linear relations or operators as the sum of the corresponding linear subspaces of $X \times Y$. This is not the same thing!

1.2. Spectral theory

A linear relation R always has an inverse relation, cf. Definition 1.2 (c), but even if R is an operator, this inverse need not be an operator. We introduce the following notions.

Definition 1.12. *A linear operator L between X and Y is called* invertible *if $L^{-1} \in \mathcal{L}(Y, X)$.*

Definition 1.13. *Let L be a linear operator in X. Then*

(a) $\sigma(L) := \{\lambda \in \mathbb{C} \mid \lambda - L$ *is not invertible*$\}$ *is the* spectrum *of L,*

(b) $\varrho(L) := \mathbb{C} \setminus \sigma(L) = \big\{\lambda \in \mathbb{C} \mid (\lambda - L)^{-1} \in \mathcal{L}(X)\big\}$ *is the* resolvent set *of L,*

(c) *the map $R(\cdot, L)\colon \varrho(L) \to \mathcal{L}(X)$ with $R(\lambda, L) := (\lambda - L)^{-1}$ for $\lambda \in \varrho(L)$ is called* resolvent *of L.*

Remark 1.14. It follows from Remark 1.5 that invertible operators are closed. Applying this to $\lambda - L$ with $\lambda \in \varrho(L)$, we see that L is closed whenever $\varrho(L) \neq \emptyset$.

Below, we collect some fundamental properties of the resolvent and the resolvent set. Vector-valued holomorphic functions — defined in complete analogy with the scalar-valued case from your complex analysis course — appear for the first time. You find further background in Appendix A.3 but as for now, you merely need to know that power series are (vector-valued) holomorphic inside their disc of convergence.

Proposition 1.15. *Let L be a linear operator in X. Then:*

(a) *For all $\lambda, \mu \in \varrho(L)$ we have the* resolvent identity

$$R(\lambda, L) - R(\mu, L) = (\mu - \lambda) R(\lambda, L) R(\mu, L).$$

(b) *The resolvent set $\varrho(L)$ is open.*

(c) *The resolvent is a holomorphic function and for all $\mu \in \varrho(L)$ and all $\lambda \in \mathbb{C}$ with $|\lambda - \mu| < \|R(\mu, L)\|^{-1}_{\mathcal{L}(X)}$ we find $\lambda \in \varrho(L)$ and the power series expansion*

$$R(\lambda, L) = \sum_{k=0}^{\infty} (\mu - \lambda)^k R(\mu, L)^{k+1}.$$

(d) *For all $\lambda \in \varrho(L)$ and $k \in \mathbb{N}$ we have*

$$\frac{\mathrm{d}^k}{\mathrm{d}\lambda^n} R(\lambda, L) = (-1)^k k! R(\lambda, L)^{k+1}.$$

Proof. In order to prove the resolvent identity, let $\lambda, \mu \in \varrho(L)$. For all $u \in X$ we have $R(\mu, L)u \in \operatorname{dom}(L)$, so we can calculate

$$\begin{aligned} R(\lambda, L)u - R(\mu, L)u &= R(\lambda, L)\big[u - \lambda R(\mu, L)u + LR(\mu, L)u\big] \\ &= R(\lambda, L)\big[u - \lambda R(\mu, L)u - (\mu - L)R(\mu, L)u + \mu R(\mu, L)u\big] \\ &= R(\lambda, L)\big[\mu R(\mu, L)u - \lambda R(\mu, L)u\big] \\ &= (\mu - \lambda) R(\lambda, L) R(\mu, L)u. \end{aligned}$$

For the proof of the other assertions, let $\mu \in \varrho(L)$ and $\lambda \in \mathbb{C}$ with $|\lambda - \mu| < \|R(\mu, L)\|_{\mathcal{L}(X)}^{-1}$. Then

$$\lambda - L = \lambda - \mu + \mu - L = \big((\lambda - \mu)R(\mu, L) + 1\big)(\mu - L)$$

is invertible if and only if $(\lambda - \mu)R(\mu, L) + 1$ is invertible. By the Neumann series criterion, this is the case if

$$\|(\lambda - \mu)R(\mu, L)\|_{\mathcal{L}(X)} < 1,$$

which holds by assumption. Thus, $\lambda \in \varrho(L)$, which already shows (b). Furthermore, the Neumann series criterion also gives the formula

$$R(\lambda, L) = R(\mu, L)\sum_{k=0}^{\infty}(\mu - \lambda)^k R(\mu, L)^k = \sum_{k=0}^{\infty}(\mu - \lambda)^k R(\mu, L)^{k+1}.$$

Now, holomorphy of the resolvent map as well as the formula for the derivatives follow by this power series representation, compare with Example A.20. □

Example 1.16. We show that the spectrum of our multiplication operator M_m in $\mathrm{L}^2(\mathbb{R}^n)$ is the *essential range* of m given by

$$\operatorname{essran}(m) \coloneqq \Big\{\lambda \in \mathbb{C} \;\Big|\; \forall \varepsilon > 0 : \big|\{x \in \mathbb{R}^n \mid |m(x) - \lambda| < \varepsilon\}\big| > 0\Big\}.$$

(If you have not seen this measure theoretic construction before, think of a continuous function m. In this case $\operatorname{essran}(m) = \overline{m(\mathbb{R}^n)}$.)

If λ is not in the essential range of m, then there is an $\varepsilon > 0$ such that $\{x \in \mathbb{R}^n \mid |m(x) - \lambda| < \varepsilon\}$ is a nullset. This means that $|m - \lambda| \geq \varepsilon$ almost everywhere, which entails that $(\lambda - m)^{-1} \in \mathrm{L}^\infty(\mathbb{R}^n)$. Thus, by Proposition 1.9 the corresponding multiplication operator $M_{(\lambda - m)^{-1}}$ is bounded and for $u \in \mathrm{L}^2(\mathbb{R}^n)$ we have $(\lambda - m)M_{(\lambda - m)^{-1}}u = u$. This proves $(\lambda - M_m)M_{(\lambda - m)^{-1}} = \mathrm{id}_{\mathrm{L}^2(\mathbb{R}^n)}$. In the same way $M_{(\lambda - m)^{-1}}(\lambda - M_m) = \mathrm{id}_{\operatorname{dom}(M_m)}$ follows. Consequently, $\lambda - M_m$ is invertible with $(\lambda - M_m)^{-1} = M_{(\lambda - m)^{-1}}$ and we have shown that $\mathbb{C} \setminus \operatorname{essran}(m) \subseteq \varrho(M_m)$.

It remains to prove $\operatorname{essran}(m) \subseteq \sigma(M_m)$. Let $\lambda \in \operatorname{essran}(m)$ and $\varepsilon > 0$. Since λ is in the essential range, the set $\{x \in \mathbb{R}^n \mid |m(x) - \lambda| < \varepsilon\}$ has strictly positive measure, so we can choose a subset A_ε with strictly positive and finite measure. Then $u_\varepsilon \coloneqq \mathbf{1}_{A_\varepsilon} \in \operatorname{dom}(M_m)$ and we have

$$\|(\lambda - M_m)u_\varepsilon\|_{\mathrm{L}^2(\mathbb{R}^n)} = \|(\lambda - m)\mathbf{1}_{A_\varepsilon}\|_{\mathrm{L}^2(\mathbb{R}^n)} \leq \varepsilon\|\mathbf{1}_{A_\varepsilon}\|_{\mathrm{L}^2(\mathbb{R}^n)} = \varepsilon\|u_\varepsilon\|_{\mathrm{L}^2(\mathbb{R}^n)}.$$

This shows that $\lambda - M_m$ cannot have a bounded inverse, so $\lambda \in \sigma(M_m)$.

1.3. Adjoints and self-adjointness

Our next aim is to generalize the notion of an adjoint of a linear operator. The main problem in the context of closed operators is dealing with the domains. A huge advantage of using the language of linear relations is that it allows us to define an adjoint relation without any complications.

For our purposes it will be sufficient (and much easier) to work in Hilbert spaces from now on.

Definition 1.17. *Let R be a linear relation between H and K. Then the* adjoint relation *of R is the linear relation $R^* \subseteq K \times H$ given by*

$$(w, z) \in R^* :\Longleftrightarrow \langle u, z\rangle_H = \langle v, w\rangle_K \text{ for all } (u, v) \in R.$$

Definition 1.17 will become more concise when both R and R^* are linear operators. At the level of relations, it helps to think about adjoints in the following way. On $H \times K$ (and similarly on $K \times H$) we consider the canonical inner product

$$\langle (u, v), (z, w)\rangle := \langle u, z\rangle_H + \langle v, w\rangle_K$$

that turns this space itself into a Hilbert space and define the unitary operator

$$\Phi\colon H \times K \to K \times H, \quad \Phi(u, v) := (-v, u). \tag{1.1}$$

Now, we have again a sharp look at Definition 1.17: It says precisely that

$$R^* = (\Phi(R))^\perp = \Phi(R^\perp), \tag{1.2}$$

where the orthogonal complement has to be taken once in $K\times H$ and once in $H\times K$.

Caution: In general the adjoint relation of a linear operator is not again a linear operator. The next result tells us precisely, when this is the case.

Proposition 1.18. *Let L be a linear operator from H to K. Then:*

(a) L^* is closed.

(b) L^* is a linear operator if and only if L is densely defined. In this case we have

$$\mathrm{dom}(L^*) = \{w \in K \mid \exists z \in H\colon \langle u, z\rangle_H = \langle Lu, w\rangle_K \text{ for all } u \in \mathrm{dom}(L)\}$$

and

$$L^* w = z.$$

In particular, we have

$$\langle u, L^* w\rangle_H = \langle Lu, w\rangle_K \qquad (u \in \mathrm{dom}(L),\ w \in \mathrm{dom}(L^*)). \tag{1.3}$$

(c) If L is closed and densely defined, then so is L^* and $(L^*)^* = L$.

Proof. (a) By (1.2), $L^* \subseteq K \times H$ is the orthogonal complement of a set, and hence closed.

(b) We have

$$\begin{aligned} z \in \ker((L^*)^{-1}) &\iff (0, z) \in L^* \\ &\overset{(1.2)}{\iff} 0 = \langle (0, z), (-v, u) \rangle = \langle z, u \rangle_H \text{ for all } (u, v) \in L \\ &\iff z \in \operatorname{dom}(L)^\perp. \end{aligned}$$

That L^* is a linear operator means that $\ker((L^*)^{-1})$ is trivial, which now is equivalent to $\operatorname{dom}(L)$ being dense. Since in this case the relations L and L^* are linear operators, we can reformulate Definition 1.17 as stated in (b).

(c) Properties (a) and (b) imply that L^* is a closed operator. In order to determine its adjoint, we need the formalism of (1.1) and (1.2) with the roles of H and K reversed, which amounts to replacing Φ by $-\Phi^{-1}$. Thus, we get

$$(L^*)^* = (-\Phi^{-1}(L^*))^\perp = \big(-\Phi^{-1}(\Phi(L^\perp))\big)^\perp = \big(L^\perp\big)^\perp = \overline{L} = L.$$

In particular, $(L^*)^*$ is a linear operator, so L^* is densely defined thanks to (b). □

Example 1.19. We have seen in Example 1.8 that the multiplication operator M_m is densely defined. So, it has an adjoint operator that we want to determine now. Our experience with bounded operators (or matrices) tells us that we should try to prove $(M_m)^* = M_{\overline{m}}$. Let's do it. We simply write $\langle \cdot , \cdot \rangle$ for the inner product on $\mathrm{L}^2(\mathbb{R}^n)$.

Given $w \in \operatorname{dom}(M_{\overline{m}})$, we obtain for all $u \in \operatorname{dom}(M_m)$ that

$$\langle M_m u, w \rangle = \int_{\mathbb{R}^n} mu \cdot \overline{w} \, \mathrm{d}x = \int_{\mathbb{R}^n} u \cdot \overline{\overline{m} w} \, \mathrm{d}x = \langle u, M_{\overline{m}} w \rangle .$$

In view of Proposition 1.18 (b) this means $w \in \operatorname{dom}((M_m)^*)$ with $(M_m)^* w = M_{\overline{m}} w$.

Conversely, assume $w \in \operatorname{dom}((M_m)^*)$ and note carefully that we cannot simply do the same calculation backwards, because we do not know yet that $\overline{m} w \in \mathrm{L}^2(\mathbb{R}^n)$. We resort to the trick from Example 1.8: Given any $u \in \mathrm{L}^2(\mathbb{R}^n)$, we have $\frac{u}{1+|m|^2} \in \operatorname{dom}(M_m)$, which allows us to compute

$$\left\langle u, \frac{(M_m)^* w}{1 + |m|^2} \right\rangle = \left\langle \frac{u}{1 + |m|^2}, (M_m)^* w \right\rangle = \left\langle M_m\Big(\frac{u}{1 + |m|^2}\Big), w \right\rangle = \left\langle u, \frac{\overline{m} w}{1 + |m|^2} \right\rangle.$$

Since $u \in \mathrm{L}^2(\mathbb{R}^n)$ is arbitrary, we conclude $\frac{(M_m)^* w}{1+|m|^2} = \frac{\overline{m} w}{1+|m|^2}$ in $\mathrm{L}^2(\mathbb{R}^n)$ and now $w \in \operatorname{dom}(M_{\overline{m}})$ with $M_{\overline{m}} w = (M_m)^* w$ follows.

Let us collect some useful computation rules for adjoints.

Proposition 1.20. *Let L, L_1, L_2 be densely defined linear operators from H to K. Then:*

(a) If $L_1 \subseteq L_2$, then $L_2^* \subseteq L_1^*$.

(b) $(L_1 + L_2)^* \supseteq L_1^* + L_2^*$ with equality if L_2 is bounded.

(c) If $H = K$, then $(\lambda - L)^* = \overline{\lambda} - L^*$ for all $\lambda \in \mathbb{C}$.

(d) If L_3 is a densely defined linear operator from K to some other Hilbert space, then $L_1^* L_3^* \subseteq (L_3 L_1)^*$ with equality if L_3 is bounded.

(e) $(L^{-1})^* = (L^*)^{-1}$ and in particular, L^* is invertible if L is invertible.

(f) $(\overline{L})^* = L^*$.

(g) $\ker(L^*) = \operatorname{ran}(L)^\perp$.

(h) If L is bounded, then L^* is bounded and $\|L\|_{\mathcal{L}(H,K)} = \|L^*\|_{\mathcal{L}(K,H)}$.

Proof. Parts (a) - (g) are good finger exercises to become acquainted with adjoints, see Exercise 1.4. We give the proof of (h). To this end, let $w \in K$ and note that $\phi(u) := \langle Lu, w\rangle_K$ is a linear functional that satisfies

$$|\phi(u)| \leq \|w\|_K \|L\|_{\mathcal{L}(H,K)} \|u\|_H \qquad (u \in H).$$

The Riesz representation theorem yields a $z \in H$ with $\|z\|_H \leq \|w\|_K \|L\|_{\mathcal{L}(H,K)}$ and $\langle u, z\rangle_H = \phi(u)$ for all $u \in H$. Hence, $L^* w = z$ and L^* is bounded with $\|L^*\|_{\mathcal{L}(K,H)} \leq \|L\|_{\mathcal{L}(H,K)}$. Applying this to L^* in place of L yields the reverse estimate thanks to Proposition 1.18 (c). □

In view of (e) and (c) of Proposition 1.20 one might hope that there is an easy link between $\sigma(L)$ and $\sigma(L^*)$ and this is indeed the case.

Proposition 1.21. *Let L be a closed and densely defined linear operator in H. Then*

$$\sigma(L^*) = \left\{\overline{\lambda} \mid \lambda \in \sigma(L)\right\}$$

and $R(\overline{\lambda}, L^*) = R(\lambda, L)^*$ for all $\lambda \in \varrho(L)$.

Proof. Let $\lambda \in \varrho(L)$. Then $\lambda - L$ is invertible and by (e), (c) and (h) of Proposition 1.20 also $(\lambda - L)^*$ is invertible with

$$R(\lambda, L)^* = \left[(\lambda - L)^{-1}\right]^* = \left[(\lambda - L)^*\right]^{-1} = (\overline{\lambda} - L^*)^{-1} = R(\overline{\lambda}, L^*).$$

Thanks to Proposition 1.18 (c) we can apply the same reasoning to L^* in place of L and get the converse implication that $\overline{\lambda} \in \varrho(L^*)$ implies $\lambda \in \varrho((L^*)^*) = \varrho(L)$. □

In many fields of mathematics and physics, a particular important class of operators are the self-adjoint ones, i.e., operators that coincide with their adjoint.

Definition 1.22. *(a) A linear operator L in H is* symmetric *if*

$$\langle Lu, v\rangle_H = \langle u, Lv\rangle_H \qquad (u, v \in \operatorname{dom}(L)).$$

(b) A linear operator L in H is self-adjoint *if $L = L^*$ holds.*

Note that self-adjoint operators are automatically densely defined — this is a consequence of Proposition 1.18 (b).

Lemma 1.23. *If L is a densely defined operator in H, then L is symmetric if and only if $L \subseteq L^*$.*

Proof. If L is symmetric and $v \in \operatorname{dom}(L)$, then we have for all $u \in \operatorname{dom}(L)$ that $\langle u, Lv\rangle_H = \langle Lu, v\rangle_H$. Thus, $v \in \operatorname{dom}(L^*)$ and $L^*v = Lv$ by Proposition 1.18 (b). Conversely, if $L \subseteq L^*$, then for all $u, v \in \operatorname{dom}(L)$ we find with the help of Proposition 1.18 (b) that

$$\langle Lu, v\rangle_H = \langle u, L^*v\rangle_H = \langle u, Lv\rangle_H,$$

so L is symmetric. □

Be aware that every self-adjoint operator is necessarily symmetric but that even 'densely defined and symmetric' does not mean 'self-adjoint'! The point is that for an operator to be self-adjoint, one not only needs the mapping behavior described by symmetry, but also that $\operatorname{dom}(L) = \operatorname{dom}(L^*)$ and this can be a severe constraint.

Example 1.24. For which functions m is our multiplication operator M_m self-adjoint? In Example 1.19 we have seen that $M_m^* = M_{\overline{m}}$, which has the same domain as M_m. So, it turns out that M_m is self-adjoint if and only if m is real-valued (almost everywhere).

We close this section with a useful criterion for a densely defined, closed and symmetric operator to be indeed self-adjoint.

Proposition 1.25. *Let L be a densely defined, closed and symmetric operator in H.*

(a) If $\varrho(L) \cap \mathbb{R} \neq \emptyset$, then L is self-adjoint.

(b) L is self-adjoint if and only if $\sigma(L) \subseteq \mathbb{R}$.

Proof. We only prove (a), the proof of (b) is left as Exercise 1.6. Let $\lambda \in \varrho(L)$ be a real number. Then, since L is symmetric, also $\lambda - L$ is symmetric and Lemma 1.23 tells us that $\lambda - L \subseteq \lambda - L^*$. Furthermore, Proposition 1.21 yields $\lambda = \overline{\lambda} \in \varrho(L^*) \cap \varrho(L)$. Now, Exercise 1.2 (b) gives the conclusion. □

1.4. Exercises

Exercise 1.1 (Why domains matter)**.** Let S and T be two linear operators from X into itself.

(a) Provide a counterexample to show that in general even if S and T are closed operators with the same domain, then $S+T$ need not be closed.

(b) Suppose that S is closed and T is bounded. Prove that ST is closed and provide a counterexample that in general TS is not.

Exercise 1.2 (Upgrading inclusions to equalities)**.** Let S, T be operators between Banach spaces X and Y with $S \subseteq T$.

(a) Show that if S is surjective and T is injective, then $S = T$.

(b) Conclude that if $X = Y$, then

$$\varrho(S) \cap \varrho(T) \neq \emptyset \quad \Longrightarrow \quad S = T.$$

(c) Suppose that T is closed and that $\mathrm{dom}(S)$ is a core for T. Prove that $\overline{S} = T$.

Remark: These criteria are (surprisingly) useful in applications.

Exercise 1.3. Let L be a linear operator in a Banach space X with $-1 \in \varrho(L)$.

(a) Argue that $L(1+L)^{-1}$ is bounded.

(b) Let $j, k \in \mathbb{N}$. Prove that

$$(L(1+L)^{-1})^j u \in \mathrm{dom}(L^k) \quad \Longleftrightarrow \quad u \in \mathrm{dom}(L^k).$$

Exercise 1.4. Prove the calculation rules (a) - (g) stated in Proposition 1.20.

Exercise 1.5. Let L be a closed operator in a Hilbert space H. Prove that for all $\lambda \in \varrho(L)$ we have

$$\|(\lambda - L)^{-1}\|_{\mathcal{L}(H)} \geq \frac{1}{\mathrm{dist}(\lambda, \sigma(L))}.$$

Exercise 1.6. Let L be an operator in a Hilbert space H.

(a) Suppose that $\langle Lu, u\rangle \in \mathbb{R}$ holds for all $u \in \mathrm{dom}(L)$. Show that L is symmetric.

(b) Let L be densely defined, closed, and symmetric. Show that L is self-adjoint if and only if $\sigma(L) \subseteq \mathbb{R}$.

2. Sectorial operators and sesquilinear forms

In this lecture, we first study sectorial operators in Hilbert spaces. These are closed operators, whose spectrum is localized in a sector of the complex plane and for which we have a specific control of the norm of the resolvent. All (or almost all) of the differential operators that we will encounter in the later parts of the ISem lectures will then turn out to be of this class.

In the second part, we introduce the form method as a way to construct sectorial operators from sesquilinear forms.

Notation 2.1. In the whole lecture, H is a complex Hilbert space.

2.1. Sectorial operators

We start right away with the central definition.

Definition 2.2. *(a) For $\varphi \in (0, \pi)$ the (open)* sector of angle φ *is denoted by* $\mathrm{S}_\varphi := \{z \in \mathbb{C} \setminus \{0\} \mid |\arg(z)| < \varphi\}$. *Furthermore, we set* $\mathrm{S}_0 := (0, \infty)$.

(b) A linear operator L in H is called sectorial of angle $\varphi \in [0, \pi)$ *if $\sigma(L) \subseteq \overline{\mathrm{S}_\varphi}$ and we have the following resolvent bounds: for all $\psi \in (\varphi, \pi)$ there is some $C_\psi \geq 0$ such that*

$$\|R(\lambda, L)\|_{\mathcal{L}(H)} \leq \frac{C_\psi}{|\lambda|} \qquad (\lambda \in \mathbb{C} \setminus \overline{\mathrm{S}_\psi}).$$

(c) If L is a sectorial operator, then

$$\varphi_L := \inf\{\varphi \in [0, \pi) \mid L \text{ is sectorial of angle } \varphi\}$$

is called sectoriality angle *of L.*

Sectorial operators are closed, because they have a non-empty resolvent set, see Remark 1.14. It is a remarkable feature of these operators that only the knowledge about the location of the spectrum and the decay rate of the resolvent gives a huge amount of information about the functional analytic properties of the operator, including density of its domain. The aim of this section is to elaborate on this connection.

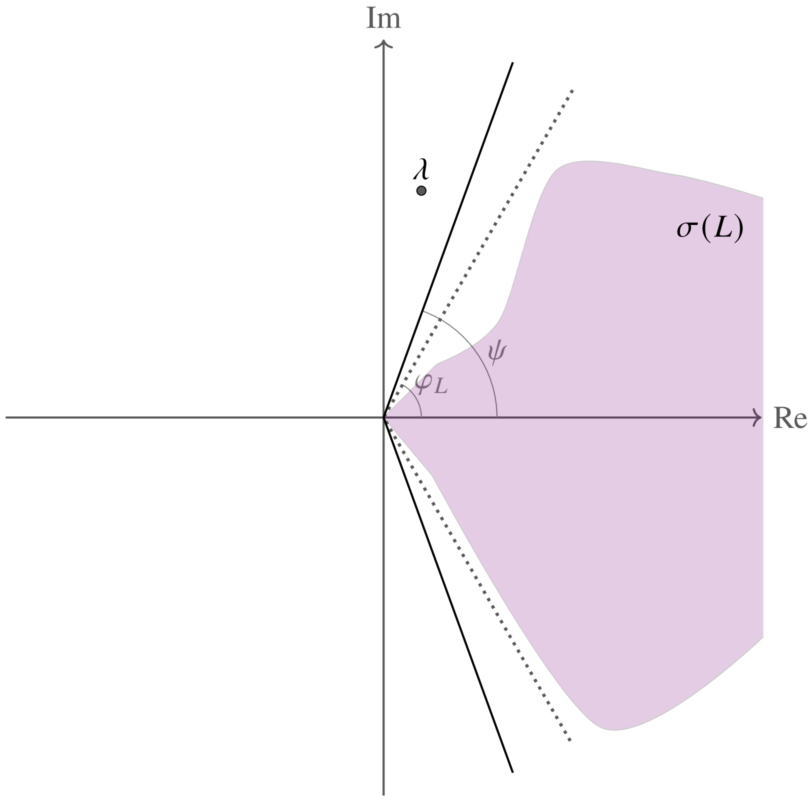

Figure 2.1.: The spectrum of a sectorial operator is trapped in a sector symmetric around the positive real axis. For λ outside of the closure of larger sectors S_ψ, a resolvent bound by $\frac{C_\psi}{|\lambda|}$ is required.

We need an elementary lemma that shows that closed subspaces of Hilbert spaces are even closed with respect to weak convergence.

Lemma 2.3. *If K is a closed subspace of H and (u_j) is a sequence in K that converges weakly in H to some $u \in H$, then $u \in K$.*

Proof. Since K is closed, it suffices to prove $u \in (K^\perp)^\perp$ $(= K)$. Given any $v \in K^\perp$, weak convergence of (u_j) yields the claim

$$\langle u, v\rangle_H = \lim_{j\to\infty} \left\langle u_j, v\right\rangle_H \overset{u_j \in K}{=} \lim_{j\to\infty} 0 = 0. \qquad \square$$

Proposition 2.4. *For a sectorial operator L in H the following properties hold:*

(a) For all $u \in \overline{\mathrm{ran}(L)}$, we have

$$(1 + tL)^{-1}u \to 0 \quad as \quad t \to \infty.$$

(b) We have a topological decomposition

$$H = \ker(L) \oplus \overline{\operatorname{ran}(L)}.$$

In particular, L is injective if and only if it has dense range.

(c) The operator L is densely defined.

(d) For all $u \in H$, we have

$$(1 + tL)^{-1}u \to u \quad \text{as} \quad t \searrow 0.$$

(e) For every $k \in \mathbb{N}$ we have

i) $\overline{\operatorname{dom}(L^k)} = H$,

ii) $\overline{\operatorname{dom}(L^k) \cap \operatorname{ran}(L^k)} = \overline{\operatorname{ran}(L)}$,

iii) $\operatorname{ran}(L^k(1 + L)^{-2k}) = \operatorname{dom}(L^k) \cap \operatorname{ran}(L^k)$.

Proof. We will repeatedly use the fact that if $(T_j) \subseteq \mathcal{L}(H)$ is a bounded sequence with the property that $(T_j u)$ converges to zero for all u in some set $U \subseteq H$, then the same is true for all $u \in \overline{U}$.[1]

(a) By sectoriality of L the operators $(1 + tL)^{-1} = t^{-1}(t^{-1} + L)^{-1}$ are uniformly bounded for $t > 0$. Hence, it suffices to prove convergence for $u \in \operatorname{ran}(L)$. Choosing $v \in \operatorname{dom}(L)$ with $Lv = u$, we get

$$(1 + tL)^{-1}u = t^{-1}(1 + tL)^{-1}tLv = t^{-1}\big(v - (1 + tL)^{-1}v\big)$$

and thus by sectoriality

$$\|(1 + tL)^{-1}u\|_H \le t^{-1}\big(\|v\|_H + C\|v\|_H\big) \xrightarrow{t\to\infty} 0.$$

(b) First, let $u \in \ker(L) \cap \overline{\operatorname{ran}(L)}$. Then, using $Lu = 0$ and (a), we find

$$u = \lim_{t\to\infty}(1 + tL)^{-1}(1 + tL)u = \lim_{t\to\infty}(1 + tL)^{-1}u = 0.$$

Hence, $\ker(L) + \overline{\operatorname{ran}(L)}$ is a direct sum. In order to show that it spans the whole space H, let $u \in H$ be given and consider the sequence $(u_j) := ((1 + jL)^{-1}u)$. Note that for $u \in \ker(L)$ this sequence is constantly u and for $u \in \overline{\operatorname{ran}(L)}$ it converges to zero by (a). This means that once we have shown that the sum spans H, then the limit exists indeed for every $u \in H$,

$$Pu := \lim_{j\to\infty}(1 + jL)^{-1}u$$

[1] This is a 2ε-argument. Convince yourself!

will be the projection onto $\ker(L)$ along $\overline{\operatorname{ran}(L)}$, and P is bounded by sectoriality of L, which by definition means that the decomposition will be topological. Surprisingly, we can pull ourselves up by our own bootstraps and use (u_j) also to prove that $u \in \ker(L) + \overline{\operatorname{ran}(L)}$ as follows.

By sectoriality of L the sequence (u_j) is bounded in H, so we can pick a weakly convergent subsequence $(u_{j_k})_{k\in\mathbb{N}}$ and we will call its limit v. Furthermore,

$$Lu_j = L(1 + jL)^{-1}u = \tfrac{1}{j}\big(u - (1 + jL)^{-1}u\big) = \tfrac{1}{j}\big(u - u_j\big) \tag{2.1}$$

tends to 0 in the limit as $j \to \infty$ because $(u - u_j)$ is a bounded sequence in H. Now, consider the sequence $((u_{j_k}, Lu_{j_k}))_{k\in\mathbb{N}}$ in $L \subseteq H \times H$. Since L is closed and $((u_{j_k}, Lu_{j_k}))_{k\in\mathbb{N}}$ converges weakly in $H \times H$ to $(v, 0)$, Lemma 2.3 implies $(v, 0) \in L$ and this means that $v \in \ker(L)$.

Moreover, $(u - u_{j_k})_{k\in\mathbb{N}}$ converges weakly in H to $u - v$ and by (2.1) we have $u - u_{j_k} = j_k Lu_{j_k} \in \operatorname{ran}(L)$. Applying Lemma 2.3 again, this time in the closed subspace $\overline{\operatorname{ran}(L)}$, we find that $u-v \in \overline{\operatorname{ran}(L)}$ and $u = v+(u-v) \in \ker(L)+\overline{\operatorname{ran}(L)}$ follows.

(c) By Exercise 2.2, $(1 + L)^{-1}$ is an injective sectorial operator. Hence,

$$H \overset{(b)}{=} \overline{\operatorname{ran}\big((1 + L)^{-1}\big)} = \overline{\operatorname{dom}(1 + L)} = \overline{\operatorname{dom}(L)}.$$

(d) Thanks to (c) it suffices to prove convergence for $u \in \operatorname{dom}(L)$. In this case

$$u = (1 + tL)^{-1}(1 + tL)u = (1 + tL)^{-1}u + t(1 + tL)^{-1}Lu$$

and sectoriality of L yields that rightmost term vanishes as $t \searrow 0$.

(e) This proof is left as Exercise 2.3 □

Example 2.5. Consider again our multiplication operator M_m in $\mathrm{L}^2(\mathbb{R}^n)$. By Example 1.16 we know that for M_m being sectorial of angle $\varphi \in [0, \pi)$, it is necessary that $\operatorname{essran}(m) = \sigma(M_m) \subseteq \overline{\mathrm{S}_\varphi}$. We will now show that this is already sufficient.

Suppose that $\operatorname{essran}(m) \subseteq \overline{\mathrm{S}_\varphi}$, let $\psi \in (\varphi, \pi)$ and take $\lambda \in \mathbb{C} \setminus \overline{\mathrm{S}_\psi}$. By Example 1.16 and Proposition 1.9 we know that

$$\|R(\lambda, M_m)\|_{\mathcal{L}(\mathrm{L}^2(\mathbb{R}^n))} = \|M_{(\lambda-m)^{-1}}\|_{\mathcal{L}(\mathrm{L}^2(\mathbb{R}^n))} = \|(\lambda - m)^{-1}\|_{\mathrm{L}^\infty(\mathbb{R}^n)} \leq \frac{1}{\operatorname{dist}(\lambda, \overline{\mathrm{S}_\varphi})}.$$

Now, consider the compact circle arc $A_\psi := \{\mathrm{e}^{\mathrm{i}\theta} \mid \psi \leq |\theta| \leq \pi\}$. Since $\psi > \varphi$, we have $\operatorname{dist}(A_\psi, \overline{\mathrm{S}_\varphi}) > 0$ and since $\overline{\mathrm{S}_\varphi}$ is a cone, we get

$$\operatorname{dist}(\lambda, \overline{\mathrm{S}_\varphi}) = |\lambda| \operatorname{dist}\bigg(\frac{\lambda}{|\lambda|}, \overline{\mathrm{S}_\varphi}\bigg) \geq |\lambda| \operatorname{dist}(A_\psi, \overline{\mathrm{S}_\varphi}).$$

Combining the previous two estimates yields the resolvent bounds needed for sectoriality.

2.2. Elliptic forms

In this and the next section we will encounter a general method to construct sectorial linear operators in the Hilbert space H via sesquilinear forms defined on suitable subspaces of H. This domain of the form will be denoted by V.

Notation 2.6. For the rest of the lecture, V will be another Hilbert space that is continuously and densely embedded into H.

Definition 2.7. *(a) A map $a\colon V \times V \to \mathbb{C}$ is called a* sesquilinear form *on V if it is linear in the first argument and anti-linear in the second, i.e., for all $u, v, w \in V$ and all $\lambda \in \mathbb{C}$ we have*

$$a(u + \lambda v, w) = a(u, w) + \lambda a(v, w) \quad \text{and} \quad a(u, v + \lambda w) = a(u, v) + \overline{\lambda} a(u, w).$$

We denote by $a(u) \coloneqq a(u, u)$ the corresponding quadratic form *on V.*

(b) A sesquilinear form $a\colon V \times V \to \mathbb{C}$ is called

i) symmetric *if $a(u, v) = \overline{a(v, u)}$ holds for all $u, v \in V$,*

ii) bounded *if there is some $C \geq 0$ such that*

$$|a(u, v)| \leq C\|u\|_V \|v\|_V \qquad (u, v \in V),$$

iii) accretive *if $\operatorname{Re}\big(a(u)\big) \geq 0$ holds for all $u \in V$,*

iv) coercive *if there is some $c > 0$ such that*

$$\operatorname{Re}\big(a(u)\big) \geq c\|u\|_V^2 \qquad (u \in V),$$

v) elliptic *if there are $\omega, \kappa > 0$ such that*

$$\operatorname{Re}\big(a(u)\big) + \omega\|u\|_H^2 \geq \kappa\|u\|_V^2 \qquad (u \in V).$$

Remark 2.8. (a) The quadratic form that corresponds to a sesquilinear form seems to contain less information than the whole form, but this is not true. The latter can be reconstructed from the former via the *polarization identity*

$$a(u, v) = \frac{1}{4}\big(a(u + v) - a(u - v) + \mathrm{i}a(u + \mathrm{i}v) - \mathrm{i}a(u - \mathrm{i}v)\big) \qquad (u, v \in V)$$

that can be verified straightforwardly by expanding the right-hand side.

(b) Likewise, a fulfills the *parallelogram identity*

$$a(u + v) + a(u - v) = 2a(u) + 2a(v) \qquad (u, v \in V).$$

Lemma 2.9. *A sesquilinear form $a\colon V \times V \to \mathbb{C}$ is symmetric if and only if $a(u) \in \mathbb{R}$ for all $u \in V$.*

Proof. If a is symmetric, then $a(u) = a(u,u) = \overline{a(u,u)} = \overline{a(u)}$, so $a(u) \in \mathbb{R}$. Conversely, suppose $a(u) \in \mathbb{R}$ for all $u \in V$. We introduce a new sesquilinear form a^* on V by $a^*(u,v) := \overline{a(v,u)}$ and need to prove that $a^* = a$. But this follows from the polarization identity since the corresponding quadratic forms satisfy $a^*(u) = \overline{a(u)} = a(u)$ for all $u \in V$. □

Lemma 2.10 (Schwarz's inequality)**.** *Let $a, b\colon V \times V \to \mathbb{C}$ be two symmetric sesquilinear forms and assume that $|a(u)| \le b(u)$ for all $u \in V$. Then we have*

$$\bigl|a(u,v)\bigr| \le b(u)^{1/2} b(v)^{1/2} \qquad (u, v \in V).$$

Proof. Let $u, v \in V$ and choose $\gamma \in \mathbb{C}$ with $|\gamma| = 1$ in such a way that $\gamma a(u,v) = a(\gamma u, v)$ is a real number. The quadratic form associated with a only takes real values, see Lemma 2.9. Thus, the polarization identity yields

$$\gamma a(u,v) = \operatorname{Re}\bigl(a(\gamma u, v)\bigr) = \frac{1}{4}\bigl(a(\gamma u + v) - a(\gamma u - v)\bigr).$$

Taking absolute values on both sides, we can use the triangle inequality, the assumption and the parallelogram identity in Remark 2.8 (b) to obtain

$$|a(u,v)| \le \frac{1}{4}\bigl(b(\gamma u + v) + b(\gamma u - v)\bigr) = \frac{1}{2}\bigl(b(\gamma u) + b(v)\bigr) = \frac{1}{2}\bigl(b(u) + b(v)\bigr).$$

Note that in the end we obtain a bound for $|a(u,v)|$ that does not contain γ anymore and is valid for all choices of $u, v \in V$. In particular, we can take any numbers $s > b(u)^{1/2}$ and $t > b(v)^{1/2}$ and use our bound for $(\frac{1}{s}u, \frac{1}{t}v)$ in place of (u, v) in order to obtain

$$\bigl|a(u,v)\bigr| = st\Bigl|a\Bigl(\frac{1}{s}u, \frac{1}{t}v\Bigr)\Bigr| \le \frac{st}{2}\Bigl[b\Bigl(\frac{1}{s}u\Bigr) + b\Bigl(\frac{1}{t}v\Bigr)\Bigr] = \frac{st}{2}\Bigl(\frac{b(u)}{s^2} + \frac{b(v)}{t^2}\Bigr) \le st.$$

Letting $s \searrow b(u)^{1/2}$ and $t \searrow b(v)^{1/2}$ yields the claim. □

We give a name to the sesquilinear form a^* that already appeared in the proof of Lemma 2.9.

Definition 2.11. *Let a be a sesquilinear form on V. Then*

$$a^*(u,v) := \overline{a(v,u)} \qquad (u, v \in V)$$

is called the adjoint form *of a.*

Remark 2.12. The adjoint form satisfies $|a^*(u,v)| = |a(v,u)|$ and $\operatorname{Re}(a^*(u)) = \operatorname{Re}(a(u))$ for all $u, v \in V$. Thus, a^* is bounded, accretive, coercive and/or elliptic, if and only if a has the respective property. Furthermore, a is symmetric if and only if $a = a^*$.

2.3. Sectorial operators from elliptic forms

Now, we start to construct an associated sectorial operator for a given bounded, elliptic and accretive sesquilinear form.

Definition 2.13. *We say that* $\phi\colon V \to \mathbb{C}$ *is* anti-linear *if for all* $u, v \in V$ *and all* $\lambda \in \mathbb{C}$ *we have*

$$\phi(u + \lambda v) = \phi(u) + \overline{\lambda}\phi(v).$$

The space

$$V' := \{\phi\colon V \to \mathbb{C} \mid \phi \textit{ anti-linear and continuous}\}$$

is called anti-dual *space of* V.

For the application of $\phi \in V'$ *to* $v \in V$ *we will write*

$$\langle \phi, v\rangle_{V',V} := \phi(v).$$

Remark 2.14. (a) The anti-dual space V' of V is a Banach space when equipped with the norm

$$\|\phi\|_{V'} := \sup_{v\in V, \|v\|_V=1} |\phi(v)|.$$

(b) Since V is dense in H, each $u \in H$ is uniquely determined by the associated anti-linear functional $\phi_u\colon V \ni v \mapsto \langle u, v\rangle_H$. Since the embedding $V \hookrightarrow H$ is continuous, we get

$$\left|\phi_u(v)\right| = \left|\langle u, v\rangle_H\right| \le \|u\|_H\|v\|_H \lesssim \|u\|_H\|v\|_V.$$

Hence, $\phi_u \in V'$ and identifying u with ϕ_u, we find that $H \hookrightarrow V'$. The three spaces V, H and V' together are sometimes called a *Gelfand triple*.

(c) Let a be a bounded sesquilinear form on V and let $u \in V$. Then $V \ni w \mapsto a(u, w)$ is an element of V' by the boundedness of a. So, assigning u to this functional gives a linear operator $\mathscr{L}$ from V to V' that is bounded once again thanks to the boundedness of a:

$$\left|(\mathscr{L}u)(w)\right| = \left|a(u, w)\right| \le C\|u\|_V\|w\|_V \qquad (u, w \in V).$$

Definition 2.15. *Let* $a\colon V \times V \to \mathbb{C}$ *be a bounded sesquilinear form. The bounded operator* $\mathscr{L}\colon V \to V'$ *from Remark 2.14 (c) that fulfills*

$$a(u, w) = \langle \mathscr{L}u, w\rangle_{V',V} \qquad (u, w \in V)$$

is called the operator associated with a on V.

The form a also provides us with an operator in H that is a close relative to $\mathscr{L}$.

Definition 2.16. *Let $a\colon V \times V \to \mathbb{C}$ be a bounded sesquilinear form. Then*

$$L := \big\{(u, v) \in H \times H \mid u \in V \text{ and } a(u, w) = \langle v, w\rangle_H \text{ for all } w \in V\big\}$$

is the operator associated with a in H.

We have to justify that L is indeed an operator and we will show that L is the maximal restriction of $\mathscr{L}$ to H.

Proposition 2.17. *Let L be the relation from Definition 2.16. Then L is a linear operator in H with*

$$\operatorname{dom}(L) = \{u \in V \mid \mathscr{L}u \in H\}, \quad Lu = \mathscr{L}u \text{ for } u \in \operatorname{dom}(L).$$

Proof. This is a direct consequence of the definition of $\mathscr{L}$ and our understanding of the embedding $H \hookrightarrow V'$, see Remark 2.14 (b):

$$\begin{aligned}(u, v) \in L \quad &\iff \quad (u, v) \in V \times H \text{ and } \langle \mathscr{L}u, w\rangle_{V',V} = \langle v, w\rangle_{V',V} \text{ for all } w \in V \\ &\iff \quad (u, v) \in V \times H \text{ and } \mathscr{L}u = v.\end{aligned}$$

□

Our next goal will be to show that the operator L becomes sectorial under suitable assumptions on a.

Lemma 2.18. *Let a be a bounded sesquilinear form on V with associated operators $\mathscr{L}$ on V and L in H. Let $\lambda \in \mathbb{C}$ and consider the sesquilinear form*

$$a_\lambda\colon V \times V \to \mathbb{C}, \quad a_\lambda(u, v) = a(u, v) + \lambda \langle u, v\rangle_H .$$

Then a_λ is also bounded and the operators associated with a_λ on V and in H equal $\lambda + \mathscr{L}$ and $\lambda + L$, respectively.

Proof. We have $|\lambda \langle u, v\rangle_H | \leq C|\lambda|\|u\|_V\|v\|_V$ for all $u, v \in V$ since V is continuously embedded into H. Hence, as the sum of two bounded sesquilinear forms on V, a_λ is itself bounded. Moreover,

$$a_\lambda(u, v) = a(u, v) + \lambda \langle u, v\rangle_H = \langle \mathscr{L}u, v\rangle_{V',V} + \lambda \langle u, v\rangle_{V',V} = \langle (\lambda + \mathscr{L})u, v\rangle_{V',V}$$

shows that the operator associated with a_λ on V is given by $\lambda + \mathscr{L}$. The identity maps V into H. Now, Proposition 2.17 yields that the maximal restriction of $\lambda + \mathscr{L}$ to H is at the same time $\lambda + L$ and the operator associated with a_λ in H. □

The ellipticity condition in 2.7 (b) yields invertibility of some shifted versions of $\mathscr{L}$. This is a consequence of the famous Lax–Milgram lemma from functional analysis that we recall for convenience:

Proposition 2.19 (Lax–Milgram, [Alt16, Thm. 6.2]). *Let $a\colon V\times V\to\mathbb{C}$ be a bounded and coercive sesquilinear form. Then the associated operator on V is an isomorphism with $\|\mathscr{L}^{-1}\|_{\mathcal{L}(V',V)}\le 1/c$, where c is the coercivity constant from Definition 2.7 (b).*

Corollary 2.20. *Let $a\colon V\times V\to\mathbb{C}$ be a bounded and elliptic sesquilinear form with ellipticity constants ω and κ. Then for all $\lambda\in\mathbb{C}$ with $\operatorname{Re}(\lambda)\ge\omega$ the operator $\lambda+\mathscr{L}$ is an isomorphism with $\|(\lambda+\mathscr{L})^{-1}\|_{\mathcal{L}(V',V)}\le 1/\kappa$.*

Proof. The operator $\lambda+\mathscr{L}$ is associated with the form a_λ by Lemma 2.18 and, thanks to ellipticity of a, we have for all $u\in V$ that

$$\operatorname{Re}\bigl(a_\lambda(u)\bigr)=\operatorname{Re}\bigl(a(u)+\omega\|u\|_H^2+(\lambda-\omega)\|u\|_H^2\bigr)\ge\kappa\|u\|_V^2+\bigl(\operatorname{Re}(\lambda)-\omega\bigr)\|u\|_H^2\ge\kappa\|u\|_V^2.$$

Thus, this form is coercive with coercivity constant κ and the claim follows from the Lax–Milgram lemma. □

Now, we can formulate a sufficient condition when a sesquilinear form on V gives rise to a sectorial operator in H.

Theorem 2.21. *If a is a bounded, accretive and elliptic sesquilinear form on V, then the associated operator L in H is sectorial of angle $\pi/2$. Moreover, we have the particular resolvent bound*

$$\|(\lambda+L)^{-1}\|_{\mathcal{L}(H)}\le\frac{1}{\operatorname{Re}(\lambda)}\qquad(\operatorname{Re}(\lambda)>0).$$

Proof. Firstly, we note that for all $\lambda\in\mathbb{C}$ with $\operatorname{Re}(\lambda)>0$ and all $u\in\operatorname{dom}(L)$ the Cauchy–Schwarz inequality and the accretivity of a yield

$$\|(\lambda+L)u\|_H\|u\|_H\ge\operatorname{Re}\bigl(\langle(\lambda+L)u,u\rangle_H\bigr)=\operatorname{Re}(\lambda)\|u\|_H^2+\operatorname{Re}\bigl(a(u)\bigr)\ge\operatorname{Re}(\lambda)\|u\|_H^2.$$

Thus, whenever $\operatorname{Re}(\lambda)>0$, we have the a priori estimate

$$\|u\|_H\le\frac{1}{\operatorname{Re}(\lambda)}\|(\lambda+L)u\|_H\qquad(u\in\operatorname{dom}(L)).\tag{2.2}$$

Let ω and κ denote again the ellipticity constants of a and let, in a first step, $\lambda\in\mathbb{C}$ with $\operatorname{Re}(\lambda)\ge\omega$ be given. Then, by Corollary 2.20, the operator $\lambda+\mathscr{L}\colon V\to V'$ is an isomorphism. We want to show that $\lambda+L$ is invertible, too, meaning that $-\lambda\in\varrho(L)$. Since $\lambda+L$ is a restriction of $\lambda+\mathscr{L}$, injectivity follows. To prove surjectivity, let $v\in H$. Interpreting v as element of V', we can set $u:=(\lambda+\mathscr{L})^{-1}v\in V$. Then $(\lambda+\mathscr{L})u=v\in H$ and Proposition 2.17 yields $u\in\operatorname{dom}(\lambda+L)=\operatorname{dom}(L)$ and $(\lambda+L)u=v$. To show that $\lambda+L$ is invertible, it remains to make sure that the inverse $(\lambda+L)^{-1}$ is bounded. But this follows by setting $u=(\lambda+L)^{-1}v$ in (2.2) and we get the resolvent bound

$$\|(\lambda+L)^{-1}\|_{\mathcal{L}(H)}\le\frac{1}{\operatorname{Re}(\lambda)}\qquad(\operatorname{Re}(\lambda)\ge\omega).\tag{2.3}$$

Now, consider $\lambda \in \mathbb{C}$ with $0 < \operatorname{Re}(\lambda) \le \omega$ and set $\mu := \omega + \mathrm{i}\operatorname{Im}(\lambda)$. By our considerations above, we have $-\mu \in \varrho(L)$ and

$$|-\lambda - (-\mu)| = \operatorname{Re}(\mu) - \operatorname{Re}(\lambda) < \operatorname{Re}(\mu) \overset{(2.3)}{\le} \frac{1}{\|(-\mu - L)^{-1}\|_{\mathcal{L}(H)}}.$$

Proposition 1.15 (c) yields $-\lambda \in \varrho(L)$ and (2.2) shows that the particular resolvent bound (2.3) in fact holds for all $\operatorname{Re}(\lambda) > 0$.

To prove sectoriality of L, let $\psi \in (\pi/2, \pi)$ and let $\mu \in \mathbb{C} \setminus \overline{\mathrm{S}_\psi}$ be given. Then $-\operatorname{Re}(\mu) \ge \sin(\psi - \pi/2) \cdot |\mu|$ and therefore

$$\|R(\mu, L)\|_{\mathcal{L}(H)} = \|(-\mu + L)^{-1}\|_{\mathcal{L}(H)} \le \frac{1}{\operatorname{Re}(-\mu)} \le \frac{1}{\sin(\psi - \pi/2)} \cdot \frac{1}{|\mu|},$$

which is the desired estimate. □

Sectorial operators of angle $\pi/2$ with the specific bound as in the theorem above have their own hallmark.

Definition 2.22. *An operator L in H with the particular resolvent bound for* $\operatorname{Re}(\lambda) > 0$ *from Theorem 2.21 is called* m-accretive.

The sectorial operators coming from a form have dense domain in H by the abstract theory of sectorial operators, but more is true.

Proposition 2.23. *Let L be an operator in H that is associated with a bounded, accretive and elliptic sesquilinear form on V. Then* $\operatorname{dom}(L)$ *is dense in V.*

Proof. Let $u \in V$ be orthogonal to $\operatorname{dom}(L)$ in V, i.e., $\langle u, v\rangle_V = 0$ for all $v \in \operatorname{dom}(L)$. This element u defines a bounded anti-linear functional $\psi_u \colon V \to \mathbb{C}$ via $\psi_u(v) := \langle u, v\rangle_V$. Let ω, κ denote the ellipticity constants of a. Then we know from the hypotheses and Remark 2.12 that the shifted adjoint form a^*_ω is coercive and, by the Lax–Milgram lemma, there exists some $w \in V$ such that

$$\langle u, v\rangle_V = \psi_u(v) = a^*_\omega(w, v) = \overline{a_\omega(v, w)} \qquad (v \in V). \tag{2.4}$$

For $v \in \operatorname{dom}(L)$ we even get

$$0 = \langle u, v\rangle_V = \overline{a_\omega(v, w)} = \overline{\langle(\omega + L)v, w\rangle_H}.$$

Now, $-\omega \in \varrho(L)$, so $(\omega + L)v$ for $v \in \operatorname{dom}(L)$ runs through all of H. This implies that $\langle z, w\rangle_H = 0$ for all $z \in H$ and thus $w = 0$. Taking a look back to (2.4), this entails $u = 0$ and we have proved that $(\operatorname{dom}(L))^\perp = \{0\}$ in V, so $\operatorname{dom}(L)$ is dense in V. □

Our notion of adjoint form from Definition 2.11 is consistent in the following sense.

Proposition 2.24. *Let a be a bounded, elliptic and accretive sesquilinear form on V with associated operator L in H. Then the operator associated with a^* in H coincides with L^*. In particular, L is self-adjoint if a is symmetric.*

Proof. Since L is sectorial of angle $\pi/2$ by Theorem 2.21, it is densely defined by Proposition 2.4 (c) and the adjoint operator exists. Furthermore, the form a^* is also bounded, elliptic and accretive by Remark 2.12. Hence, the associated operator, which we denote by $\widetilde{L}$, is sectorial of angle $\pi/2$, again by Theorem 2.21.

For all $u \in \operatorname{dom}(L)$ and all $v \in \operatorname{dom}(\widetilde{L})$ we have

$$\langle Lu, v\rangle_H = a(u, v) = \overline{a^*(v, u)} = \overline{\langle \widetilde{L}v, u\rangle_H} = \langle u, \widetilde{L}v\rangle_H.$$

Hence, by definition of the adjoint operator, $v \in \operatorname{dom}(L^*)$ and $L^*v = \widetilde{L}v$, which means $\widetilde{L} \subseteq L^*$. Due to Proposition 1.21 we have $-1 \in \varrho(L^*) \cap \varrho(\widetilde{L})$, hence the equality $\widetilde{L} = L^*$ follows from Exercise 1.2 (b). □

2.4. Sectorial forms

We have seen that accretive forms give rise to sectorial operators of angle $\pi/2$. Accretivity of a precisely means that we have $a(u) \in \overline{\mathrm{S}_{\pi/2}}$ for all $u \in V$, so it seems that we are on to something . . .

Definition 2.25. *A sesquilinear form a on V is* sectorial of angle $\varphi \in [0, \pi/2)$ *if $a(u) \in \overline{\mathrm{S}_\varphi}$ for all $u \in V$.*

Note that a sectorial form is automatically accretive and that we could have equivalently defined sectoriality as follows.

Lemma 2.26. *A sesquilinear form a on V is sectorial of angle $\varphi \in [0, \pi/2)$ if and only if $|\operatorname{Im}(a(u))| \leq \tan(\varphi)\operatorname{Re}(a(u))$ for all $u \in V$.*

In the same manner as we did in Lemma 2.18 for shifts of forms, one proves the following version for rotated forms.

Lemma 2.27. *Let a be a bounded sesquilinear form on V with associated operators $\mathscr{L}$ on V and L in H. For every $\theta \in (-\pi, \pi]$ the sesquilinear form $(\mathrm{e}^{\mathrm{i}\theta}a)(\cdot\,,\cdot) := \mathrm{e}^{\mathrm{i}\theta}a(\cdot\,,\cdot)$ is bounded and the associated operators are $\mathrm{e}^{\mathrm{i}\theta}\mathscr{L}$ and $\mathrm{e}^{\mathrm{i}\theta}L$, respectively.*

Imposing the stronger condition of sectoriality instead of accretivity, we get the following enhanced version of Theorem 2.21.

Theorem 2.28. *Let a be a bounded and elliptic sesquilinear form on V that is sectorial of angle $\varphi \in [0, \pi/2)$. Then the associated operator L in H is also sectorial of angle φ.*

Proof. Let $\theta \in [0, \pi/2 - \varphi)$ and let γ be one of $e^{\pm i\theta}$. In a first step, we show that the rotated form γa is accretive and elliptic.

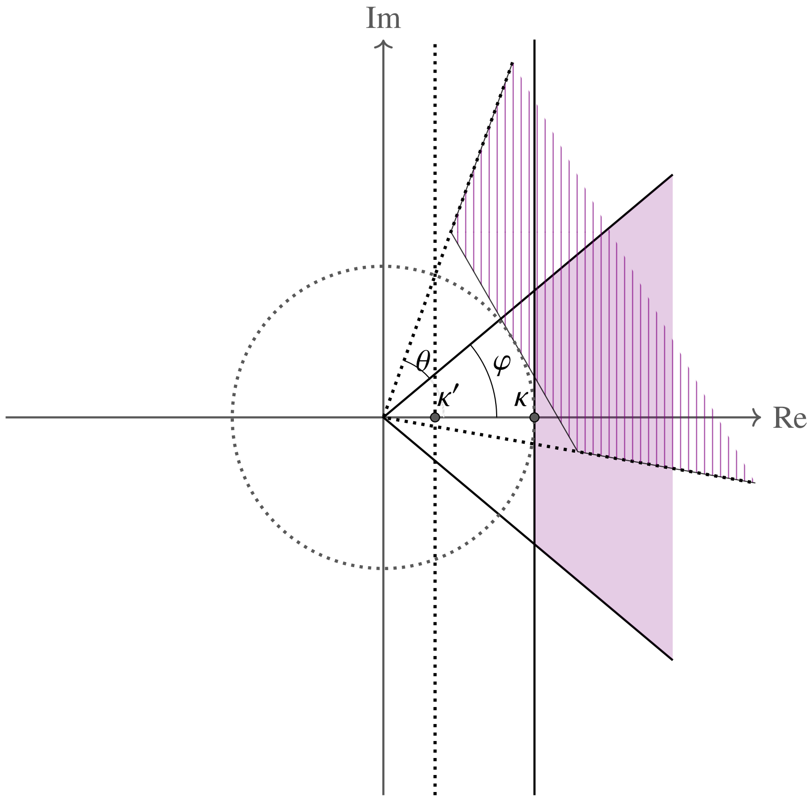

Figure 2.2.: Rotation by an angle $\theta \in (0, \pi/2 - \varphi)$. The part of S_φ in the half plane defined by $\mathrm{Re}(z) \geq \kappa$ is rotated into a new half plane defined by $\mathrm{Re}(z) \geq \kappa'$.

By sectoriality of a we get for all $v \in V$ that

$$\gamma a(v) \in \gamma \overline{S_\varphi} \subseteq \overline{S_{\varphi+\theta}} \subseteq \{z \mid \mathrm{Re}(z) \geq 0\},$$

meaning that γa is accretive. Next, we let $\omega, \kappa > 0$ be as in the definition of ellipticity for a. Ellipticity and sectoriality imply that for all $v \in V$ with $\|v\|_V = 1$ we have

$$a(v) + \omega \|v\|_H^2 \in \left\{z \in \overline{S_\varphi} \mid \mathrm{Re}(z) \geq \kappa\right\} =: W.$$

We have again $\gamma W \subseteq \overline{S_{\varphi+\theta}}$. Rotation preserves the distance to the origin, so we also have $\mathrm{dist}(0, \gamma W) = \mathrm{dist}(0, W) = \kappa$ and thus $\gamma W \subseteq \{z \in \mathbb{C} \mid \mathrm{Re}(z) \geq \kappa'\}$ for some $\kappa' > 0$, see also Figure 2.2.

Hence, for every $v \in V$ normalized to $\|v\|_V = 1$, we have shown

$$\mathrm{Re}\big(\gamma a(v)\big) + \cos(\theta)\omega\|v\|_H^2 = \mathrm{Re}\big(\gamma\big(a(v) + \omega\|v\|_H^2\big)\big) \geq \kappa',$$

which means that γa is elliptic.

Based on this result we can apply Theorem 2.21 to the rotated forms γa and obtain with the help of Lemma 2.27 that the associated operators γL in H are m-accretive.

Now, let $\varepsilon > 0$ with $\varphi + 2\varepsilon < \pi/2$. We prove the resolvent estimate required for sectoriality for $\lambda \in \mathbb{C} \setminus \overline{\mathrm{S}_{\varphi+2\varepsilon}}$. We fix $\theta := \pi/2 - \varphi - \varepsilon$ and take the correct sign in the definition of γ to make sure that $\gamma\lambda \in \mathbb{C} \setminus \overline{\mathrm{S}_{\pi/2+\varepsilon}}$. By m-accretivity of γL, we get that $\lambda - L = \gamma^{-1}(\gamma\lambda - \gamma L)$ is invertible with

$$\|R(\lambda, L)\|_{\mathcal{L}(H)} = \|(\gamma\lambda - \gamma L)^{-1}\|_{\mathcal{L}(H)} \leq \frac{1}{\mathrm{Re}(-\gamma\lambda)} \leq \frac{1}{\sin(\varepsilon)|\gamma\lambda|} = \frac{1}{\sin(\varepsilon)|\lambda|}. \qquad \square$$

2.5. Exercises

Exercise 2.1 (Cauchy–Schwarz for sectorial forms)**.** Let $a\colon V \times V \to \mathbb{C}$ be a sectorial sesquilinear form. Prove that there exists $C > 0$ such that

$$\big|a(u, v)\big| \leq C\big(\mathrm{Re}(a(u))\big)^{1/2}\big(\mathrm{Re}(a(v))\big)^{1/2} \qquad (u, v \in V).$$

Hint: Write $a = a_{\mathrm{Re}} + \mathrm{i}a_{\mathrm{Im}}$, with symmetric sesquilinear forms a_{Re} and a_{Im} on V.

Exercise 2.2. Let L be an injective linear operator in a Hilbert space H and let $-\lambda \in \varrho(L)$.

(a) Prove that

$$1 - \lambda(\lambda + L)^{-1} = (1 + \lambda L^{-1})^{-1}.$$

(b) Show that if L is a sectorial operator, then so is L^{-1}.

(c) Conclude that if L is sectorial and $\varepsilon > 0$, then $(\varepsilon + L)^{-1}$ is an injective sectorial operator.

Exercise 2.3. Prove Proposition 2.4 (e).

Hint: You have already seen the proof of i) and ii) — find it!

Exercise 2.4 (Perturbed Dirac operators – I)**.** Let D be a self-adjoint operator in a Hilbert space H, let $B \in \mathcal{L}(H)$ and assume there exists $\kappa > 0$ such that $\mathrm{Re}\langle Bu, u\rangle_H \geq \kappa\|u\|_H^2$ for all $u \in \mathrm{ran}(D)$. By Proposition 1.20 (g) you know that D induces the topological (orthogonal) splitting

$$H = \ker(D) \oplus \overline{\mathrm{ran}(D)}.$$

In this exercise, you will learn how this splitting is perturbed through B.

(a) Argue that B^*D and DB are closed, densely defined, and adjoint to each other.

(b) Prove that $\ker(D) \oplus \overline{\operatorname{ran}(B^*D)}$ is a topological decomposition in H.

(c) Show that $\overline{\operatorname{ran}(D)} \cap \ker(DB) = \{0\}$.

(d) Conclude that topologically

$$H = \ker(D) \oplus \overline{\operatorname{ran}(B^*D)} = \overline{\operatorname{ran}(DB)} \oplus \ker(DB).$$

(e) Prove the following identities for the subspaces appearing in (d):

$$\ker(D) = \ker(B^*D), \quad B^* \overline{\operatorname{ran}(D)} = \overline{\operatorname{ran}(B^*D)}, \quad \operatorname{ran}(DB) = \operatorname{ran}(D).$$

Exercise 2.5 (Perturbed Dirac operators – II)**.** We continue the study of the operators D and B from Exercise 2.4 and prove resolvent estimates for the perturbed Dirac operators DB and B^*D. These operators play a fundamental role in recent solvability theory of boundary value problems for elliptic systems [AA11]. In the project phase you will have the possibility to delve into that topic.

Let $t \in \mathbb{R} \setminus \{0\}$.

(a) Show that

$$\|(\mathrm{i}t - DB)u\|_H \geq \varepsilon |t| \|u\|_H, \qquad (u \in \operatorname{dom}(DB) \cap \overline{\operatorname{ran}(D)}),$$

where $\varepsilon = \kappa / \|B\|_{\mathcal{L}(H)}$.

Hint: Find some inspiration from Exercise 1.6.

(b) Show that

$$\|(\mathrm{i}t - B^*D)u\|_H \geq \varepsilon |t| \|u\|_H, \qquad (u \in \operatorname{dom}(B^*D) \cap \overline{\operatorname{ran}(B^*D)}),$$

for some different constant $\varepsilon > 0$ depending on the same parameters.

(c) Conclude that $\mathrm{i}\mathbb{R} \setminus \{0\} \subseteq \varrho(DB) \cap \varrho(B^*D)$ and that there exists $C > 0$ such that for all $t \in \mathbb{R} \setminus \{0\}$ we have

$$\|t(\mathrm{i}t - DB)^{-1}\|_{\mathcal{L}(H)} \leq C \quad \text{and} \quad \|t(\mathrm{i}t - B^*D)^{-1}\|_{\mathcal{L}(H)} \leq C.$$

(d) Conclude that $(DB)^2$ and $(B^*D)^2$ are sectorial.

3. The form method for elliptic operators

The objective of this lecture is the Hilbert space treatment of the Dirichlet-Laplace operator and more general elliptic differential operators in divergence form as a particular example of the form method from Lecture 2. This requires a new scale of spaces, called Sobolev spaces: They are Hilbert spaces of differentiable functions, where derivatives will be understood in a new, weaker sense.

Establishing the basic properties of Sobolev spaces will allow us to recapitulate tools from measure and integration theory that we consider as prerequisites for our lectures: convolution, mollifiers, L^p-approximation by smooth functions, integration by parts. Readers, who are not familiar with these concepts will find proofs and further background in standard textbooks such as [AE09].

Notation 3.1. Throughout the lecture, $\Omega \subseteq \mathbb{R}^n$ is an open, non-empty set.

3.1. Test functions and convolutions

We use the usual spaces $\mathrm{C}_\mathrm{c}^\infty(\Omega)$ of smooth functions with compact support in Ω, sometimes called *test functions*, and $\mathrm{L}^1_{\mathrm{loc}}(\Omega)$ of measurable functions[1] that are integrable over all compact subsets of Ω. For measurable $u, v\colon \mathbb{R}^n \to \mathbb{C}$ the *convolution* is defined by

$$(u * v)(x) := \int_{\mathbb{R}^n} u(x-y)v(y)\,\mathrm{d}y \qquad (x \in \mathbb{R}^n),$$

provided this integral exists for a.e. $x \in \mathbb{R}^n$, e.g., if $u \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$ and $v \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n)$.

Definition 3.2. *Let $u\colon \mathbb{R}^n \to \mathbb{C}$ be measurable. The family $(u_t)_{t>0}$ defined by*

$$u_t(x) := \frac{1}{t^n} u\Big(\frac{x}{t}\Big) \qquad (x \in \mathbb{R}^n)$$

is called mollifier *associated with u.*

[1] We will try to be precise when the difference between 'functions' and 'equivalence classes of functions' matters on a mathematical level, but we do not make a linguistic distinction.

We will mainly need the following properties of test functions. Throughout the lectures we write $B(x, r)$ for the Euclidean ball with center x and radius r.

Proposition 3.3 (Toolkit for test functions)**.**

(a) *There is a radial $\theta \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$ such that $\mathbf{1}_{B(0,1)} \leq \theta \leq \mathbf{1}_{B(0,2)}$.*

(b) *If $u \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $\theta \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$, then $\theta * u \in \mathrm{C}^\infty(\mathbb{R}^n)$ and for every multi-index $\alpha \in \mathbb{N}_0^n$ we have*

$$\partial^\alpha(\theta * u) = (\partial^\alpha \theta) * u.$$

(c) *If $1 \leq p \leq \infty$, $u \in \mathrm{L}^p(\mathbb{R}^n)$ and $\theta \in \mathrm{L}^1(\mathbb{R}^n)$, then*

$$\|\theta * u\|_{\mathrm{L}^p(\mathbb{R}^n)} \leq \|\theta\|_{\mathrm{L}^1(\mathbb{R}^n)} \|u\|_{\mathrm{L}^p(\mathbb{R}^n)}.$$

Moreover, if $p < \infty$ and θ is normalized to $\int_{\mathbb{R}^n} \theta \, \mathrm{d}x = 1$, then

$$\lim_{t \searrow 0} \theta_t * u = u \qquad (\textit{in } \mathrm{L}^p(\mathbb{R}^n)).$$

(d) *If $1 \leq p < \infty$, then $\mathrm{C}_\mathrm{c}^\infty(\Omega)$ is dense in $\mathrm{L}^p(\Omega)$.*

3.2. Weak derivatives

Suppose that $u \in \mathrm{C}^k(\Omega)$ and that $\alpha \in \mathbb{N}_0^n$ is a multi-index with $|\alpha| \leq k$. Integration by parts yields

$$\int_\Omega u \cdot \partial^\alpha \phi \, \mathrm{d}x = (-1)^{|\alpha|} \int_\Omega \partial^\alpha u \cdot \phi \, \mathrm{d}x \qquad (\phi \in \mathrm{C}_\mathrm{c}^\infty(\Omega)). \tag{3.1}$$

The partial derivative $\partial^\alpha u$ is the unique element in $\mathrm{L}^1_{\mathrm{loc}}(\Omega)$ with this property, because of the following principle.

Lemma 3.4 (Fundamental lemma in the calculus of variations)**.** *Let $v \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$. Then $v = 0$ if and only if*

$$\int_\Omega v \cdot \phi \, \mathrm{d}x = 0 \qquad (\phi \in \mathrm{C}_\mathrm{c}^\infty(\Omega)). \tag{3.2}$$

Proof. For the interesting direction let us assume that v satisfies (3.2). Writing Ω as a countable union of increasing bounded open sets Ω_j, $j = 1, 2, \ldots$, with $\overline{\Omega_j} \subseteq \Omega_{j+1}$, for example

$$\Omega_j := \{x \in \Omega \mid |x| < j \text{ and } \operatorname{dist}(x, \partial\Omega) \geq j^{-1}\},$$

it suffices to prove $v = 0$ a.e. on Ω_j for every j.

To this end, we set $v_j := \mathbf{1}_{\Omega_{j+1}} v$, which we think of being defined as zero on $\mathbb{R}^n \setminus \Omega_{j+1}$.[2] Since $\overline{\Omega_{j+1}}$ is a compact subset of Ω, we have $v_j \in \mathrm{L}^1(\mathbb{R}^n)$. Let $\theta \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$ be supported in the unit ball and with integral 1. We want to use (3.2) with $\phi(x) = \theta_t(y-x)$. This choice is admissible for $y \in \Omega_j$ and $t < \operatorname{dist}(\overline{\Omega_j}, \mathbb{R}^n \setminus \Omega_{j+1}) =: \varepsilon_j$ since in this case ϕ vanishes outside of $\Omega_{j+1} \subseteq \Omega$. Thus, we get

$$0 = \int_\Omega v(x) \cdot \theta_t(y-x) \,\mathrm{d}x = \int_{\mathbb{R}^n} v_j(x) \cdot \theta_t(y-x) \,\mathrm{d}x = (\theta_t * v_j)(y) \qquad (y \in \Omega_j),$$

whenever $t < \varepsilon_j$. In the limit as $t \searrow 0$ we have $\theta_t * v_j \to v_j$ in $\mathrm{L}^1(\mathbb{R}^n)$, see Proposition 3.3 (c). Thus, $\theta_t * v_j \to v$ in $\mathrm{L}^1(\Omega_j)$ and therefore $v = 0$ a.e. on Ω_j. □

Now, we change our perspective on (3.1): Instead of assuming that u is differentiable, we assume that there are functions u_α in place of $\partial^\alpha u$, which make the integration by parts formulæ work.

Definition 3.5. *Let $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$. Given a multi-index $\alpha \in \mathbb{N}_0^n$, we say that $\partial^\alpha u$* exists in the weak sense *if there is some $u_\alpha \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ such that*

$$\int_\Omega u \cdot \partial^\alpha \phi \,\mathrm{d}x = (-1)^{|\alpha|} \int_\Omega u_\alpha \cdot \phi \,\mathrm{d}x \qquad (\phi \in \mathrm{C}_\mathrm{c}^\infty(\Omega)).$$

In this case we write $\partial^\alpha u := u_\alpha$ and call it weak α-th derivative *of u. If all weak derivatives of order $|\alpha| = 1$ exist, we call $\nabla u := (\partial_1 u, \ldots, \partial_n u)$ the* weak gradient *of u and say that u is* weakly differentiable *in Ω.*

Remark 3.6. We have seen above that if u is continuously differentiable, then u is also differentiable in the weak sense and the two notions of derivatives coincide.

Let us illustrate the concepts with some (non-)examples.

Example 3.7 (The weak derivative of the absolute value)**.** The absolute value function $u\colon \mathbb{R} \to \mathbb{C}$, $u(x) = |x|$ is not differentiable in $x = 0$, but it feels like in some sense its derivative should still be the sign function

$$v(x) := \begin{cases} -1 & \text{if } x < 0, \\ 1 & \text{if } x \geq 0. \end{cases}$$

We claim that u is weakly differentiable with weak derivative $u' = v$. Indeed, u and v are locally integrable and for any $\phi \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R})$ integration by parts yields

$$\begin{aligned} \int_\mathbb{R} u \cdot \phi' \,\mathrm{d}x &= \int_0^\infty x\phi'(x) \,\mathrm{d}x - \int_{-\infty}^0 x\phi'(x) \,\mathrm{d}x \\ &= -\int_0^\infty \phi(x) \,\mathrm{d}x + \int_{-\infty}^0 \phi(x) \,\mathrm{d}x = -\int_\mathbb{R} v \cdot \phi \,\mathrm{d}x. \end{aligned}$$

[2] In other words: We will agree on ‘$0 \cdot \text{undefined} = 0$’ in such a situation.

Example 3.8 (Weak derivative of the sign function)**.** The sign function v from the previous example is not weakly differentiable. Indeed, suppose to the contrary that $v' = w$ exists in the weak sense. Then, given $\phi \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R})$, we obtain

$$\int_{\mathbb{R}} w \cdot \phi \,\mathrm{d}x = -\int_{\mathbb{R}} v \cdot \phi' \,\mathrm{d}x = -\int_0^\infty \phi'(x)\,\mathrm{d}x + \int_{-\infty}^0 \phi'(x)\,\mathrm{d}x = 2\phi(0).$$

Lemma 3.4 applied on $\mathbb{R} \setminus \{0\}$ yields $w = 0$ a.e. on $\mathbb{R} \setminus \{0\}$, but then the left-hand side is zero even for all $\phi \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R})$, which leads to a contradiction if we take a test function with $\phi(0) \neq 0$.

Further below we shall see the deeper reason why weak differentiation 'over a jump on the real line', as in the case of the sign function, does not work. However, this is specific to the one-dimensional setting and in higher dimensions we can even 'differentiate over a (suitable) singularity', see Exercise 3.4 for a particularly naughty example.

To gain further trust in the concept of weak derivatives, we verify that many properties of classical derivatives remain valid in the weak sense.

Lemma 3.9. *Let $k \in \mathbb{N}$, $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq k$ and $\lambda \in \mathbb{C}$. Let $u, v \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ have all weak derivatives up to order k in the weak sense and let $\eta \in \mathrm{C}^\infty(\Omega)$. Then we have*

(a) $\partial^\alpha(u + \lambda v) = \partial^\alpha u + \lambda \partial^\alpha v$,

(b) $\partial^{\alpha-\beta}(\partial^\beta u) = \partial^\alpha u$ $(= \partial^\beta(\partial^{\alpha-\beta} u))$ *provided that* $\beta \leq \alpha$,

(c) $\partial^\alpha(\eta u) = \sum_{0 \leq \beta \leq \alpha} \binom{\alpha}{\beta} \partial^{\alpha-\beta} u \cdot \partial^\beta \eta$ *(*Leibniz' rule*).*

Proof. Part (a) follows from the definition of weak derivatives. Next, we let $\beta \leq \alpha$ and $\phi \in \mathrm{C}_\mathrm{c}^\infty(\Omega)$. Since $\partial^{\alpha-\beta}\phi \in \mathrm{C}_\mathrm{c}^\infty(\Omega)$, we obtain

$$(-1)^{|\alpha-\beta|} \int_\Omega \partial^\beta u \cdot \partial^{\alpha-\beta}\phi \,\mathrm{d}x = (-1)^{|\alpha|} \int_\Omega u \cdot \partial^\alpha \phi \,\mathrm{d}x = \int_\Omega \partial^\alpha u \cdot \phi \,\mathrm{d}x,$$

which proves (b). As for (c) it suffices to prove the product rule corresponding to $|\alpha| = 1$ since this implies Leibniz' rule by induction just as in the case of classical derivatives. Given α with $|\alpha| = 1$, we get for all $\phi \in \mathrm{C}_\mathrm{c}^\infty(\Omega)$ that

$$\begin{aligned}\int_\Omega \eta u \cdot \partial^\alpha \phi \,\mathrm{d}x &= \int_\Omega u \cdot \eta \partial^\alpha \phi \,\mathrm{d}x = \int_\Omega u \cdot \big(\partial^\alpha(\eta\phi) - \phi \partial^\alpha \eta\big)\,\mathrm{d}x \\ &= -\int_\Omega \big(\partial^\alpha u \cdot \eta\phi + u \cdot \phi \partial^\alpha \eta\big)\,\mathrm{d}x = -\int_\Omega \big(\eta \partial^\alpha u + u \partial^\alpha \eta\big) \cdot \phi \,\mathrm{d}x,\end{aligned}$$

where in the third step we have used $\eta\phi \in \mathrm{C}_\mathrm{c}^\infty(\Omega)$. □

While reading, you might have already wondered if and when we are going to announce the following result.

Proposition 3.10. *Let Ω be connected and $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ be weakly differentiable. Then $\nabla u = 0$ in the weak sense if and only if there is a constant c such that $u = c$ a.e. in Ω.*

For the proof we need a version of Proposition 3.3 (b) for weak derivatives and functions that are only defined locally. On a first reading the reader should think of $\Omega = \mathbb{R}^n = \Omega_\varepsilon$ for all ε in the lemma below. Given $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ and $V \subseteq \Omega$, we always think of $\mathbf{1}_V u$ as being defined as zero on $\mathbb{R}^n \setminus V$.

Lemma 3.11. *Let $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$, let $\theta \in \mathrm{C}^\infty_{\mathrm{c}}(\mathbb{R}^n)$ have support in the unit ball and let $\alpha \in \mathbb{N}_0^n$. For $\varepsilon > 0$ define the open sets*

$$\Omega_\varepsilon := \{x \in \Omega \mid \operatorname{dist}(x, \partial\Omega) > \varepsilon\}.$$

If $\partial^\alpha u$ exists in the weak sense and $0 < t < \varepsilon$, then within $\Omega_{2\varepsilon}$ we have

$$\partial^\alpha(\theta_t * (\mathbf{1}_{\Omega_\varepsilon} u)) = \theta_t * (\mathbf{1}_{\Omega_\varepsilon} \partial^\alpha u) \ (= \partial^\alpha \theta_t * (\mathbf{1}_{\Omega_\varepsilon} u)).$$

Proof. The function $\theta_t * (\mathbf{1}_{\Omega_\varepsilon} u)$ is smooth with $\partial^\alpha(\theta_t * (\mathbf{1}_{\Omega_\varepsilon} u)) = (\partial^\alpha \theta_t) * (\mathbf{1}_{\Omega_\varepsilon} u)$, see Proposition 3.3 (b). Given $x \in \Omega_{2\varepsilon}$ and $t < \varepsilon$, we set $\phi(y) := \theta_t(x - y)$. Since θ_t has support in $B(0,t)$, we see that ϕ has support in $\Omega_{2\varepsilon - t} \subseteq \Omega_\varepsilon$. In particular, it is a test function for the weak derivative $\partial^\alpha u$ and we can compute

$$\begin{aligned}
\partial^\alpha\big(\theta_t * (\mathbf{1}_{\Omega_\varepsilon} u)\big)(x) &= \int_{\mathbb{R}^n} (\partial^\alpha \theta_t)(x - y) \cdot \mathbf{1}_{\Omega_\varepsilon}(y) u(y) \,\mathrm{d}y \\
&= \int_{\Omega_\varepsilon} (-1)^{|\alpha|} \partial^\alpha \phi(y) \cdot u(y) \,\mathrm{d}y \\
&= \int_{\Omega_\varepsilon} \phi(y) \cdot \partial^\alpha u(y) \,\mathrm{d}y \\
&= \int_{\mathbb{R}^n} \theta_t(x - y) \cdot \mathbf{1}_{\Omega_\varepsilon}(y) \partial^\alpha u(y) \,\mathrm{d}y \\
&= \theta_t * (\mathbf{1}_{\Omega_\varepsilon} \partial^\alpha u)(x).
\end{aligned}$$ □

Proof of Proposition 3.10. If $u = c$ a.e. in Ω, then $\nabla u = 0$ follows by integration by parts (or Remark 3.6).

Conversely, suppose $\nabla u = 0$ in the weak sense. We fix $\theta \in \mathrm{C}^\infty_{\mathrm{c}}(\mathbb{R}^n)$ with support in the unit ball and integral 1 and first work inside an arbitrary ball $B = B(x,r)$ with $B(x,3r) \subseteq \Omega$. Lemma 3.11 applied to $u \in \mathrm{L}^1_{\mathrm{loc}}(B(x,3r))$ with $\varepsilon = r$ yields

$$\partial_j(\theta_t * (\mathbf{1}_{B(x,2r)} u)) = \theta_t * (\mathbf{1}_{B(x,2r)} \partial_j u) = \theta_t * 0 = 0 \qquad (1 \le j \le n)$$

in B, provided that $t < r$. The upshot is that on the left-hand side we are computing a classical derivative of a smooth function! Hence, $\theta_t * (\mathbf{1}_{B(x,2r)} u)$ is constant in B. In the limit as $t \searrow 0$, these functions converge in $\mathrm{L}^1(\mathbb{R}^n)$ to $\mathbf{1}_{B(x,2r)} u$, see Proposition 3.3.

Since a convergent sequence in L^1 has an a.e. convergent subsequence, we conclude that $u = c(B)$ a.e. on B for some constant $c(B)$.

Now, we fix any ball B_0 as above and set $c := c(B_0)$. We introduce

$$M := \{x \in \Omega \mid \text{there is a ball } B(x,r) \subseteq \Omega \text{ on which } u = c \text{ a.e.}\}$$

and complete the proof by showing M is open and closed in Ω. Indeed, if this is true, then $M = \Omega$ by connectedness. Consequently, Ω can be covered by balls on which $u = c$ a.e. and since every open cover has a countable subcover, we get $u = c$ a.e. on Ω.

Since balls are open, M is open. For closedness, let $(x_j) \subseteq M$ be a sequence with limit $x \in \Omega$. By the first part $u = C$ is constant on some ball $B(x,r)$. Since eventually $x_j \in B(x,r)$, we must have $C = c$, thereby proving $x \in M$. □

In one dimension, there is even a fundamental theorem of calculus for weakly differentiable functions.

Theorem 3.12 (Fundamental theorem of calculus)**.** *Let (a,b) be a bounded interval and let $u \in \mathrm{L}^1((a,b))$. The following are equivalent:*

(a) u is weakly differentiable in (a,b) with $u' \in \mathrm{L}^1((a,b))$.

(b) There is some $f \in \mathrm{L}^1((a,b))$ and a constant C such that for a.e. $x \in (a,b)$ we have

$$u(x) = C + \int_a^x f(y)\,\mathrm{d}y. \tag{3.3}$$

In this case $u' = f$ and the right-hand side in (3.3) *is a (unique) representative of u of class* $\mathrm{C}([a,b])$.

Proof. First, assume (b) and define v by the right-hand side of (3.3). Since f is integrable, v is continuous on $[a,b]$ by dominated convergence. For any $\phi \in \mathrm{C}_\mathrm{c}^\infty((a,b))$ we have

$$\begin{aligned}\int_a^b v(x)\phi'(x)\,\mathrm{d}x &= C\int_a^b \phi'(x)\,\mathrm{d}x + \int_a^b \int_a^x f(y)\,\mathrm{d}y\,\phi'(x)\,\mathrm{d}x \\ &\overset{\text{Fubini}}{=} 0 + \int_a^b f(y)\int_y^b \phi'(x)\,\mathrm{d}x\,\mathrm{d}y = -\int_a^b f(y)\phi(y)\,\mathrm{d}y,\end{aligned}$$

that is, $v' = f$ in the weak sense.

Conversely, suppose that u is weakly differentiable with $u' \in \mathrm{L}^1((a,b))$ and define $v(x) := \int_a^x u'(y)\,\mathrm{d}y$. We have $v' = u'$ by the first part of the proof and Proposition 3.10 yields a constant C such that $u - v = C$ a.e. on (a,b). □

Remark 3.13. Functions u of the form (3.3) are called *absolutely continuous*. Theorem 3.12 is the deeper reason why differentiating the sign function in Example 3.8 had to go wrong.

As corollary we obtain a first example of an embedding theorem for a space of weakly differentiable functions, later called *Sobolev embedding*.

Corollary 3.14. *Let (a,b) be a bounded interval and let $u \in \mathrm{L}^1((a,b))$ have a weak derivative $u' \in \mathrm{L}^1((a,b))$. Then u has a unique representative $\widetilde{u} \in \mathrm{C}([a,b])$ and*

$$\|\widetilde{u}\|_{\mathrm{C}([a,b])} \leq (b-a)^{-1}\|u\|_{\mathrm{L}^1((a,b))} + \|u'\|_{\mathrm{L}^1((a,b))}.$$

Proof. The continuous representative is given by the right-hand side of (3.3) with $f = u'$. Hence, we have for all $x, y \in [a,b]$ that[3]

$$|\widetilde{u}(x)| = \left|\widetilde{u}(y) + \int_y^x u'(s)\,\mathrm{d}s\right| \leq |\widetilde{u}(y)| + \|u'\|_{\mathrm{L}^1((a,b))}.$$

We want to bound $\widetilde{u}(x)$ uniformly in x, so y is a free variable that we can get rid of by averaging (i.e., integrating with respect to $\frac{\mathrm{d}y}{b-a}$):

$$|\widetilde{u}(x)| \leq \int_a^b \Big(|\widetilde{u}(y)| + \|u'\|_{\mathrm{L}^1((a,b))}\Big)\frac{\mathrm{d}y}{b-a} = (b-a)^{-1}\|u\|_{\mathrm{L}^1((a,b))} + \|u'\|_{\mathrm{L}^1((a,b))}.$$

This bound holds for every $x \in [a,b]$ and the proof is complete. □

3.3. The Sobolev spaces $\mathrm{H}^k(\Omega)$

We introduce Hilbert spaces of differentiable functions by using weak derivatives that we measure in L^2-norm.

Definition 3.15. *Let $k \in \mathbb{N}$. The k-th order* Sobolev space *is defined as*

$$\mathrm{H}^k(\Omega) := \Big\{u \in \mathrm{L}^2(\Omega) \mid \partial^\alpha u \in \mathrm{L}^2(\Omega) \text{ for all } \alpha \in \mathbb{N}_0^n \text{ with } |\alpha| \leq k\Big\},$$

equipped with the inner product

$$\langle u, v\rangle_{\mathrm{H}^k(\Omega)} := \sum_{|\alpha| \leq k} \langle \partial^\alpha u, \partial^\alpha v\rangle_{\mathrm{L}^2(\Omega)}\,.$$

Theorem 3.16. $\mathrm{H}^k(\Omega)$ *is a separable Hilbert space for every $k \in \mathbb{N}$.*

[3] For $x < y$ we use the convention that $\int_y^x := -\int_x^y := -\int_{[x,y]}$.

Proof. We see immediately that $\mathrm{H}^k(\Omega)$ is a pre-Hilbert space (definiteness follows by looking at the term for $\alpha = 0$).

Let (u_j) be a Cauchy sequence in $\mathrm{H}^k(\Omega)$. For any multi-index with $|\alpha| \le k$ we see by definition of the $\mathrm{H}^k(\Omega)$-norm that $(\partial^\alpha u_j)$ is a Cauchy sequence in $\mathrm{L}^2(\Omega)$, hence by completeness it has a limit in $\mathrm{L}^2(\Omega)$ that we call u_α. Let us set $u := u_0$. For any $\phi \in \mathrm{C}_\mathrm{c}^\infty(\Omega)$ we have

$$\begin{aligned}\langle u, \partial^\alpha \phi\rangle_{\mathrm{L}^2(\Omega)} &= \lim_{j\to\infty} \left\langle u_j, \partial^\alpha \phi\right\rangle_{\mathrm{L}^2(\Omega)} \\ &= \lim_{j\to\infty} (-1)^{|\alpha|} \left\langle \partial^\alpha u_j, \phi\right\rangle_{\mathrm{L}^2(\Omega)} = (-1)^{|\alpha|} \langle u_\alpha, \phi\rangle_{\mathrm{L}^2(\Omega)} .\end{aligned}$$

By definition, this means $\partial^\alpha u = u_\alpha$ in the weak sense and therefore $u \in \mathrm{H}^k(\Omega)$ and $u_j \to u$ in $\mathrm{H}^k(\Omega)$ in the limit as $j \to \infty$.

Finally, we consider the isometry

$$\Phi\colon \mathrm{H}^k(\Omega) \to \bigtimes_{|\alpha|\le k} \mathrm{L}^2(\Omega), \quad u \mapsto (\partial^\alpha u)_{|\alpha|\le k}.$$

Since $\mathrm{H}^k(\Omega)$ is complete, the image of Φ is closed and hence separable as being a closed subspace of a separable space. Therefore $\mathrm{H}^k(\Omega)$ is separable, too. □

Warning 3.17. While it is true that every $u \in \mathrm{C}^k(\Omega)$ has weak derivatives up to order k, this does not mean that $\mathrm{C}^k(\Omega) \subseteq \mathrm{H}^k(\Omega)$ — for instance, a non-zero constant function u is not contained in $\mathrm{H}^1(\mathbb{R}^n)$, because it misses the integrability condition $u \in \mathrm{L}^2(\mathbb{R}^n)$.

Test functions are dense in $\mathrm{H}^k(\mathbb{R}^n)$, but we will learn in the next section that — unlike for L^2-spaces — this property is very specific to working on the whole space.

Proposition 3.18. $\mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$ *is dense in* $\mathrm{H}^k(\mathbb{R}^n)$ *for every* $k \in \mathbb{N}$.

Proof. Let $u \in \mathrm{H}^k(\mathbb{R}^n)$ and $\varepsilon > 0$. We approximate u in two steps.

Step 1: We find $v \in \mathrm{C}^\infty(\mathbb{R}^n) \cap \mathrm{H}^k(\mathbb{R}^n)$ *with* $\|u - v\|_{\mathrm{H}^k(\mathbb{R}^n)} \le \varepsilon$.

We use a mollifier based on $\theta \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$ with integral 1 and support in the unit ball. We set $v := \theta_t * u$ with $t > 0$ yet to be chosen. Lemma 3.11 yields $\partial^\alpha v = \theta_t * \partial^\alpha u$ for all multi-indices with $|\alpha| \le k$ and Proposition 3.3 (c) tells us that v has the required property provided we take $t > 0$ sufficiently small.

Step 2: We find $w \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$ *with* $\|v - w\|_{\mathrm{H}^k(\mathbb{R}^n)} \le \varepsilon$.

We take $\eta \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$ with $\eta(0) = 1$ and consider the *smooth truncations* $w(x) := \eta(tx)v(x)$ with $t > 0$ yet to be chosen. For any multi-index with $|\alpha| \le k$ we have

$$\partial^\alpha w(x) = \eta(tx)\partial^\alpha v(x) + \sum_{0<\beta\le\alpha} \Big(\tbinom{\alpha}{\beta} t^{|\beta|} (\partial^\beta \eta)(tx)\Big) \partial^{\alpha-\beta} v(x).$$

Let us investigate what happens in $\mathrm{L}^2(\mathbb{R}^n)$ when we pass to the limit as $t \to 0$. The first term on the right tends to $\partial^\alpha v$ by dominated convergence. In the sum over β the prefactors of $\partial^{\alpha-\beta} v(x)$ are uniformly bounded and come with a positive power of t. Hence, the full sum tends to 0 and we are done. □

We end this section with a useful integration by parts formula.

Corollary 3.19. *Let $k \in \mathbb{N}$ and $u, v \in \mathrm{H}^k(\mathbb{R}^n)$. Then for every multi-index $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq k$ we have*

$$\int_{\mathbb{R}^n} \partial^\alpha u \cdot v \,\mathrm{d}x = (-1)^{|\alpha|} \int_{\mathbb{R}^n} u \cdot \partial^\alpha v \,\mathrm{d}x.$$

Proof. If $v \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$, then this holds by definition of the weak derivative $\partial^\alpha u$. The identity extends to $v \in \mathrm{H}^k(\mathbb{R}^n)$ by density, see Proposition 3.18. □

3.4. The Sobolev space $\mathrm{H}_0^1(\Omega)$

Eventually, we want to use Sobolev spaces in order to solve boundary value problems and in particular, given a Sobolev function $u \in \mathrm{H}^1(\Omega)$, we want to give a meaning to the *Dirichlet boundary condition* '$u = 0$ on $\partial\Omega$'.

In the one-dimensional case we can use Corollary 3.14 and the continuous inclusion $\mathrm{L}^2((a,b)) \subseteq \mathrm{L}^1((a,b))$ on bounded intervals in order to proceed as follows.

Definition 3.20. *Let $(a,b) \subset \mathbb{R}$ be a bounded interval. Define a bounded* trace map

$$\mathrm{tr}\colon \mathrm{H}^1((a,b)) \to \mathbb{C}^2, \quad u \mapsto (\widetilde{u}(a), \widetilde{u}(b)),$$

where $\widetilde{u} \in \mathrm{C}([a,b])$ is the unique continuous representative of u.

Functions with trace zero can be characterized as follows. The proof is left as Exercise 3.2.

Theorem 3.21. *Let $(a,b) \subseteq \mathbb{R}$ be a bounded interval and $u \in \mathrm{H}^1((a,b))$. Then $\mathrm{tr}(u) = 0$ if and only if u is contained in the $\mathrm{H}^1((a,b))$-closure of $\mathrm{C}_\mathrm{c}^\infty((a,b))$.*

There is no obvious way how to define a meaningful trace in higher dimensions and in fact it is impossible in general, see Exercise 3.4. Characterizing sets with a reasonable trace operation quickly becomes a hard problem in potential theory [AH96, Chapter 6]. Inspired by Theorem 3.21 we do not define traces in higher dimensions, but we *postulate* the subset of functions with zero boundary values.

Definition 3.22. *We define $\mathrm{H}_0^1(\Omega)$ as the closure of $\mathrm{C}_\mathrm{c}^\infty(\Omega)$ in $\mathrm{H}^1(\Omega)$ and equip it with the $\mathrm{H}^1(\Omega)$-norm. We say that a function $u \in \mathrm{H}^1(\Omega)$* vanishes on $\partial\Omega$ in the H^1-sense *if $u \in \mathrm{H}_0^1(\Omega)$.*

On $\Omega = \mathbb{R}^n$ we have shown $\mathrm{H}^1(\mathbb{R}^n) = \mathrm{H}^1_0(\mathbb{R}^n)$ in Proposition 3.18, which is no surprise since $\mathbb{R}^n$ has empty boundary. However, this is not the only example: In Lecture 10 you will learn that $\mathrm{H}^1(\Omega) = \mathrm{H}^1_0(\Omega)$ for $\Omega = \mathbb{R}^n \setminus \{0\}$ in dimension $n \geq 2$. The interpretation is that this set has a boundary that is 'too small to be seen' by Sobolev functions. However, if Ω has finite measure, then the two spaces always fall apart as the next lemma shows. Its proof is left for you as Exercise 3.3.

Lemma 3.23. *If $u \in \mathrm{H}^1_0(\Omega)$ satisfies $\nabla u = 0$, then already $u = 0$. In particular, $\mathrm{H}^1_0(\Omega)$ is a proper subspace of $\mathrm{H}^1(\Omega)$ if Ω has finite measure.*

A more quantitative question aiming in a similar direction is whether the norm of ∇u controls (up to multiplicative constants) the norm of u. Such an estimate holds for instance if Ω has the following property.

Definition 3.24. *We say that Ω is* contained in a strip *(of height h) if there exists a direction j and some $a \in \mathbb{R}$ such that $x_j \in [a, a+h]$ for every $x \in \Omega$.*

Proposition 3.25 (Poincaré inequality)**.** *If Ω is contained in a strip of height h, then*

$$\|u\|_{\mathrm{L}^2(\Omega)} \leq \frac{h}{\sqrt{2}} \|\nabla u\|_{\mathrm{L}^2(\Omega)} \qquad (u \in \mathrm{H}^1_0(\Omega)).$$

Proof. Since both sides of the inequality depend continuously on $u \in \mathrm{H}^1_0(\Omega)$, it suffices to prove the estimate for $u \in \mathrm{C}^\infty_\mathrm{c}(\Omega)$. We write points in $\mathbb{R}^n$ as $x = (x_1, x')$, where $x' \in \mathbb{R}^{n-1}$ and x' should be ignored when $n = 1$, and think of u being extended by 0 outside of Ω. After translation and relabeling of coordinates we can assume $x_1 \in [0, h]$ for every $x \in \Omega$. Thanks to $u(0, x') = 0$, we can use the fundamental theorem of calculus and the Cauchy–Schwarz inequality to estimate

$$|u(x_1, x')|^2 = \left| \int_0^{x_1} \partial_1 u(s, x') \, \mathrm{d}s \right|^2 \leq x_1 \int_0^h |\partial_1 u(s, x')|^2 \, \mathrm{d}s.$$

Integrating both sides in (x_1, x') gives the result

$$\|u\|^2_{\mathrm{L}^2(\Omega)} \leq \left(\int_0^h x_1 \, \mathrm{d}x_1 \right) \|\partial_1 u\|^2_{\mathrm{L}^2(\Omega)} \leq \frac{h^2}{2} \|\nabla u\|^2_{\mathrm{L}^2(\Omega)}. \qquad \square$$

3.5. Elliptic operators in divergence form

In this section we introduce the main actors of the lecture series. Throughout, $A \colon \Omega \to \mathcal{L}(\mathbb{C}^n)$ is a measurable, essentially bounded, matrix-valued function and we set

$$\Lambda := \operatorname{esssup}_{x \in \Omega} \|A(x)\|_{\mathcal{L}(\mathbb{C}^n)} < \infty. \tag{3.4}$$

The equations and boundary value problems that we are going to consider take the formal form

$$-\operatorname{div}(A\nabla u) = f \qquad \text{in } \Omega, \tag{3.5}$$
$$u = 0 \qquad \text{on } \partial\Omega. \tag{3.6}$$

If $A(x) = \mathrm{id}_{\mathbb{C}^n}$ is the identity matrix for every $x \in \Omega$, then (3.5) is Poisson's equation $-\Delta u = f$. At this point, variable coefficients A do not pose any additional complication for the theory. We will come back to their meaning and use in some of the later lectures. The Hilbert space approach to the boundary value problem via the form method from Lecture 2 starts by interpreting the Dirichlet condition (3.6) in the H^1-sense (Definition 3.22). Hence, we naturally work with $u \in \mathrm{H}_0^1(\Omega)$. We set

$$V := \mathrm{H}_0^1(\Omega),$$
$$H := \mathrm{L}^2(\Omega),$$

and note that $V \subseteq H$ with dense inclusion since $\mathrm{C}_c^\infty(\Omega)$ is dense in $\mathrm{L}^2(\Omega)$. In this setting it is common to write

$$V' =: \mathrm{H}^{-1}(\Omega).$$

We define the sesquilinear form

$$a\colon \mathrm{H}_0^1(\Omega) \times \mathrm{H}_0^1(\Omega) \to \mathbb{C}, \quad a(u,v) = \int_\Omega A\nabla u \cdot \overline{\nabla v}\, \mathrm{d}x \tag{3.7}$$

and by assumption (3.4) and the Cauchy–Schwarz inequality we obtain that a is bounded:

$$|a(u,v)| \leq \Lambda \|\nabla u\|_{\mathrm{L}^2(\Omega)} \|\nabla v\|_{\mathrm{L}^2(\Omega)} \leq \Lambda \|u\|_{\mathrm{H}_0^1(\Omega)} \|v\|_{\mathrm{H}_0^1(\Omega)} \qquad (u, v \in \mathrm{H}_0^1(\Omega)). \tag{3.8}$$

Hence, we can associate with a a bounded operator $\mathscr{L}\colon \mathrm{H}_0^1(\Omega) \to \mathrm{H}^{-1}(\Omega)$ and an unbounded operator L in $\mathrm{L}^2(\Omega)$, see Definitions 2.15 and 2.16. By definition, we have $Lu = f$ if and only if $f \in \mathrm{L}^2(\Omega)$, $u \in \mathrm{H}_0^1(\Omega)$ and

$$\int_\Omega A\nabla u \cdot \overline{\nabla v}\, \mathrm{d}x = \int_\Omega f \cdot \overline{v}\, \mathrm{d}x \qquad (v \in \mathrm{H}_0^1(\Omega)).$$

We see that (3.5) is hidden in the equation $Lu = f$ up to undoing one integration by parts that cannot be justified in general. But if $A\nabla u$ was weakly differentiable (componentwise), then indeed we would have $f = -\operatorname{div}(A\nabla u)$ in the weak sense.

We have seen in the previous lecture that all further theory for L rests on some sort of lower bound assumption for the form a. In our concrete setting it is most natural to work with positive definiteness assumptions directly on the coefficients A.

Definition 3.26. *The coefficients A are called* elliptic *if there exists $\lambda > 0$ such that for a.e. $x \in \Omega$ we have for all $\xi \in \mathbb{C}^n$ that*

$$\operatorname{Re}(A(x)\xi \cdot \overline{\xi}) \geq \lambda |\xi|^2. \tag{3.9}$$

In this case we call L an elliptic operator in divergence form *with Dirichlet boundary conditions on Ω and we write $L = -\operatorname{div}(A\nabla\,\cdot)$ if the context is clear. The operator associated with $A = \mathrm{id}_{\mathbb{C}^n}$ is the* negative Dirichlet Laplacian *on Ω, denoted by $(-\Delta)_{\mathrm{H}_0^1(\Omega)}$.*

This terminology is consistent with abstract form theory because of the following lemma.

Lemma 3.27. *If A is elliptic, then a defined in* (3.7) *is bounded and elliptic. Moreover, a is sectorial of angle* $\arctan(\Lambda/\lambda)$.

Proof. Let $u \in \mathrm{H}_0^1(\Omega)$. We have seen boundedness in (3.8). As for ellipticity, (3.9) yields

$$\operatorname{Re}\bigl(a(u)\bigr) = \int_\Omega \operatorname{Re}\bigl((A(x)\nabla u(x) \cdot \overline{\nabla u(x)}\bigr)\,\mathrm{d}x \geq \int_\Omega \lambda |\nabla u(x)|^2\,\mathrm{d}x = \lambda \|\nabla u\|^2_{\mathrm{L}^2(\Omega)}.$$

Thus, we get at once

$$\operatorname{Re}\bigl(a(u)\bigr) + \lambda\,\langle u, u\rangle_{\mathrm{L}^2(\Omega)} \geq \lambda \|u\|^2_{\mathrm{H}_0^1(\Omega)}$$

and

$$|\operatorname{Im}\bigl(a(u)\bigr)| \leq |a(u)| \overset{(3.4)}{\leq} \Lambda \|\nabla u\|^2_{\mathrm{L}^2(\Omega)} \leq \frac{\Lambda}{\lambda} \operatorname{Re}\bigl(a(u)\bigr). \qquad \square$$

Remark 3.28. The angle in Lemma 3.27 can be improved in terms of λ and Λ, see Exercise 3.6.

In the following theorem we summarize the operator theoretic properties of divergence form operators that we can prove so far.

Theorem 3.29. *An elliptic operator $L = -\operatorname{div}(A\nabla\,\cdot)$ in divergence form with Dirichlet boundary conditions on Ω has the following properties:*

(a) It is sectorial of angle $\varphi_L < \pi/2$.

(b) It is injective.

(c) Its adjoint is $L^* = -\operatorname{div}(A^*\nabla\,\cdot)$. In particular, L is self-adjoint if $A(x)$ is Hermitean for a.e. $x \in \Omega$.

(d) It is invertible if Ω is contained in a strip.

Proof. Part (a) is a consequence of Lemma 3.27 and Theorem 2.28 and (c) follows from Proposition 2.24 since

$$a^*(u,v) = \overline{a(v,u)} = \overline{\langle A\nabla v, \nabla u\rangle_{\mathrm{L}^2(\Omega)}} = \overline{\int_\Omega \big\langle A(x)\nabla v(x), \nabla u(x)\big\rangle_{\mathbb{C}^n}\,\mathrm{d}x}$$
$$= \overline{\int_\Omega \big\langle \nabla v(x), A(x)^*\nabla u(x)\big\rangle_{\mathbb{C}^n}\,\mathrm{d}x} = \overline{\langle \nabla v, A^*\nabla u\rangle_{\mathrm{L}^2(\Omega)}} = \langle A^*\nabla u, \nabla v\rangle_{\mathrm{L}^2(\Omega)}$$

holds for all $u, v \in V$. The other two assertions are left as Exercise 3.5. □

3.6. Exercises

Exercise 3.1. Let $(a,b) \subseteq \mathbb{R}$ be a bounded interval and let $u \in \mathrm{H}^1((a,b))$. Prove that the continuous representative $\widetilde{u}$ satisfies the Hölder estimate

$$|\widetilde{u}(x) - \widetilde{u}(y)| \le |x-y|^{1/2}\|u'\|_{\mathrm{L}^2((a,b))} \qquad (x, y \in [a,b]).$$

Exercise 3.2 (Trace zero functions on a bounded interval)**.** We suggest a proof of Theorem 3.21 via the following steps.

(a) Convince yourself that $u \in \mathrm{C}_\mathrm{c}^\infty((a,b))$ implies $\mathrm{tr}(u) = 0$.

For the converse, we assume from now on that $u \in \mathrm{H}^1((a,b))$ satisfies $\mathrm{tr}(u) = 0$.

(b) Let $0 < \varepsilon < (b-a)/4$. Construct a function $\eta_\varepsilon \in \mathrm{C}_\mathrm{c}^\infty((a,b))$ such that

- $\eta_\varepsilon = 1$ on $(a + 2\varepsilon, b - 2\varepsilon)$ and $\eta_\varepsilon = 0$ on $(a,b) \setminus (a+\varepsilon, b-\varepsilon)$,
- $\|\eta_\varepsilon\|_{\mathrm{L}^\infty((a,b))} + \varepsilon\|(\eta_\varepsilon)'\|_{\mathrm{L}^\infty((a,b))} \le C$ for a constant C independent of ε.

Hint: Start with the indicator function of $(a + 3\varepsilon/2, b - 3\varepsilon/2)$ and make it smooth.

(c) Prove that $\eta_\varepsilon u \in \mathrm{H}_0^1((a,b))$.

(d) Prove that $\eta_\varepsilon u \to u$ in $\mathrm{H}^1((a,b))$ in the limit as $\varepsilon \to 0$. Conclude.

Exercise 3.3. Prove Lemma 3.23.

Hint: Suppose that $u \in \mathrm{H}_0^1(\Omega)$ satisfies $\nabla u = 0$ and study its extension to $\mathbb{R}^n \setminus \Omega$ by zero.

Exercise 3.4 (A naughty H^1-function)**.** In this exercise we are going to explore how badly the embedding in Corollary 3.14 fails in higher dimensions. We work on the unit ball $B := B(0,1)$ in dimension $n \ge 3$ and proceed as follows.

(a) Let $\alpha > 0$. On $\mathbb{R}^n \setminus \{0\}$ define $u(x) := |x|^{-\alpha}$ and let $v_1, \ldots, v_n$ be its classical partial derivatives. Show that α can be chosen such that these $(n+1)$ functions belong to $\mathrm{L}^2(B)$.

(b) Prove $\nabla u = (v_1, \dots, v_n)$ in the weak sense on B.

Hint: Let $\varepsilon \in (0,1)$ and apply the Gauß–Green theorem on $B \setminus B(0,\varepsilon)$. Then pass to the limit as $\varepsilon \to 0$.

The example above shows that $u \in \mathrm{H}^1(B)$ does not imply $\operatorname{esssup}_B |u| < \infty$. Let us make the situation even worse.

(c) Let $q_1, q_2, \dots$ be a countable dense subset of B and define $u_j(x) := u(x - q_j)$, a.e. on B. Check that there is a constant C such that $\|u_j\|_{\mathrm{H}^1(B)} \leq C$ for all j.

(d) Use the u_j to construct some $w \in \mathrm{H}^1(B)$ with the property that $\operatorname{esssup}_\Omega |w| = \infty$ holds for *every* non-empty open set $\Omega \subseteq B$.

Exercise 3.5. In this exercise you will complete the proof of Theorem 3.29.

(a) Show that

$$\ker(L) = \{u \in \mathrm{H}^1_0(\Omega) \mid \nabla u = 0\}.$$

(b) Conclude that L is injective.

(c) Suppose that in addition Ω is contained in a strip. Prove that $\mathscr{L}$ and L are invertible.

Exercise 3.6. In the setup of Section 3.5 prove that a is sectorial of angle $\arccos(\lambda/\Lambda)$, which is a better bound than in Lemma 3.27.

Hint: Draw the values of $a(u)$ for $u \in \mathrm{H}^1_0(\Omega)$ with $\|\nabla u\|_{\mathrm{L}^2(\Omega)} = 1$ in the complex plane.

4. Fourier analysis and the Laplacian

The Fourier transform is the fundamental tool of harmonic analysis. It provides an isometry on $\mathrm{L}^2(\mathbb{R}^n)$ that translates derivatives into multiplication by polynomials and vice versa, allowing to transform questions about differentiability and differential operators into algebraic properties. In return, this will open a new perspective on Sobolev spaces and the negative Laplacian on $\mathbb{R}^n$ that will allow us to generalize the concept of smoothness to non-integer parameters. In particular, we will see

- fractional powers of the negative Laplacian, for example $(-\Delta)^{1/4}$,
- fractional Sobolev spaces, for example $\mathrm{H}^{1/2}(\mathbb{R}^n)$,

and study their relationship. We will also see a first instance of a so-called functional calculus. Generalizing the emerging concepts to elliptic operators other than the negative Laplacian will be our guiding principle for the upcoming lectures.

4.1. The Fourier transform

We assume familiarity with the basic properties of the Fourier transform. For convenience and later reference we begin by summarizing the toolkit that will be needed in our course. For proofs and further background we refer, for instance, to [Gra14] and the exercises.

Definition 4.1. *The* Fourier transform *of $u \in \mathrm{L}^1(\mathbb{R}^n)$ is defined by*

$$\mathcal{F}u(\xi) := \widehat{u}(\xi) := \int_{\mathbb{R}^n} \mathrm{e}^{-2\pi \mathrm{i} x \cdot \xi} u(x) \,\mathrm{d}x \qquad (\xi \in \mathbb{R}^n)$$

and the inverse Fourier transform *by*

$$\mathcal{F}^{-1}u(x) := \check{u}(x) := \int_{\mathbb{R}^n} \mathrm{e}^{2\pi \mathrm{i} \xi \cdot x} u(\xi) \,\mathrm{d}\xi \qquad (x \in \mathbb{R}^n).$$

A particularly useful space when working with the Fourier transform is the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ of rapidly decaying smooth functions. We recall the multi-index notation $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, whenever $x \in \mathbb{R}^n$ and $\alpha \in \mathbb{N}_0^n$, and that $u \in \mathcal{S}(\mathbb{R}^n)$ if and only if $\sup_{x \in \mathbb{R}^n} |x^\alpha \partial^\beta u(x)|$ is finite for all $\alpha, \beta \in \mathbb{N}_0^n$.

Proposition 4.2 (Toolkit for the Fourier transform)**.** *Let $u, v \in \mathcal{S}(\mathbb{R}^n)$, $y \in \mathbb{R}^n$, $\alpha \in \mathbb{N}_0^n$ and $t > 0$.*

(a) $\mathcal{F}$ and $\mathcal{F}^{-1}$ map the Schwartz space into itself and are inverse of one another.

(b) We have $\widehat{\partial^\alpha u}(\xi) = (2\pi \mathrm{i}\xi)^\alpha \widehat{u}(\xi)$.

(c) Define the dilation $\delta_t u(x) \coloneqq u(tx)$ *and the* translation $\tau_y u(x) \coloneqq u(x-y)$*. Their Fourier transforms are given by*

$$\widehat{\delta_t u}(\xi) = t^{-n}\delta_{t^{-1}}\widehat{u}(\xi) = \widehat{u}_t(\xi) \quad \text{and} \quad \widehat{\tau_y u}(\xi) = \mathrm{e}^{-2\pi \mathrm{i} y\cdot\xi}\widehat{u}(\xi),$$

where we also use the mollifier notation from Definition 3.2.

(d) We have $\widehat{u * v} = \widehat{u} \cdot \widehat{v}$.

(e) Parseval's formula $\langle u, v\rangle_{\mathrm{L}^2(\mathbb{R}^n)} = \langle \widehat{u}, \widehat{v}\rangle_{\mathrm{L}^2(\mathbb{R}^n)}$ *holds.*

(f) We have Plancherel's identity $\|u\|_{\mathrm{L}^2(\mathbb{R}^n)} = \|\widehat{u}\|_{\mathrm{L}^2(\mathbb{R}^n)} = \|\check{u}\|_{\mathrm{L}^2(\mathbb{R}^n)}$.

Remark 4.3. Rules similar to (b), (c) and (d) also hold for the inverse Fourier transform, because we have $\check{u}(y) = \widehat{u}(-y)$ by definition.

Remark 4.4. Since $\mathcal{S}(\mathbb{R}^n)$ is dense in $\mathrm{L}^2(\mathbb{R}^n)$ (it already contains $\mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$), we obtain from (a) and (f) that $\mathcal{F}$ uniquely extends to a bijective isometry on $\mathrm{L}^2(\mathbb{R}^n)$. We use the same symbols for the extension. This observation is usually called *Plancherel's theorem*.

By density, formulas (c), (e), (f) in Proposition 4.2 remain valid for all $u, v \in \mathrm{L}^2(\mathbb{R}^n)$. In (d) we can allow $u \in \mathrm{L}^1(\mathbb{R}^n)$, $v \in \mathrm{L}^2(\mathbb{R}^n)$ and we ask you to provide the details in Exercise 4.1. In order to generalize (b), we have to work with functions that admit derivatives in $\mathrm{L}^2(\mathbb{R}^n)$. This will naturally lead to a different perspective on Sobolev spaces in the next section.

4.2. The domain of the negative Laplacian on $\mathbb{R}^n$

We begin with a Fourier analytic characterization of Sobolev spaces.

Proposition 4.5. *For $k \in \mathbb{N}$ we have that*

$$\mathrm{H}^k(\mathbb{R}^n) = \Big\{u \in \mathrm{L}^2(\mathbb{R}^n) \mid (\xi \mapsto |\xi|^k \widehat{u}(\xi)) \in \mathrm{L}^2(\mathbb{R}^n)\Big\}$$

with equivalence of norms

$$\|u\|_{\mathrm{H}^k(\mathbb{R}^n)} \simeq \|u\|_{\mathrm{L}^2(\mathbb{R}^n)} + \big\|\xi \mapsto |\xi|^k \widehat{u}(\xi)\big\|_{\mathrm{L}^2(\mathbb{R}^n)} \qquad (u \in \mathrm{H}^k(\mathbb{R}^n)).$$

Moreover, if $u \in \mathrm{H}^k(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq k$, then $\widehat{\partial^\alpha u}(\xi) = (2\pi\mathrm{i}\xi)^\alpha\widehat{u}(\xi)$ for almost every $\xi \in \mathbb{R}^n$.

Proof. Given $u \in \mathrm{L}^2(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}_0^n$, Parseval's formula yields for all $v \in \mathcal{S}(\mathbb{R}^n)$ that

$$\begin{aligned}(-1)^{|\alpha|} \int_{\mathbb{R}^n} u(x) \cdot \overline{\partial^\alpha v(x)}\, \mathrm{d}x &= (-1)^{|\alpha|} \int_{\mathbb{R}^n} \widehat{u}(\xi) \cdot \overline{\widehat{\partial^\alpha v}(\xi)}\, \mathrm{d}\xi \\ &= \int_{\mathbb{R}^n} (2\pi \mathrm{i} \xi)^\alpha \widehat{u}(\xi) \cdot \overline{\widehat{v}(\xi)}\, \mathrm{d}\xi.\end{aligned} \tag{4.1}$$

'$\subseteq$': *We assume* $u \in \mathrm{H}^k(\mathbb{R}^n)$.

Let $|\alpha| \leq k$. We can integrate by parts on the left-hand side of (4.1), see Corollary 3.19, and use Parseval's formula once more, to obtain

$$\int_{\mathbb{R}^n} (2\pi \mathrm{i} \xi)^\alpha \widehat{u}(\xi) \cdot \overline{\widehat{v}(\xi)}\, \mathrm{d}\xi = \int_{\mathbb{R}^n} \partial^\alpha u(x) \cdot \overline{v(x)}\, \mathrm{d}x = \int_{\mathbb{R}^n} \widehat{\partial^\alpha u}(\xi) \cdot \overline{\widehat{v}(\xi)}\, \mathrm{d}\xi.$$

Given any $\varphi \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$, we can take $v = \check{\overline{\varphi}}$, so that $\overline{\widehat{v}} = \varphi$. The fundamental lemma in the calculus of variations (Lemma 3.4) implies $\widehat{\partial^\alpha u}(\xi) = (2\pi \mathrm{i}\xi)^\alpha \widehat{u}(\xi)$ for a.e. $\xi \in \mathbb{R}^n$. With these Fourier formulæ for the weak derivatives at hand, we obtain

$$\|u\|_{\mathrm{H}^k(\mathbb{R}^n)}^2 = \sum_{|\alpha| \leq k} \int_{\mathbb{R}^n} \big|\widehat{\partial^\alpha u}(\xi)\big|^2\, \mathrm{d}\xi = \int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 \sum_{|\alpha| \leq k} |(2\pi\xi)^\alpha|^2\, \mathrm{d}\xi. \tag{4.2}$$

The powers of ξ can be estimated from below by looking at the terms for $\alpha = 0$ and $\alpha = k e_j$, where $e_1, \ldots, e_n$ are the standard unit vectors:

$$\sum_{|\alpha| \leq k} |(2\pi\xi)^\alpha|^2 \geq 1 + \sum_{j=1}^n |\xi_j|^{2k} \geq 1 + n^{1-k} \Big(\sum_{j=1}^n |\xi_j|^2 \Big)^k \geq n^{1-k}(1 + |\xi|^{2k}),$$

where in the second step we have used Hölder's inequality $(A_1 + \ldots + A_n)^k \leq n^{k-1}(A_1^k + \ldots + A_n^k)$ for the non-negative numbers $A_j = |\xi_j|^2$. In total, we have shown the required estimate

$$\|u\|_{\mathrm{L}^2(\mathbb{R}^n)}^2 + \big\|\xi \mapsto |\xi|^k \widehat{u}(\xi)\big\|_{\mathrm{L}^2(\mathbb{R}^n)}^2 = \int_{\mathbb{R}^n} \big(1 + |\xi|^{2k}\big) \big|\widehat{u}(\xi)\big|^2\, \mathrm{d}\xi \leq n^{k-1} \|u\|_{\mathrm{H}^k(\mathbb{R}^n)}^2.$$

'$\supseteq$': *We assume* $(\xi \mapsto |\xi|^k \widehat{u}(\xi)) \in \mathrm{L}^2(\mathbb{R}^n)$.

Let again $|\alpha| \leq k$. As

$$\big|(2\pi \mathrm{i} \xi)^\alpha\big| \leq (2\pi)^{|\alpha|} |\xi|^{|\alpha|} \leq (2\pi)^{|\alpha|}(1 + |\xi|^k) \qquad (\xi \in \mathbb{R}^n), \tag{4.3}$$

we can define a function $u_\alpha \in \mathrm{L}^2(\mathbb{R}^n)$ on the Fourier-side via $\widehat{u_\alpha}(\xi) := (2\pi \mathrm{i}\xi)^\alpha \widehat{u}(\xi)$. Applying Parseval's formula on the right-hand side of (4.1) leads us to

$$\int_{\mathbb{R}^n} u(x) \cdot \overline{\partial^\alpha v(x)}\, \mathrm{d}x = (-1)^{|\alpha|} \int_{\mathbb{R}^n} u_\alpha(x) \cdot \overline{v(x)}\, \mathrm{d}x$$

for any $v \in \mathcal{S}(\mathbb{R}^n)$. Thus, we have $u_\alpha = \partial^\alpha u$ in the weak sense and, consequently, $u \in \mathrm{H}^k(\mathbb{R}^n)$. By an estimate similar to (4.3), we obtain from (4.2) the reverse norm estimate

$$\|u\|^2_{\mathrm{H}^k(\mathbb{R}^n)} \leq C\big(\|u\|^2_{\mathrm{L}^2(\mathbb{R}^n)} + \big\|\xi \mapsto |\xi|^k \widehat{u}(\xi)\big\|^2_{\mathrm{L}^2(\mathbb{R}^n)}\big). \qquad \square$$

We use this proposition to identify the domain of the negative Dirichlet Laplacian on $\mathbb{R}^n$ that was introduced in Definition 3.26 via a sesquilinear form on $\mathrm{H}^1_0(\mathbb{R}^n)$. Since $\mathbb{R}^n$ has no boundary and as $\mathrm{H}^1_0(\mathbb{R}^n) = \mathrm{H}^1(\mathbb{R}^n)$ by Proposition 3.18, we find it less peculiar to call this operator 'negative Laplacian' and, for the moment, denote it by $(-\Delta)_{\mathrm{H}^1(\mathbb{R}^n)}$.

Theorem 4.6. *Let $(-\Delta)_{\mathrm{H}^1(\mathbb{R}^n)}$ be the negative Laplacian on $\mathbb{R}^n$. We have*

$$\mathrm{dom}((-\Delta)_{\mathrm{H}^1(\mathbb{R}^n)}) = \mathrm{H}^2(\mathbb{R}^n)$$

with equivalent norms

$$\|u\|_{\mathrm{L}^2(\mathbb{R}^n)} + \|(-\Delta)_{\mathrm{H}^1(\mathbb{R}^n)} u\|_{\mathrm{L}^2(\mathbb{R}^n)} \simeq \|u\|_{\mathrm{H}^2(\mathbb{R}^n)}$$

and for all $u \in \mathrm{dom}((-\Delta)_{\mathrm{H}^1(\mathbb{R}^n)})$ we have $(-\Delta)_{\mathrm{H}^1(\mathbb{R}^n)} u = -\Delta u$, where the right-hand side is understood in the H^2-sense.

Remark 4.7. (a) Theorem 4.6 is an example of a smoothing property of a differential operator: The knowledge of $u \in \mathrm{H}^1(\mathbb{R}^n)$ and $(-\Delta)_{\mathrm{H}^1(\mathbb{R}^n)} u \in \mathrm{L}^2(\mathbb{R}^n)$ is enough to control all weak derivatives of u up to order 2.

(b) The smoothing property is related to the smoothness of the coefficients $A = \mathrm{id}_{\mathbb{C}^n}$, see [Eva10] for further background. For general elliptic operators in divergence form the smoothing property can fail badly, see Exercise 4.6.

(c) In dimension $n = 1$ a smoothing property for the negative Dirichlet Laplacian $(-\Delta)_{\mathrm{H}^1_0(\Omega)}$ can be proved on any non-empty open set Ω, see Exercise 4.3. This argument breaks down in higher dimensions.

Proof of Theorem 4.6. First, suppose that $(u, v) \in (-\Delta)_{\mathrm{H}^1(\mathbb{R}^n)}$. By definition via the form method, see Section 3.5, Parseval's formula and Proposition 4.5, we obtain for all $w \in \mathrm{H}^1(\mathbb{R}^n)$ that

$$\begin{aligned} \int_{\mathbb{R}^n} \widehat{v}(\xi) \cdot \overline{\widehat{w}(\xi)} \,\mathrm{d}\xi = \int_{\mathbb{R}^n} v(x) \cdot \overline{w(x)} \,\mathrm{d}x &= \int_{\mathbb{R}^n} \nabla u(x) \cdot \overline{\nabla w(x)} \,\mathrm{d}x \\ &= \int_{\mathbb{R}^n} 4\pi^2 |\xi|^2 \widehat{u}(\xi) \cdot \overline{\widehat{w}(\xi)} \,\mathrm{d}\xi. \end{aligned} \tag{4.4}$$

From this we conclude $4\pi^2|\xi|^2\widehat{u}(\xi) = \widehat{v}(\xi)$ by Lemma 3.4 and then $u \in \mathrm{H}^2(\mathbb{R}^n)$ with $-\Delta u = v$ $(= (-\Delta)_{\mathrm{H}^1(\mathbb{R}^n)} u)$ by Proposition 4.5. Moreover, the same proposition yields

$$\|u\|_{\mathrm{H}^2(\mathbb{R}^n)} \simeq \|u\|_{\mathrm{L}^2(\mathbb{R}^n)} + \|\Delta u\|_{\mathrm{L}^2(\mathbb{R}^n)} = \|u\|_{\mathrm{L}^2(\mathbb{R}^n)} + \|(-\Delta)_{\mathrm{H}^1(\mathbb{R}^n)} u\|_{\mathrm{L}^2(\mathbb{R}^n)}.$$

Conversely, given $u \in \mathrm{H}^2(\mathbb{R}^n)$, the calculation (4.4) works for any $w \in \mathrm{H}^1(\mathbb{R}^n)$ and the choice $v := -\Delta u$. By definition, this means $(u, -\Delta u) \in (-\Delta)_{\mathrm{H}^1(\mathbb{R}^n)}$. □

With Theorem 4.6 at hand, we will simply write $-\Delta$ instead of $(-\Delta)_{\mathrm{H}^1(\mathbb{R}^n)}$ in the following. Re-inspecting the above proof reveals the following result.

Corollary 4.8. *The negative Laplacian is unitarily equivalent (via the Fourier transform) to the multiplication operator on* $\mathrm{L}^2(\mathbb{R}^n)$ *associated with* $m(\xi) = 4\pi^2|\xi|^2$*, that is,*

$$-\Delta = \mathcal{F}^{-1} M_m \mathcal{F}.$$

This result is of course a special case of the spectral theorem for self-adjoint operators. It can be visualized by the following commutative diagram:

$$\begin{tikzcd} \mathrm{L}^2(\mathbb{R}^n) \supseteq \mathrm{dom}(M_m) \arrow[r, "M_m"] & \mathrm{L}^2(\mathbb{R}^n) \\ \mathrm{L}^2(\mathbb{R}^n) \supseteq \mathrm{dom}(-\Delta) \arrow[u, "\mathcal{F}"] \arrow[r, "-\Delta"] & \mathrm{L}^2(\mathbb{R}^n) \arrow[u, "\mathcal{F}"] \end{tikzcd}$$

Figure 4.1.: Unitary equivalence of $-\Delta$ with a multiplication operator via the Fourier transform, where $m(\xi) = 4\pi^2|\xi|^2$.

4.3. A first glimpse at functional calculus

For the unbounded operator $L = -\Delta$ we know how to define integer powers L^k and we can now give a description of their domains.

Proposition 4.9. *For all* $k \in \mathbb{N}$ *we have*

$$\mathrm{dom}\big((-\Delta)^k\big) = \mathrm{H}^{2k}(\mathbb{R}^n)$$

with equivalent (graph) norms.

Proof. In view of Corollary 4.8 the operator $(-\Delta)^k$ is equivalent to the multiplication operator on $\mathrm{L}^2(\mathbb{R}^n)$ associated with $m^k(\xi) = (4\pi^2)^k|\xi|^{2k}$. Proposition 4.5 and Plancherel's theorem yield $\mathrm{H}^{2k}(\mathbb{R}^n) = \mathrm{dom}\big((-\Delta)^k\big)$ and

$$\|u\|_{\mathrm{H}^k(\mathbb{R}^n)} \simeq \|u\|_{\mathrm{L}^2(\mathbb{R}^n)} + \big\|\xi \mapsto |\xi|^{2k}\widehat{u}(\xi)\big\|_{\mathrm{L}^2(\mathbb{R}^n)} \simeq \|u\|_{\mathrm{L}^2(\mathbb{R}^n)} + \|(-\Delta)^k u\|_{\mathrm{L}^2(\mathbb{R}^n)}$$

for all $u \in \mathrm{H}^{2k}(\mathbb{R}^n)$. □

It is not so obvious how to define more complicated functions of L, like $\sqrt{L}$, e^{-L} or $L^{1/4}\mathrm{e}^{-\sqrt{L}}$. Answering this question leads to the concept of a functional calculus, which will play an important role in our course. At this point we can use Figure 4.1 in order to proceed for the negative Laplacian, because if f is a measurable function defined on the range of m, i.e., on $[0,\infty)$, then $f(-\Delta)$ should be equivalent to the multiplication operator associated with $f \circ m$.

Definition 4.10. *Let $f\colon [0,\infty) \to \mathbb{C}$ be a measurable function. Define $f(-\Delta)$ via the commutative diagram*

$$\begin{array}{ccc} \mathrm{L}^2(\mathbb{R}^n) \supseteq \operatorname{dom}(M_{f\circ m}) & \xrightarrow{M_{f\circ m}} & \mathrm{L}^2(\mathbb{R}^n) \\ \mathcal{F}\big\uparrow & & \mathcal{F}\big\uparrow \\ \mathrm{L}^2(\mathbb{R}^n) \supseteq \operatorname{dom}(f(-\Delta)) & \xrightarrow{f(-\Delta)} & \mathrm{L}^2(\mathbb{R}^n), \end{array}$$

that is, $f(-\Delta) = \mathcal{F}^{-1} M_{f\circ m}\mathcal{F}$, where the multiplication operator $M_{f\circ m}$ is defined with maximal domain in $\mathrm{L}^2(\mathbb{R}^n)$, see Example 1.8.

The map

$$\begin{aligned} \big\{\text{measurable functions on } [0,\infty)\big\} &\to \big\{\text{closed operators in } \mathrm{L}^2(\mathbb{R}^n)\big\} \\ f &\mapsto f(-\Delta) \end{aligned}$$

becomes some sort of algebra homomorphism. This theory will be developed in greater generality in the upcoming two lectures, see also Exercise 4.4.

A particularly interesting choice of functions f is $f(t) = t^\alpha$ for $\alpha > 0$, leading to the operators $(-\Delta)^\alpha$, which we call *fractional powers* of the negative Laplacian. Their domains generalize the spaces appearing on the left-hand side in Proposition 4.9 to non-integer k and provide a natural definition for Sobolev spaces of fractional order!

Definition 4.11. *Let $\alpha > 0$. The space*

$$\mathrm{H}^\alpha(\mathbb{R}^n) \coloneqq \operatorname{dom}\big((-\Delta)^{\alpha/2}\big)$$

equipped with the graph norm is called fractional Sobolev space *of* order α.

A particularly attentive reader might have noticed that this definition is troublesome when α is an odd integer, because in this case H^α has been defined before and up to now we did not check that the two definitions coincide. Fortunately, this is just a repetition of the proof of Proposition 4.9, replacing $2k$ by k. The special case $\alpha = 1$ reveals a connection between fractional powers and the form domain of the negative Laplacian that we state explicitly.

Proposition 4.12 (Kato property for the negative Laplacian on $\mathbb{R}^n$)**.** *Let L be the negative Laplacian on $\mathbb{R}^n$. Then* $\operatorname{dom}(\sqrt{L}) = \mathrm{H}^1(\mathbb{R}^n)$ *and*

$$\|\sqrt{L}u\|_{\mathrm{L}^2(\mathbb{R}^n)} = \|\nabla u\|_{\mathrm{L}^2(\mathbb{R}^n)} \qquad (u \in \mathrm{H}^1(\mathbb{R}^n)).$$

The following proposition shows that the fractional Sobolev spaces $\mathrm{H}^\alpha(\mathbb{R}^n)$, $0 < \alpha < 1$, fill the 'gap' between $\mathrm{L}^2(\mathbb{R}^n)$ and $\mathrm{H}^1(\mathbb{R}^n)$ in a way that resembles how Hölder spaces $\mathrm{C}^\alpha(\mathbb{R}^n)$ fill the gap between continuous and continuously differentiable functions.

Proposition 4.13. *Let* $0 < \alpha < 1$. *We have* $u \in \mathrm{H}^\alpha(\mathbb{R}^n)$ *if and only if* $u \in \mathrm{L}^2(\mathbb{R}^n)$ *and*

$$[u]_{\alpha,2} \coloneqq \left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|u(x)-u(y)|^2}{|x-y|^{2\alpha}} \frac{\mathrm{d}x\,\mathrm{d}y}{|x-y|^n}\right)^{1/2} < \infty.$$

Moreover $\left(\|\cdot\|^2_{\mathrm{L}^2(\mathbb{R}^n)} + [\,\cdot\,]^2_{\alpha,2}\right)^{1/2}$ *defines an equivalent norm on* $\mathrm{H}^\alpha(\mathbb{R}^n)$ *and there is a constant* $C > 0$ *such that*

$$\|(-\Delta)^{\alpha/2}u\|_{\mathrm{L}^2(\mathbb{R}^n)} = C[u]_{\alpha,2} \qquad (u \in \mathrm{H}^\alpha(\mathbb{R}^n)).$$

Proof. For $u \in \mathrm{L}^2(\mathbb{R}^n)$ we can compute

$$\begin{aligned}
\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|u(x)-u(y)|^2}{|x-y|^{2\alpha}} \frac{\mathrm{d}x\,\mathrm{d}y}{|x-y|^n} &\overset{y=x-h}{=} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|u(x)-u(x-h)|^2}{|h|^{2\alpha}}\,\mathrm{d}x\,\frac{\mathrm{d}h}{|h|^n} \\
&\overset{\text{Plancherel}}{=} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|(1-\mathrm{e}^{-2\pi\mathrm{i}h\cdot\xi})\widehat{u}(\xi)|^2}{|h|^{2\alpha}}\,\mathrm{d}\xi\,\frac{\mathrm{d}h}{|h|^n} \\
&\overset{\text{Tonelli}}{=} \int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 \int_{\mathbb{R}^n} \frac{|1-\mathrm{e}^{-2\pi\mathrm{i}h\cdot\xi}|^2}{|h|^{2\alpha}}\,\frac{\mathrm{d}h}{|h|^n}\,\mathrm{d}\xi.
\end{aligned}$$

Let us investigate the inner integral more carefully. We introduce

$$I\colon \mathbb{R}^n \setminus \{0\} \to (0,\infty), \quad I(\xi) \coloneqq \int_{\mathbb{R}^n} \frac{|1-\mathrm{e}^{-2\pi\mathrm{i}h\cdot\xi}|^2}{|2\pi\xi|^{2\alpha}|h|^{2\alpha}}\,\frac{\mathrm{d}h}{|h|^n}.$$

To see that these integrals are finite for fixed ξ, we use Taylor's theorem for $|h|$ small and the boundedness of the complex exponential for $|h|$ large to bound the integrand by $C(\xi)\min(|h|^{2-2\alpha}, |h|^{-2\alpha})$, which is integrable with respect to $\frac{\mathrm{d}h}{|h|^n}$. If $O \in \mathbb{R}^{n\times n}$ is an orthogonal matrix and $r > 0$ is a scalar, then $I(rO\xi) = I(\xi)$ by the change of variable $rO^T h = \zeta$. For any $\xi \neq 0$ there is a rotation O such that $|\xi|^{-1}O\xi = e_1$ is the first unit vector. Thus, $I(\xi) = I(e_1) =: C^2$ is independent of ξ. Altogether, we have shown

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|u(x)-u(y)|^2}{|x-y|^{2\alpha}} \frac{\mathrm{d}x\,\mathrm{d}y}{|x-y|^n} = C^2 \int_{\mathbb{R}^n} \left||2\pi\xi|^\alpha \widehat{u}(\xi)\right|^2 \mathrm{d}\xi,$$

regardless of whether this expression is finite or not. All further assertions follow from this by definition of the operator $(-\Delta)^{\alpha/2}$. $\square$

Let us explore further the functional calculus of the negative Laplacian via the Fourier transform. Proposition 1.9 immediately gives a description of those measurable functions f that correspond to bounded operators $f(-\Delta)$. We state it in the following theorem. In order to avoid measure theoretic technicalities when comparing f on $(0,\infty)$ and $f \circ m$ on $\mathbb{R}^n \setminus \{0\}$, we only consider continuous functions f. In this case, we clearly have that $\|f\|_\infty = \|f \circ m\|_{\mathrm{L}^\infty(\mathbb{R}^n)}$.

Theorem 4.14. *Let $f\colon (0,\infty) \to \mathbb{C}$ be continuous. Then $f(-\Delta)$ is a bounded operator if and only if f is bounded. In this case we have*

$$\|f(-\Delta)\|_{\mathcal{L}(\mathrm{L}^2(\mathbb{R}^n))} = \|f\|_\infty.$$

We continue with two classical results of Fourier analysis that allow us to reconstruct a given function $u \in \mathrm{L}^2(\mathbb{R}^n)$ and its norm from the knowledge of the convolutions $\theta_t * u$, where θ_t is the mollifier associated with a suitable smooth function θ that has the cancellation property $\widehat{\theta}(0) = \int_{\mathbb{R}^n} \theta(x) \, \mathrm{d}x = 0$. We recall that a function $\theta\colon \mathbb{R}^n \to \mathbb{C}$ is *radial* if for any orthogonal matrix $O \in \mathbb{R}^{n\times n}$ we have $\theta(x) = \theta(Ox)$ for all $x \in \mathbb{R}^n$. In this case, we can write $\theta(x) = \vartheta(|x|)$ for some function $\vartheta\colon [0,\infty) \to \mathbb{C}$, for example, $\vartheta(r) := \theta(re)$, where e is any fixed unit vector in $\mathbb{R}^n$. Moreover, the Fourier transform of a radial Schwartz function is again of that type, see Exercise 4.2.

Proposition 4.15 (Reproducing formula)**.** *Let $\theta \in \mathcal{S}(\mathbb{R}^n)$ be radial with $\widehat{\theta}(0) = 0$. Then in the sense of* $\mathrm{L}^2(\mathbb{R}^n)$*-valued Bochner integrals we have*

$$C_1(\theta)u = \lim_{\substack{\varepsilon\to 0\\ R\to\infty}} \int_\varepsilon^R \theta_{\sqrt{t}} * u \, \frac{\mathrm{d}t}{t} \qquad (u \in \mathrm{L}^2(\mathbb{R}^n)) \tag{4.5}$$

with the normalization factor $C_1(\theta) := 2\int_0^\infty \widehat{\theta}(se_1) \, \frac{\mathrm{d}s}{s}$.

Proof. Let $u \in \mathrm{L}^2(\mathbb{R}^n)$. In a first step we explain why the various objects in the statement are well-defined and then we prove the actual reproducing formula.

Step 1: Preliminaries.

We have $\widehat{\theta}(0) = 0$ by assumption and $\widehat{\theta} \in \mathcal{S}(\mathbb{R}^n)$ by Proposition 4.2 (a). Hence there is a constant $C > 0$ such that $|\widehat{\theta}(\xi)| \le C\min(|\xi|, |\xi|^{-1})$ for all $\xi \in \mathbb{R}^n$ and thus we have

$$\int_0^\infty |\widehat{\theta}(se_1)| \, \frac{\mathrm{d}s}{s} < \infty. \tag{4.6}$$

In particular, the constant $C_1(\theta)$ is finite.

In order to see that the integrals in (4.5) are defined in the sense of Bochner, it suffices to check that

$$f\colon (0,\infty) \to \mathrm{L}^2(\mathbb{R}^n), \quad f(t) := \theta_{\sqrt{t}} * u \tag{4.7}$$

is continuous, see Example A.12. Since the Fourier transform $\mathcal{F}$ is an isomorphism on $\mathrm{L}^2(\mathbb{R}^n)$, it suffices in fact to check that $\mathcal{F}f\colon (0,\infty) \to \mathrm{L}^2(\mathbb{R}^n)$ is continuous. By properties (d) and (c) in Proposition 4.2 and radiality of $\widehat{\theta}$, we can write $(\mathcal{F}f)(t)(\xi)$ as

$$\widehat{\theta_{\sqrt{t}} * u}(\xi) = \widehat{\theta}(\sqrt{t}\xi) \cdot \widehat{u}(\xi) = \widehat{\theta}(\sqrt{t}|\xi|e_1) \cdot \widehat{u}(\xi) \qquad (\text{a.e. } \xi \in \mathbb{R}^n). \tag{4.8}$$

Here, $\widehat{\theta}$ is bounded and continuous and $\widehat{u} \in \mathrm{L}^2(\mathbb{R}^n)$. Thus, continuity of $\mathcal{F}f$ follows directly from the dominated convergence theorem.

Step 2: A formula for the integrals on $[\varepsilon, R]$.

Let $0 < \varepsilon < R < \infty$. We shall prove

$$\int_\varepsilon^R \mathcal{F} f(t) \, \frac{\mathrm{d}t}{t} = I_{\varepsilon,R} \cdot \widehat{u}, \tag{4.9}$$

where

$$I_{\varepsilon,R}(\xi) := \int_{\sqrt{\varepsilon}|\xi|}^{\sqrt{R}|\xi|} 2\widehat{\theta}(se_1) \, \frac{\mathrm{d}s}{s} \qquad (\xi \in \mathbb{R}^n).$$

In (4.9), both sides depend continuously on $\widehat{u} \in \mathrm{L}^2(\mathbb{R}^n)$ — for the term on the left we use (4.8) and dominated convergence, whereas on the right we note that the functions $I_{\varepsilon,R}$ are bounded by twice the value of the integral in (4.6). Thus, by density, we can assume that $\widehat{u}$ is continuous with compact support $K := \operatorname{supp}(u)$. Calling $X \hookrightarrow \mathrm{L}^2(\mathbb{R}^n)$ the Banach space of continuous functions on $\mathbb{R}^n$ that vanish outside of K, we obtain from (4.8) that $\mathcal{F} f \colon [\varepsilon, R] \to X$ is continuous. Now, the consistency properties of the Bochner integral (Proposition A.13 and Corollary A.15) allow us to compute

$$\left(\int_\varepsilon^R \mathcal{F} f(t) \, \frac{\mathrm{d}t}{t}\right)(\xi) \quad = \quad \int_\varepsilon^R \widehat{\theta}(\sqrt{t}|\xi|e_1) \cdot \widehat{u}(\xi) \, \frac{\mathrm{d}t}{t} \overset{s = \sqrt{t}|\xi|}{=} I_{\varepsilon,R}(\xi) \cdot \widehat{u}(\xi).$$

Step 3: Proof of the reproducing formula.

Since $I_{\varepsilon,R}(\xi)$ tends to $C_1(\theta)$ in the limit as $\varepsilon \to 0$, $R \to \infty$ for every $\xi \neq 0$, dominated convergence yields

$$\lim_{\substack{\varepsilon \to 0 \\ R \to \infty}} \int_\varepsilon^R \mathcal{F} f(t) \, \frac{\mathrm{d}t}{t} = C_1(\theta)\widehat{u} \qquad (\text{in } \mathrm{L}^2(\mathbb{R}^n)).$$

The claim follows by applying $\mathcal{F}^{-1}$ on both sides and using its continuity on $\mathrm{L}^2(\mathbb{R}^n)$ to interchange it first with the limit and then with the integral (Proposition A.13). □

Proposition 4.16 (Square function norm)**.** *Let* $\theta \in \mathcal{S}(\mathbb{R}^n)$ *be radially symmetric with* $\widehat{\theta}(0) = 0$. *Then*

$$C_2(\theta)\|u\|_{\mathrm{L}^2(\mathbb{R}^n)}^2 = \int_0^\infty \|\theta_{\sqrt{t}} * u\|_{\mathrm{L}^2(\mathbb{R}^n)}^2 \, \frac{\mathrm{d}t}{t} \qquad (u \in \mathrm{L}^2(\mathbb{R}^n))$$

with the normalization factor $C_2(\theta) := 2\int_0^\infty |\widehat{\theta}(se_1)|^2 \, \frac{\mathrm{d}s}{s}$.

Proof. Using (4.8), we compute

$$
\begin{aligned}
\int_0^\infty \|\theta_{\sqrt{t}} * u\|_{\mathrm{L}^2(\mathbb{R}^n)}^2 \frac{\mathrm{d}t}{t} &\overset{\text{Plancherel}}{=} \int_0^\infty \int_{\mathbb{R}^n} |\widehat{\theta}(\sqrt{t}|\xi|e_1)|^2 |\widehat{u}(\xi)|^2 \,\mathrm{d}\xi \,\frac{\mathrm{d}t}{t} \\
&\overset{\text{Tonelli}}{=} \int_{\mathbb{R}^n} \int_0^\infty |\widehat{\theta}(\sqrt{t}|\xi|e_1)|^2 |\widehat{u}(\xi)|^2 \frac{\mathrm{d}t}{t} \,\mathrm{d}\xi \\
&\overset{s = \sqrt{t}|\xi|}{=} \int_{\mathbb{R}^n} \int_0^\infty 2|\widehat{\theta}(se_1)|^2 |\widehat{u}(\xi)|^2 \frac{\mathrm{d}s}{s} \,\mathrm{d}\xi \\
&= C_2(\theta) \int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 \,\mathrm{d}\xi \\
&\overset{\text{Plancherel}}{=} C_2(\theta) \|u\|_{\mathrm{L}^2(\mathbb{R}^n)}^2. \qquad \square
\end{aligned}
$$

Parts of Fourier analysis can be understood as properties of the functional calculus for the negative Laplacian and vice versa. We will illustrate the principle by reformulating the previous two results. Later on, in Lecture 8, we will benefit from this point of view in order to develop 'Fourier analysis' for more general m-accretive operators and in particular for elliptic operators in divergence form.

To do so, we slightly change our perspective and start with a Schwartz function $f \in \mathcal{S}(\mathbb{R})$ on the real line with $f(0) = 0$. To f we associate a function ϕ on $\mathbb{R}^n$ by setting $\phi(\xi) := f(4\pi^2|\xi|^2)$. We see that ϕ is a Schwartz function (Exercise 4.2) that is radial and satisfies $\phi(0) = 0$. Hence, $\theta := \widehat{\phi}$ satisfies the assumptions of the previous two propositions and we have

$$\widehat{\theta}(\sqrt{t}\xi) = f(4\pi^2 t|\xi|^2) \qquad (\xi \in \mathbb{R}^n,\ t > 0).$$

Now, (4.8) yields

$$\theta_{\sqrt{t}} * u = f(-t\Delta)u \qquad (u \in \mathrm{L}^2(\mathbb{R}^n),\ t > 0),$$

where the right-hand side is defined by plugging $-\Delta$ into the function $f(t\,\boldsymbol{\cdot})$, see Theorem 4.14. In this setting, Propositions 4.15 and 4.16 take the following form.

Corollary 4.17. *Let $f \in \mathcal{S}(\mathbb{R})$ be such that $f(0) = 0$. Define the normalizing factors $C_1(f) := \int_0^\infty f(s) \frac{\mathrm{d}s}{s}$ and $C_2(f) := \int_0^\infty |f(s)|^2 \frac{\mathrm{d}s}{s}$. Then for all $u \in \mathrm{L}^2(\mathbb{R}^n)$ we have*

$$C_1(f)u = \lim_{\substack{\varepsilon \to 0 \\ R \to \infty}} \int_\varepsilon^R f(-t\Delta)u \,\frac{\mathrm{d}t}{t} \tag{4.10}$$

as a limit in $\mathrm{L}^2(\mathbb{R}^n)$ and

$$C_2(f)\|u\|_{\mathrm{L}^2(\mathbb{R}^n)}^2 = \int_0^\infty \|f(-t\Delta)u\|_{\mathrm{L}^2(\mathbb{R}^n)}^2 \frac{\mathrm{d}t}{t}. \tag{4.11}$$

4.4. Exercises

Exercise 4.1. Extend Proposition 4.2 (d) to $u \in \mathrm{L}^1(\mathbb{R}^n)$, $v \in \mathrm{L}^2(\mathbb{R}^n)$. To this end, recall (or quickly prove) that

- the Fourier transform $\mathcal{F}\colon \mathrm{L}^1(\mathbb{R}^n) \to \mathrm{C}_\mathrm{b}(\mathbb{R}^n)$ is bounded, where $\mathrm{C}_\mathrm{b}(\mathbb{R}^n)$ is the Banach space of bounded and continuous functions on $\mathbb{R}^n$,
- the bilinear map
$$\mathrm{L}^1(\mathbb{R}^n) \times \mathrm{L}^2(\mathbb{R}^n) \to \mathrm{L}^2(\mathbb{R}^n), \quad (u, v) \mapsto u * v$$
is well-defined and bounded.

In which space does the identity of Proposition 4.2 (d) eventually hold?

Exercise 4.2 (Radial functions and the Fourier transform)**.**

(a) Show that the Fourier transform of a radial function $f \in \mathcal{S}(\mathbb{R}^n)$ is again radial.

(b) Let $f \in \mathcal{S}(\mathbb{R})$ and define ϕ on $\mathbb{R}^n$ by $\phi(x) := f(|x|^2)$. Prove that $\phi \in \mathcal{S}(\mathbb{R}^n)$.

Exercise 4.3 (The smoothing property in dimension $n = 1$)**.** Let $\Omega \subseteq \mathbb{R}$ be an open set. Prove that

$$\operatorname{dom}((-\Delta)_{\mathrm{H}^1_0(\Omega)}) = \mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega).$$

Why does the same proof not work in higher dimensions?

Exercise 4.4. Prove that the map

$$\mathrm{C}_\mathrm{b}((0,\infty)) \to \mathcal{L}(\mathrm{L}^2(\mathbb{R}^n)), \quad f \mapsto f(-\Delta)$$

is a contractive homomorphism of algebras.

Exercise 4.5 (The Dirichlet problem via functional calculus)**.** In this exercise we propose a way of solving in a weak sense the following *Dirichlet problem* in the upper half-space $(0,\infty) \times \mathbb{R}^n$:

$$-(\partial_t^2 + \partial_{x_1}^2 + \ldots + \partial_{x_n}^2)u = 0 \quad \text{in } (0,\infty) \times \mathbb{R}^n, \tag{4.12}$$
$$\lim_{t \to 0} u(t, \cdot) = f \quad \text{in } \mathrm{L}^2(\mathbb{R}^n), \tag{4.13}$$

where $f \in \mathrm{L}^2(\mathbb{R}^n)$ is given and $\partial_t^2 + \partial_{x_1}^2 + \ldots + \partial_{x_n}^2 = \partial_t^2 + \Delta$ is the Laplacian in dimension $(n+1)$, whereas Δ is the Laplacian in dimension n as in the lecture. To this end, we define

$$u(t,x) := (\mathrm{e}^{-t\sqrt{-\Delta}} f)(x) \qquad ((t,x) \in (0,\infty) \times \mathbb{R}^n),$$

where the right-hand side is understood in virtue of Theorem 4.14.

(a) Make a formal calculation to convince yourself that u should solve (4.12) and (4.13).

(b) Prove that u can be understood as a continuous function on $(0, \infty) \times \mathbb{R}^n$.

Hint: You will be able to use the Fourier transform on $\mathrm{L}^1(\mathbb{R}^n)$.

(c) Verify that u satisfies the boundary condition (4.13).

(d) Prove that in the weak sense we have

$$\Delta u = (\Delta \mathrm{e}^{-t\sqrt{-\Delta}})f.$$

(e) Let $\phi \in \mathrm{C}_\mathrm{c}^\infty((0, \infty) \times \mathbb{R}^n)$. Show that

$$\partial_t \left[\widehat{\phi(t, \cdot)}\right] = \widehat{\partial_t \phi(t, \cdot)},$$

where the Fourier transform is the one on $\mathbb{R}^n$ and the derivative is in the classical sense.

(f) Prove that in the weak sense we have

$$\partial_t u = \left(-\sqrt{-\Delta}\mathrm{e}^{-t\sqrt{-\Delta}}\right)f \quad \text{and} \quad \partial_t^2 u = \left(-\Delta \mathrm{e}^{-t\sqrt{-\Delta}}\right)f.$$

(g) Conclude that u solves (4.12) with derivatives understood in the weak sense.

(h) Prove that all first-order weak derivatives of u are in $\mathrm{L}^2((0, \infty) \times \mathbb{R}^n)$ if and only if $f \in \mathrm{H}^{1/2}(\mathbb{R}^n)$.

Part (h) shows that fractional Sobolev spaces are connected to boundary values of functions with weak gradient in $\mathrm{L}^2((0, \infty) \times \mathbb{R}^n)$ and in fact $\mathrm{H}^{1/2}(\mathbb{R}^n)$ can also be characterized in such way, see, e.g., [DNPV12, Prop. 4.5] for further background.

Exercise 4.6 (A divergence form operator with strange domain)**.** The goal of this exercise is to construct in $\mathrm{L}^2((0, 1))$ an elliptic operator in divergence form $L = -\operatorname{div}(A\nabla\,\cdot)$ with Dirichlet boundary conditions such that the smoothing property from Remark 4.7 fails as hard as it possibly can, namely

$$\operatorname{dom}(L) \cap \mathrm{H}^2((0, 1)) = \{0\}. \tag{4.14}$$

You may proceed as follows.

(a) Prove that for any L as above we have

$$\operatorname{dom}(L) = \left\{u \in \mathrm{H}_0^1((0, 1)) \mid Au' \in \mathrm{H}^1((0, 1))\right\}.$$

The construction of the bad L is based on a *well-distributed* measurable set $E \subseteq [0, 1]$, that is, a measurable set with the property that for any non-empty open interval $I \subseteq [0, 1]$ we have $0 < |E \cap I| < |I|$.

(b) Define $A(x) = 1 + \mathbf{1}_E(x)$. Prove that the corresponding operator L has indeed the property (4.14).

For the interested readers we propose as a supplementary exercise one possible construction of E following an idea of Rudin [Rud83]. We define inductively

$$\begin{aligned}
C_0 &\coloneqq [0,1] \\
C_1 &\coloneqq [0,\tfrac{3}{8}] \cup [\tfrac{5}{8},1] \\
C_2 &\coloneqq [0,\tfrac{5}{32}] \cup [\tfrac{7}{32},\tfrac{3}{8}] \cup [\tfrac{5}{8},\tfrac{25}{32}] \cup [\tfrac{27}{32},1] \\
&\ \ \vdots \qquad\qquad\qquad\qquad \vdots
\end{aligned}$$

that is, C_{j-1} is the union of 2^{j-1} intervals and we remove an open interval of length 4^{-j} from the middle of each to obtain C_j. The set $C \coloneqq \bigcap_{j=0}^{\infty} C_j$ is called *fat Cantor set*.

(c) Prove that $C \subseteq [0,1]$ is closed, has empty interior and measure $1/2$.

In the following, closed intervals with positive measure are called *segments* and a closed subset of a set I that has empty interior and positive measure is called *bad* subset of I. Hence, C is a bad subset of $[0,1]$. By scaling and translation, every segment has a bad subset.

(d) We let $I_0, I_1, \ldots$ be an enumeration of all segments in $[0,1]$ with rational endpoints. Justify that the following algorithm produces sequences $E_0, E_1, \ldots$ and $F_0, F_1, \ldots$ of bad subsets $E_j, F_j \subseteq I_j$:

`Start with disjoint bad subsets` E_0 `and` F_0 `of` I_0.

`Once` $E_0, F_0, \ldots E_{j-1}, F_{j-1}$ `are chosen, let` G `be their union.`

`Pick a segment` $J \subseteq I_j \setminus G$.

`Choose a pair of disjoint bad subsets` $E_j, F_j \subseteq J$.

(e) Prove that $E \coloneqq \bigcup_{j=0}^{\infty} E_j$ is well-distributed.

5. Functional calculus for sectorial operators

In the previous lecture, we have developed a functional calculus for the negative Laplacian on $\mathrm{L}^2(\mathbb{R}^n)$: Making essential use of its unitary equivalence to a multiplication operator (the spectral theorem), we have created a map

$$f \mapsto f(-\Delta)$$

that 'plugs' this operator into certain admissible functions, the measurable functions on $[0, \infty)$ in this case. In this lecture, we are going to develop this idea in more generality, focusing on sectorial operators. We follow Markus Haase's book [Haa06] from which three of us virtual lecturers have learned about functional calculus for the first time. The substitute for the spectral theorem comes from the work of Dunford and Riesz (see [DS58, Sect. VII.11] for a historical account). It is inspired by the reproducing structure

$$f(\lambda) = \frac{1}{2\pi \mathrm{i}} \int_\gamma f(z)(z-\lambda)^{-1} \, \mathrm{d}z$$

of the Cauchy integral formula for holomorphic functions. Replacing formally λ by an operator L, all it takes is to interpret the right-hand side in terms of resolvents in order to define

$$f(L) = \frac{1}{2\pi \mathrm{i}} \int_\gamma f(z)(z-L)^{-1} \, \mathrm{d}z. \tag{5.1}$$

At this point this is only a formal calculation. Convergence of the integral, choice of the path γ and existence of the resolvent remains nebulous, to say the least, but we are going to get there soon . . .

Notation 5.1. Throughout the entire lecture, L denotes a sectorial operator in a Hilbert space H and $\varphi_L \in [0, \pi)$ its sectoriality angle, see Definition 2.2. In the complex plane we will write from now on $\mathbf{1}$ and $\mathbf{z}$ for the functions $z \mapsto 1$ and $z \mapsto z$, respectively.

5.1. Elementary functional calculus

We will work with the following classes of holomorphic functions.

Definition 5.2. *Let $\varphi \in (0, \pi)$ and consider the following spaces of holomorphic functions $f\colon \mathrm{S}_\varphi \to \mathbb{C}$ endowed with the supremum norm*

$$\|f\|_{\infty,\varphi} := \sup_{z \in \mathrm{S}_\varphi} |f(z)|.$$

(a) The bounded holomorphic functions $\mathrm{H}^\infty(\mathrm{S}_\varphi)$.

(b) The subspace of functions with regular decay *as $|z| \to 0$ and $|z| \to \infty$:*

$$\mathrm{H}_0^\infty(\mathrm{S}_\varphi) := \Big\{ f \in \mathrm{H}^\infty(\mathrm{S}_\varphi) \mid \exists C, s > 0 \, \forall z \in \mathrm{S}_\varphi : |f(z)| \le C \min(|z|^s, |z|^{-s}) \Big\}.$$

(c) The subspace spanned by $\mathrm{H}_0^\infty(\mathrm{S}_\varphi)$, the rational function $(1+\mathbf{z})^{-1}$ and the constant function $\mathbf{1}$, called Dunford–Riesz class

$$\mathcal{E}(\mathrm{S}_\varphi) := \mathrm{H}_0^\infty(\mathrm{S}_\varphi) \oplus \langle (1+\mathbf{z})^{-1} \rangle \oplus \langle \mathbf{1} \rangle.$$

Remark 5.3. (a) An important element of $\mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ is $\mathbf{z}(1+\mathbf{z})^{-2}$. The class $\mathcal{E}(\mathrm{S}_\varphi)$ contains all '*resolvent functions*' $(\lambda - \mathbf{z})^{-1}$ with $\pi \ge |\arg \lambda| > \varphi$, because

$$\frac{1}{\lambda - \mathbf{z}} = \frac{(1+\lambda^{-1})\mathbf{z}}{(\lambda - \mathbf{z})(1+\mathbf{z})} + \frac{\lambda^{-1}}{1+\mathbf{z}} + 0.$$

If $\varphi < \pi/2$, then it also contains the exponential functions $\mathrm{e}^{-t\mathbf{z}}$, where $t > 0$ is fixed, because

$$\mathrm{e}^{-t\mathbf{z}} = \left(\mathrm{e}^{-t\mathbf{z}} - \frac{1}{1+\mathbf{z}} \right) + \frac{1}{1+\mathbf{z}} + 0.$$

(b) All three spaces are even algebras, i.e., they are closed under pointwise multiplication of functions. For the Dunford–Riesz class this observation relies on the identity

$$\frac{1}{(1+\mathbf{z})^2} = \frac{-\mathbf{z}}{(1+\mathbf{z})^2} + \frac{1}{1+\mathbf{z}} + 0.$$

(c) The Dunford–Riesz class can be characterized as a subalgebra of $\mathrm{H}^\infty(\mathrm{S}_\varphi)$ through the existence of certain limits as $|z| \to 0$ and $|z| \to \infty$. This characterization then also ensures that the three subspaces in its definition are linearly independent, see Exercise 5.3.

For functions in $\mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ we get a Cauchy integral formula for sector-shaped infinite paths that even touch $z = 0$ lying at the boundary of their domain.

Lemma 5.4. *Let* $0 < \psi < \varphi < \pi$, $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$, $a \in \mathbb{C} \setminus \partial \mathrm{S}_\psi$, *and consider the path*

$$\gamma_\psi(t) := \begin{cases} -t\mathrm{e}^{\mathrm{i}\psi}, & t \in (-\infty, 0], \\ t\mathrm{e}^{-\mathrm{i}\psi}, & t \in [0, \infty). \end{cases}$$

Then

$$\frac{1}{2\pi\mathrm{i}} \int_{\gamma_\psi} \frac{f(z)}{z-a} \,\mathrm{d}z = \begin{cases} f(a), & \text{if } a \in \mathrm{S}_\psi, \\ 0, & \text{if } a \notin \mathrm{S}_\psi. \end{cases}$$

Proof. For $0 < \varepsilon < |a|/2 < 2|a| < R$ we consider the closed path $\gamma_{\psi,\varepsilon,R}$ that emerges as the boundary of $\mathrm{S}_\psi \cap (B(0,R) \setminus B(0,\varepsilon))$, see Figure 5.1. Thanks to the estimate on f, we find

$$\left| \int_{-\psi}^{\psi} \frac{f(R\mathrm{e}^{\mathrm{i}t})}{R\mathrm{e}^{\mathrm{i}t} - a} \mathrm{i}R\mathrm{e}^{\mathrm{i}t} \,\mathrm{d}t \right| \leq R \int_{-\psi}^{\psi} \frac{|f(R\mathrm{e}^{\mathrm{i}t})|}{|R\mathrm{e}^{\mathrm{i}t} - a|} \,\mathrm{d}t \leq R \int_{-\psi}^{\psi} \frac{CR^{-s}}{\frac{R}{2}} \,\mathrm{d}t = 4C\psi R^{-s}$$

and similarly

$$\left| \int_{-\psi}^{\psi} \frac{f(\varepsilon\mathrm{e}^{\mathrm{i}t})}{\varepsilon\mathrm{e}^{\mathrm{i}t} - a} \mathrm{i}\varepsilon\mathrm{e}^{\mathrm{i}t} \,\mathrm{d}t \right| \leq \frac{4C\psi}{|a|} \varepsilon^{1+s}.$$

Hence, the contribution of the closing arcs tends to zero as $\varepsilon \to 0$ and $R \to \infty$. Since the decay of f also entails that the integral over γ_ψ in question is absolutely convergent, this means that

$$\frac{1}{2\pi\mathrm{i}} \int_{\gamma_\psi} \frac{f(z)}{z-a} \,\mathrm{d}z = \lim_{\substack{\varepsilon \to 0 \\ R \to \infty}} \frac{1}{2\pi\mathrm{i}} \int_{\gamma_{\psi,\varepsilon,R}} \frac{f(z)}{z-a} \,\mathrm{d}z.$$

For all ε, R this final integral has the asserted value by the classical Cauchy integral formula. □

As a first step towards a functional calculus, we check that the integral in (5.1) is absolutely convergent in $\mathcal{L}(H)$ whenever $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ and $\gamma = \gamma_\psi$ is the path from Lemma 5.4. Of course, the angles should satisfy $\varphi_L < \psi < \varphi < \pi$, so that we are integrating in counterclockwise direction around the spectrum of L, see Figure 5.1. Since L is sectorial, we have a resolvent estimate $\|(z-L)^{-1}\|_{\mathcal{L}(H)} \lesssim |z|^{-1}$ for all $z \in \gamma_\psi$, see Definition 2.2. Therefore,

$$\int_{\mathbb{R}} \|\gamma_\psi'(t) f(\gamma_\psi(t))(\gamma_\psi(t) - L)^{-1}\|_{\mathcal{L}(H)} \,\mathrm{d}t \lesssim \int_0^\infty |f(t\mathrm{e}^{\mathrm{i}\psi})| + |f(t\mathrm{e}^{-\mathrm{i}\psi})| \,\frac{\mathrm{d}t}{t},$$

where the decay of f at the origin and infinity is just enough to make the integral on the right convergent. In fact, this is the reason why we work with the class $\mathrm{H}_0^\infty(\mathrm{S}_\varphi)$. Thus, the integral

$$\int_{\gamma_\psi} f(z)(z-L)^{-1} \,\mathrm{d}z$$

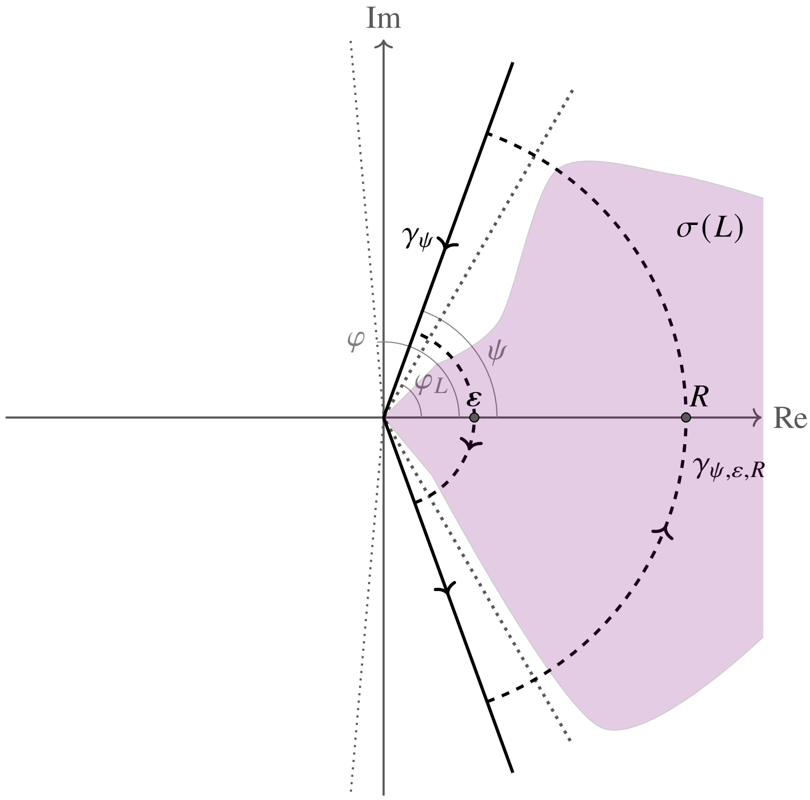

Figure 5.1.: The path γ_ψ around the spectrum of L, including the dashed circle arcs that are used to form the path $\gamma_{\psi,\varepsilon,R}$ in the proof of Lemma 5.4.

is absolutely convergent in $\mathcal{L}(H)$ and the so-defined operator is itself in $\mathcal{L}(H)$. Again by Cauchy's integral formula, the value of this integral does not depend on the angle ψ, see Exercise 5.1. This justifies the following definition.

Definition 5.5. *Let $\varphi \in (\varphi_L, \pi)$ and choose $\psi \in (\varphi_L, \varphi)$. For $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ set*

$$f(L) \coloneqq \frac{1}{2\pi \mathrm{i}} \int_{\gamma_\psi} f(z)(z-L)^{-1} \,\mathrm{d}z \quad \textit{with} \quad \gamma_\psi(t) = \begin{cases} -t\mathrm{e}^{\mathrm{i}\psi}, & t \in (-\infty, 0], \\ t\mathrm{e}^{-\mathrm{i}\psi}, & t \in [0,\infty). \end{cases}$$

The functions $(1+\mathbf{z})^{-1}$ and $\mathbf{1}$ are not contained in $\mathrm{H}_0^\infty(\mathrm{S}_\varphi)$, but we have a very precise idea what $(1+\mathbf{z})^{-1}(L)$ and $\mathbf{1}(L)$ should be. It is for this reason that we use the Dunford–Riesz class and extend the former definition as follows.

Definition 5.6. *Let $\varphi \in (\varphi_L, \pi)$. For $f = h + c(1+\mathbf{z})^{-1} + d \in \mathcal{E}(\mathrm{S}_\varphi)$ with $h \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ and $c, d \in \mathbb{C}$, set*

$$f(L) \coloneqq h(L) + c(1+L)^{-1} + d.$$

The map $\Phi_L \colon \mathcal{E}(\mathrm{S}_\varphi) \ni f \mapsto f(L) \in \mathcal{L}(H)$ is called elementary (sectorial) functional calculus *of L.*

Warning 5.7. The notation $f(L)$ is intuitive and convenient, as it describes perfectly that L is somehow plugged into the function f. Nevertheless, L does not act as something like a variable here. In fact, it is exactly the other way round. One should always think of $f(L)$ as the image of f under Φ_L. So, L is fixed and f varies!

Example 5.8. Let $L = -\Delta$ be the negative Laplacian on $\mathbb{R}^n$ from Lecture 4. Theorem 2.28 yields that L is sectorial with $\varphi_L = 0$. Let $f \in \mathcal{E}(\mathrm{S}_\varphi)$, where $\varphi \in (0, \pi)$. We check that the constructions of $f(-\Delta)$ in Definitions 5.6 and 4.10 coincide.

We know from Corollary 4.8 that (through the Fourier transform) L is equivalent to a multiplication operator M_m. Both definitions agree on that $\mathbf{1}(L) = \mathrm{id}_{\mathrm{L}^2(\mathbb{R}^n)}$. For $\lambda \in \varrho(L)$ we have

$$(\lambda - M_m)^{-1} = M_{(\lambda - m)^{-1}},$$

see Example 1.16. Hence, both definitions also agree on that $(1 + \mathbf{z})^{-1}(L) = (1 + L)^{-1}$. Finally, let $h \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ and $h(L)$ be as in Definition 5.5. Given $u \in \mathcal{S}(\mathbb{R}^n)$ with $\widehat{u} \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$, we pull bounded linear operators into the Bochner integral, see Proposition A.13, to get

$$\mathcal{F}(h(L)u) = \frac{1}{2\pi \mathrm{i}} \int_{\gamma_\psi} h(z) M_{(z-m)^{-1}} \widehat{u} \, \mathrm{d}z.$$

Due to the compact support of $\widehat{u}$, the Bochner integral can be understood in the space of continuous functions that vanish outside $K := \operatorname{supp}(\widehat{u})$. The argument presented below (4.9) yields for a.e. $\xi \in \mathbb{R}^n$ that

$$(\mathcal{F} h(L)u)(\xi) = \frac{1}{2\pi \mathrm{i}} \int_{\gamma_\psi} h(z)(z - m(\xi))^{-1} \widehat{u}(\xi) \, \mathrm{d}z = h(m(\xi)) \widehat{u}(\xi),$$

where we have used Lemma 5.4 in the final step. By density, this identity remains true for all $u \in \mathrm{L}^2(\mathbb{R}^n)$. Consequently, $h(L)$ is equivalent to the multiplication operator $M_{h \circ m}$ through the Fourier transform and this is precisely how it was defined in Lecture 4.

Let us prove that the algebraic structure of the Dunford–Riesz class is preserved by the elementary functional calculus.

Proposition 5.9. *Let $\varphi \in (\varphi_L, \pi)$. Then Φ_L is an algebra homomorphism, i.e., for all $f, g \in \mathcal{E}(\mathrm{S}_\varphi)$ and all $\lambda \in \mathbb{C}$ we have*

$$(\lambda f + g)(L) = \lambda f(L) + g(L) \quad \text{and} \quad (fg)(L) = f(L)g(L)$$

and in particular $f(L)g(L) = g(L)f(L)$. Moreover, we have $\mathbf{z}(1 + \mathbf{z})^{-2}(L) = L(1 + L)^{-2}$.

Proof. Linearity follows immediately from the definition of Φ_L and linearity of the Bochner integral. Because of linearity and symmetry, it suffices to check multiplicativity separately in the following three cases:

(a) $f, g \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$,

(b) $f = (1+\mathbf{z})^{-1}$ and $g \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$,

(c) $f = g = (1+\mathbf{z})^{-1}$.

Once we have multiplicativity, commutativity of $f(L)$ and $g(L)$ follows easily as $\mathcal{E}(\mathrm{S}_\varphi)$ is a commutative algebra:

$$f(L)g(L) = (fg)(L) = (gf)(L) = g(L)f(L).$$

Proof of (a). We choose $\varphi_L < \psi < \nu < \varphi$. By the resolvent identity and Fubini's theorem, we find

$$\begin{aligned}
f(L)g(L) &= \frac{1}{(2\pi\mathrm{i})^2}\int_{\gamma_\psi}\int_{\gamma_\nu} f(z)g(w)(z-L)^{-1}(w-L)^{-1}\,\mathrm{d}w\,\mathrm{d}z \\
&= \frac{1}{(2\pi\mathrm{i})^2}\int_{\gamma_\psi}\int_{\gamma_\nu} \frac{f(z)g(w)}{w-z}\big((z-L)^{-1}-(w-L)^{-1}\big)\,\mathrm{d}w\,\mathrm{d}z \\
&= \frac{1}{2\pi\mathrm{i}}\int_{\gamma_\psi} f(z)\bigg(\frac{1}{2\pi\mathrm{i}}\int_{\gamma_\nu}\frac{g(w)}{w-z}\,\mathrm{d}w\bigg)(z-L)^{-1}\,\mathrm{d}z \\
&\quad - \frac{1}{2\pi\mathrm{i}}\int_{\gamma_\nu} g(w)\bigg(\frac{1}{2\pi\mathrm{i}}\int_{\gamma_\psi}\frac{f(z)}{w-z}\,\mathrm{d}z\bigg)(w-L)^{-1}\,\mathrm{d}w.
\end{aligned}$$

The inner integrals can be evaluated by means of Lemma 5.4 unless $z = 0$ or $w = 0$, of course, see also Figure 5.2. The result is

$$= \frac{1}{2\pi\mathrm{i}}\int_{\gamma_\psi} f(z)g(z)(z-L)^{-1}\,\mathrm{d}z - 0$$

and we find by the definition of the functional calculus

$$= (fg)(L).$$

Proof of (b).

By definition and the resolvent identity we have

$$\begin{aligned}
f(L)g(L) &= \frac{1}{2\pi\mathrm{i}}\int_{\gamma_\psi} g(z)(1+L)^{-1}(z-L)^{-1}\,\mathrm{d}z \\
&= \frac{1}{2\pi\mathrm{i}}\int_{\gamma_\psi} \frac{g(z)}{1+z}\Big((1+L)^{-1}+(z-L)^{-1}\Big)\,\mathrm{d}z \\
&= (1+L)^{-1}\frac{1}{2\pi\mathrm{i}}\int_{\gamma_\psi}\frac{g(z)}{1+z}\,\mathrm{d}z + \frac{1}{2\pi\mathrm{i}}\int_{\gamma_\psi}\frac{g(z)}{1+z}(z-L)^{-1}\,\mathrm{d}z.
\end{aligned}$$

Now, we use Lemma 5.4 with $a = -1$ for the first integral and recall the definition of f when looking at the second one, to conclude

$$= 0 + (fg)(L).$$

Proof of (c). By definition we have $f(L)g(L) = (1 + L)^{-2}$. On the other hand, $fg = -h + (1 + \mathbf{z})^{-1}$ with $h = \mathbf{z}(1 + \mathbf{z})^{-2} \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$. Hence,

$$(fg)(L) - f(L)g(L) = -h(L) + (1 + L)^{-1} - (1 + L)^{-2} = L(1 + L)^{-2} - h(L),$$

and it remains to prove the additional claim in Proposition 5.9, namely that $h(L) = L(1 + L)^{-2}$.

To this end, we consider for $0 < \varepsilon < 1 < R < \infty$ the closed path $\widetilde{\gamma}_{\psi,\varepsilon,R}$ that emerges as the boundary of $(\mathbb{C} \setminus \mathrm{S}_\psi) \cap (B(0, R) \setminus B(0, \varepsilon))$. Note that this is exactly the opposite ('Obelix-sized') cake-piece of the one we considered in the proof of Lemma 5.4, see Figure 5.2. However, using an analogous reasoning as back then, we find

$$h(L) = \frac{1}{2\pi \mathrm{i}} \int_{\gamma_\psi} \frac{z}{(1 + z)^2}(z - L)^{-1}\,\mathrm{d}z = \lim_{\substack{\varepsilon\to 0\\ R\to\infty}} \frac{1}{2\pi \mathrm{i}} \int_{\widetilde{\gamma}_{\psi,\varepsilon,R}} \frac{z}{(1 + z)^2}(z - L)^{-1}\,\mathrm{d}z.$$

This integral can now be evaluated by the ($\mathcal{L}(H)$-valued) Cauchy integral formula for the derivative proved in the appendix, see Example A.23, and is independent of ε, R. Proposition 1.15 gives us

$$\begin{aligned}\big(\mathbf{z}(\mathbf{z} - L)^{-1}\big)' &= (\mathbf{z} - L)^{-1} - \mathbf{z}(\mathbf{z} - L)^{-2}\\ &= (\mathbf{z} - L)^{-1} - (\mathbf{z} - L)(\mathbf{z} - L)^{-2} - L(\mathbf{z} - L)^{-2} = -L(\mathbf{z} - L)^{-2},\end{aligned}$$

so, noting that $\widetilde{\gamma}_{\psi,\varepsilon,R}$ surrounds -1 clockwisely, we finally establish

$$h(L) = -\big(\mathbf{z}(\mathbf{z} - L)^{-1}\big)'(-1) = L(1 + L)^{-2}. \qquad \square$$

5.2. Regularization

There are many important functions f, for which we know intuitively what $f(L)$ should be, but which are not contained in $\mathcal{E}(\mathrm{S}_\varphi)$. The easiest example is the following: Do we have $\mathbf{z}(L) = L$? We already see that enlarging the class of functions for the functional calculus to suitable unbounded functions, will, in general, make us leave the comfort zone of bounded operators and we have to include closed unbounded operators in our calculus.

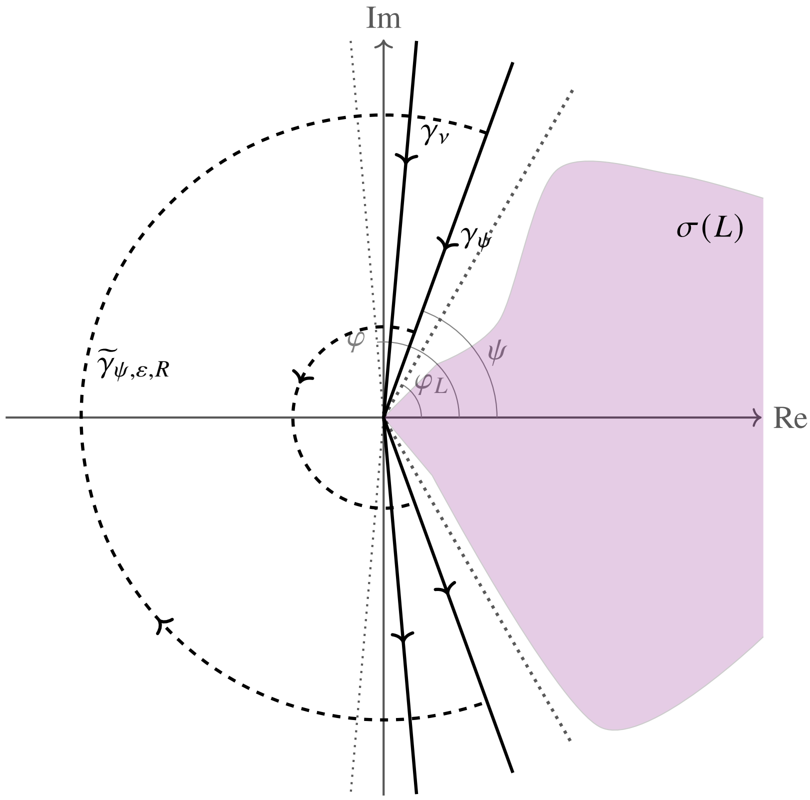

Figure 5.2.: The Obelix-sized cake-piece $\widetilde{\gamma}_{\psi,\varepsilon,R}$ in the proof of Proposition 5.9 and the two paths γ_ψ and γ_ν with $\varphi_L < \psi < \nu < \varphi$.

Definition 5.10. *Let* $\varphi \in (\varphi_L, \pi)$.

(a) The elements of the set

$$\mathrm{Reg}_L(\mathrm{S}_\varphi) := \{e \in \mathcal{E}(\mathrm{S}_\varphi) \mid e(L) \textit{ is injective}\}$$

are called regularizers.

(b) Let $\mathrm{M}(\mathrm{S}_\varphi)$ *denote the set of all meromorphic functions on* S_φ *and consider*

$$\mathrm{M}_L(\mathrm{S}_\varphi) := \{f \in \mathrm{M}(\mathrm{S}_\varphi) \mid \exists e \in \mathrm{Reg}_L(\mathrm{S}_\varphi) : ef \in \mathcal{E}(\mathrm{S}_\varphi)\}.$$

The elements of $\mathrm{M}_L(\mathrm{S}_\varphi)$ *are called* regularizable *and for* $f \in \mathrm{M}_L(\mathrm{S}_\varphi)$ *and a regularizer* e *with* $ef \in \mathcal{E}(\mathrm{S}_\varphi)$ *we say that* e regularizes f.

Remark 5.11. (a) The set $\mathrm{Reg}_L(\mathrm{S}_\varphi)$ contains at least $\mathbf{1}$ and the non-trivial regularizer $(1+\mathbf{z})^{-1}$. If L is injective, then also $\mathbf{z}(1+\mathbf{z})^{-2}$ is a regularizer thanks to Proposition 5.9.

(b) Since $\mathcal{E}(\mathrm{S}_\varphi)$ is an algebra and the composition of two injective operators is injective, the product of two regularizers is again a regularizer.

Regularizers are precious objects, since they allow us to extend our functional calculus in the following way. For $f \in \mathrm{M}_L(\mathrm{S}_\varphi)$ pick some regularizer e. Then the elementary functional calculus provides us with bounded operators $(ef)(L)$ and $e(L)$ and we know that $e(L)$ is injective. Thus, $e(L)^{-1}$ is a closed operator that heuristically corresponds to '$e^{-1}(L)$'. It is then rather natural to try the definition

$$\begin{aligned} f(L) &\coloneqq e(L)^{-1}\,(ef)(L) \\ \operatorname{dom}(f(L)) &\coloneqq \{u \in H \mid (ef)(L)u \in \operatorname{ran}(e(L))\} \end{aligned} \tag{5.2}$$

in order to extend the functional calculus to $\mathrm{M}_L(\mathrm{S}_\varphi)$.

We check that this idea gives rise to something well-defined.

Lemma 5.12. *The definition of $f(L)$ in* (5.2) *gives a closed operator that is independent of the particular choice of the regularizer e and consistent with the elementary functional calculus.*

Proof. Let $f \in \mathrm{M}_L(\mathrm{S}_\varphi)$ and let e_1 and e_2 both regularize f. By Remark 5.11 their product e_1e_2 is again a regularizer, and by Proposition 5.9 it holds that

$$(e_1e_2)(L)^{-1}\,(e_1e_2f)(L) = e_2(L)^{-1}\,e_1(L)^{-1}\,e_1(L)\,(e_2f)(L) = e_2(L)^{-1}\,(e_2f)(L).$$

Since $e_1e_2 = e_2e_1$, we analogously get $(e_1e_2)(L)^{-1}\,(e_1e_2f)(L) = e_1(L)^{-1}\,(e_1f)(L)$, so the two operators that are obtained with different regularizers are the same.

Finally, $f(L)$ is closed by Exercise 1.1 and consistency with the elementary functional calculus follows since for $f \in \mathcal{E}(\mathrm{S}_\varphi)$ we can take $e = \mathbf{1}$. □

Definition 5.13. *Let $\varphi \in (\varphi_L, \pi)$. For $f \in \mathrm{M}_L(\mathrm{S}_\varphi)$ define $f(L)$ as in* (5.2). *The emerging* functional calculus $\mathrm{M}_L(\mathrm{S}_\varphi) \ni f \mapsto f(L) \in \{$*closed operators in H*$\}$ *will (with a slight abuse of notation) again be denoted by Φ_L.*

The following consistency considerations based on Example 5.8 are left as Exercise 5.6.

Example 5.14. The functional calculus for the negative Laplacian from Chapter 4 also coincides with the extended functional calculus on $\mathrm{M}_{-\Delta}(\mathrm{S}_\varphi)$ for any $\varphi \in (0, \pi)$.

If the extended calculus should be useful in any way, we need to have something like an algebra homomorphism. Now the closed operators do not provide this algebraic structure (Exercise 1.1), so we have to be a little more modest in the next theorem, but we still get a powerful result.

Theorem 5.15. *Let $\varphi \in (\varphi_L, \pi)$. The set $\mathrm{M}_L(\mathrm{S}_\varphi)$ forms an algebra and for all $f, g \in \mathrm{M}_L(\mathrm{S}_\varphi)$ and $\lambda \in \mathbb{C}$ we have*

$$\lambda f(L) + g(L) \subseteq (\lambda f + g)(L) \quad \textit{and} \quad f(L)g(L) \subseteq (fg)(L).$$

Furthermore, if $g(L)$ is bounded, then there is equality in both inclusions and in the general case we have $\operatorname{dom}(f(L)g(L)) = \operatorname{dom}((fg)(L)) \cap \operatorname{dom}(g(L))$.

Proof. Let $f, g \in \mathrm{M}_L(\mathrm{S}_\varphi)$ and choose regularizers e_f and e_g for f and g, respectively. Then, by Remark 5.11, also $e := e_f e_g$ is a regularizer. Furthermore, since e_f, λe_g, $e_f f$ and $e_g g$ are all in $\mathcal{E}(\mathrm{S}_\varphi)$, we find

$$e(\lambda f + g) = (\lambda e_g)(e_f f) + e_f(e_g g) \in \mathcal{E}(\mathrm{S}_\varphi)$$

and

$$e(fg) = (e_f f)(e_g g) \in \mathcal{E}(\mathrm{S}_\varphi).$$

This means that e regularizes $\lambda f + g$ and fg, so these two functions are in $\mathrm{M}_L(\mathrm{S}_\varphi)$ and we can plug L into them via (5.2). Doing so, we find by the properties of the elementary functional calculus

$$\begin{aligned} \lambda f(L) + g(L) &= \lambda e(L)^{-1}(ef)(L) + e(L)^{-1}(eg)(L) \\ &\subseteq e(L)^{-1}\big(\lambda(ef)(L) + (eg)(L)\big) \\ &= e(L)^{-1}\big(e(\lambda f + g)\big)(L) = (\lambda f + g)(L). \end{aligned} \tag{5.3}$$

In the same manner we find

$$\begin{aligned} f(L)g(L) &= e_f(L)^{-1}\,(e_f f)(L)\,e_g(L)^{-1}\,(e_g g)(L) \\ &= e_f(L)^{-1}\left[e_g(L)^{-1} e_g(L)\,(e_f f)(L)\right] e_g(L)^{-1}\,(e_g g)(L) \\ &= e_f(L)^{-1}\,e_g(L)^{-1}\,(e_f f)(L)\left[e_g(L)\,e_g(L)^{-1}\right](e_g g)(L) \\ &\subseteq e(L)^{-1}\,(efg)(L) = (fg)(L). \end{aligned} \tag{5.4}$$

We take a closer look at the domains. The inclusion in (5.3) shows that

$$\operatorname{dom}(f(L)) \cap \operatorname{dom}(g(L)) = \operatorname{dom}\big(\lambda f(L) + g(L)\big) \subseteq \operatorname{dom}\big((\lambda f + g)(L)\big)$$

and

$$\begin{aligned} \operatorname{dom}\big((\lambda f + g)(L)\big) \cap \operatorname{dom}(g(L)) &= \operatorname{dom}\big((\lambda f + g)(L) - g(L)\big) \\ &\subseteq \operatorname{dom}\big(\lambda f + g - g)(L)\big) = \operatorname{dom}(f(L)). \end{aligned}$$

If we suppose that $g(L)$ is bounded, then $\operatorname{dom}(g(L)) = H$ and these two inclusions can be combined to

$$\operatorname{dom}(f(L)) = \operatorname{dom}\big(\lambda f(L) + g(L)\big) \subseteq \operatorname{dom}\big((\lambda f + g)(L)\big) \subseteq \operatorname{dom}(f(L)),$$

so all four sets must be equal.

For the domain of the product it suffices to prove the general assertion

$$\operatorname{dom}\big(f(L)g(L)\big) = \operatorname{dom}\big((fg)(L)\big) \cap \operatorname{dom}(g(L)). \tag{5.5}$$

Then, if $\mathrm{dom}(g(L)) = H$, we immediately get the asserted equality for the domains.

The inclusion in (5.4) yields $\mathrm{dom}(f(L)g(L)) \subseteq \mathrm{dom}((fg)(L))$ and we always have $\mathrm{dom}(f(L)g(L)) \subseteq \mathrm{dom}(g(L))$, so we have already proved '$\subseteq$' in (5.5).

For the reverse inclusion let $u \in \mathrm{dom}\big((fg)(L)\big) \cap \mathrm{dom}(g(L))$ be given. We have to show that

$$v := g(L)u \in \mathrm{dom}(f(L)) = \big\{w \in H \mid (e_f f)(L)w \in \mathrm{ran}(e_f(L))\big\}.$$

Using the definition of $g(L)$ as well as $u \in \mathrm{dom}((fg)(L))$, we find

$$\begin{aligned} e_g(L)\,(e_f f)(L)v &= (e_f f)(L)\,e_g(L)\,g(L)u = (e_f f)(L)\,(e_g g)(L)u \\ &= (e_g e_f f g)(L)u = e_g(L)\,e_f(L)\,(fg)(L)u. \end{aligned}$$

Since $e_g(L)$ is injective, we conclude

$$(e_f f)(L)v = e_f(L)\,(fg)(L)u \in \mathrm{ran}(e_f(L))$$

and we are done. □

As a first example of the extended functional calculus, we can answer the initial question about $f = \mathbf{z}$.

Corollary 5.16. *We have $(\lambda + \mathbf{z})(L) = \lambda + L$ for every $\lambda \in \mathbb{C}$.*

Proof. Let $\lambda \in \mathbb{C}$ and $f = 1 + \mathbf{z}$. Then $e = (1 + \mathbf{z})^{-1}$ regularizes f and by construction of the elementary functional calculus we get that

$$f(L) = e(L)^{-1}(ef)(L) = e(L)^{-1}\mathbf{1}(L) = ((1 + L)^{-1})^{-1} = 1 + L.$$

Since $\mathbf{1}(L) = \mathrm{id}_H$ is bounded, Theorem 5.15 yields

$$(\lambda + \mathbf{z})(L) = (f + (\lambda - 1)\mathbf{1})(L) = f(L) + (\lambda - 1) = \lambda + L.$$ □

As a second example, we show a commutation property.

Corollary 5.17. *Let $\varphi \in (\varphi_L, \pi)$ and $f \in \mathrm{M}_L(\mathrm{S}_\varphi)$ be such that $f(L) \in \mathcal{L}(H)$. Then $f(L)L \subseteq Lf(L)$ and in particular,* $\mathrm{dom}(L)$ *is invariant under $f(L)$.*

Proof. We use $\mathbf{z}(L) = L$ from Corollary 5.16 and apply Theorem 5.15 twice:

$$f(L)L = f(L)\,\mathbf{z}(L) \subseteq (f\,\mathbf{z})(L) = (\mathbf{z}\,f)(L) = \mathbf{z}(L)\,f(L) = Lf(L).$$

Note that in the second application of Theorem 5.15 we indeed get equality since $f(L)$ is bounded. The 'in particular' statement follows by the above inclusion and since $\mathrm{dom}(L) = \mathrm{dom}(f(L)L)$. □

Up to now we do not know how many new functions we actually get from extending the elementary functional calculus. Fortunately, already the regularizers known to us from Remark 5.11 reveal that $\mathrm{M}_L(\mathrm{S}_\varphi)$ is rather big.

Proposition 5.18. *Let $\varphi \in (\varphi_L, \pi)$ and $f\colon \mathrm{S}_\varphi \to \mathbb{C}$ be holomorphic. The following growth conditions guarantee that $f \in \mathrm{M}_L(\mathrm{S}_\varphi)$.*

(a) Polynomial limit at zero and polynomial control at infinity*: There exist $c \in \mathbb{C}$ and exponents $s > 0$ and $t \in \mathbb{R}$ such that*

$$\sup_{z \in \mathrm{S}_\varphi,\, |z| \le 1} \frac{|f(z) - c|}{|z|^s} < \infty \quad \textit{and} \quad \sup_{z \in \mathrm{S}_\varphi,\, |z| \ge 1} \frac{|f(z)|}{|z|^t} < \infty.$$

(b) Polynomial control at zero and infinity*: L is injective and there exist exponents $s, t \in \mathbb{R}$ such that*

$$\sup_{z \in \mathrm{S}_\varphi,\, |z| \le 1} |z|^s |f(z)| < \infty \quad \textit{and} \quad \sup_{z \in \mathrm{S}_\varphi,\, |z| \ge 1} \frac{|f(z)|}{|z|^t} < \infty.$$

Proof. In case (b) we know from Remark 5.11 that $e \coloneqq (\mathbf{z}(1+\mathbf{z})^{-2})^k$ is a regularizer for any $k \in \mathbb{N}$ and if $k > \max(s,t)$, then $ef \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$, meaning that e regularizes f. Likewise, in case (a) we take $e \coloneqq (1+\mathbf{z})^{-k}$ with $k \in \mathbb{N}$ such that $k > t$ and write

$$ef = (1+\mathbf{z})^{-k}(f - c) + c(1+\mathbf{z})^{-k}.$$

The first function on the right belongs to $\mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ thanks to the assumption on f, whereas the second one is in the algebra $\mathcal{E}(\mathrm{S}_\varphi)$. □

Example 5.19. An example for (a) is $f = \mathbf{z}^\alpha$ with $\operatorname{Re}\alpha > 0$. In (b) we can take any $f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)$. These functions will play a key role in the next lectures.

We close the lecture with a property for pairs of mutually inverse functions.

Proposition 5.20. *Let $\varphi \in (\varphi_L, \pi)$ and $f \in \mathrm{M}_L(\mathrm{S}_\varphi)$, $g \in \mathrm{M}(\mathrm{S}_\varphi)$ be such that $fg = \mathbf{1}$ on S_φ. Then $g \in \mathrm{M}_L(\mathrm{S}_\varphi)$ if and only if $f(L)$ is injective and in this case $g(L) = f(L)^{-1}$.*

Proof. '$\Longrightarrow$': We assume that $g \in \mathrm{M}_L(\mathrm{S}_\varphi)$.

Theorem 5.15 implies

$$g(L)\, f(L) \subseteq (gf)(L) = \mathbf{1}(L) = \mathrm{id}_H$$

with

$$\operatorname{dom}\bigl(g(L)\, f(L)\bigr) = \operatorname{dom}(\mathrm{id}_H) \cap \operatorname{dom}(f(L)) = \operatorname{dom}(f(L)).$$

This means that $f(L)$ is injective and that $f(L)^{-1} \subseteq g(L)$. Analogously, we also have

$$f(L)\, g(L) \subseteq \mathrm{id}_H \quad \text{with} \quad \operatorname{dom}\bigl(f(L)\, g(L)\bigr) = \operatorname{dom}(g(L)),$$

from which we conclude $g(L) \subseteq f(L)^{-1}$.

'$\Longleftarrow$': *We assume that $f(L)$ is injective.*

Let e regularize f. Then, thanks to Theorem 5.15 and since $e(L)$ is bounded, we get $(ef)(L) = (fe)(L) = f(L)\, e(L)$ and as a product of two injective operators, this is an injective operator. This shows that ef is again a regularizer and it even regularizes g, since $efg = e \in \mathcal{E}(\mathrm{S}_\varphi)$. $\square$

Example 5.21. Let $\varphi \in (\varphi_L, \pi)$ and $\lambda \in \varrho(L)$. We define $f, g \in \mathrm{M}(\mathrm{S}_\varphi)$ by $f := \lambda - \mathbf{z}$ and $g := (\lambda - \mathbf{z})^{-1}$, where g could have its pole inside of S_φ. We have $f \in \mathrm{M}_L(\mathrm{S}_\varphi)$ by Proposition 5.18. According to Corollary 5.16, we have $f(L) = \lambda - L$ and this operator is injective, since $\lambda \in \varrho(L)$. Thus, Proposition 5.20 applies and we find $g \in \mathrm{M}_L(\mathrm{S}_\varphi)$ and, as expected,

$$(\lambda - \mathbf{z})^{-1}(L) = \big(\lambda - \mathbf{z}(L)\big)^{-1} = (\lambda - L)^{-1}.$$

If $\lambda \in \mathbb{C} \setminus \overline{\mathrm{S}_\varphi}$, then $g \in \mathcal{E}(\mathrm{S}_\varphi)$ and the above can also be proved directly from the definition of the elementary functional calculus, see Exercise 5.2.

5.3. Exercises

Exercise 5.1. Modify the proof of Lemma 5.4 in order to show that the definition of $f(L)$ in Definition 5.5 does not depend on the particular choice of ψ.

Exercise 5.2. Let $\varphi \in (\varphi_L, \pi)$ and $\lambda \in \mathbb{C} \setminus \overline{\mathrm{S}_\varphi}$. Prove directly from Definition 5.6 and without resorting to the extended functional calculus that $(\lambda - \mathbf{z})^{-1}(L) = (\lambda - L)^{-1}$.

Exercise 5.3 (A characterization of the Dunford–Riesz class)**.** Let $\varphi \in (0, \pi)$ and let $f \colon \mathrm{S}_\varphi \to \mathbb{C}$ be holomorphic. We say that f has *polynomial limits at zero and infinity* if there are $c, d \in \mathbb{C}$ and exponents $s, t > 0$ such that

$$\sup_{z \in \mathrm{S}_\varphi,\, |z| \le 1} \frac{|f(z) - c|}{|z|^s} < \infty \quad \text{and} \quad \sup_{z \in \mathrm{S}_\varphi,\, |z| \ge 1} |z|^t |f(z) - d| < \infty,$$

compare with Proposition 5.18.

(a) Prove that $f \in \mathcal{E}(\mathrm{S}_\varphi)$ if and only if f has polynomial limits at zero and infinity.

(b) Conclude that the three subspaces that build up the Dunford–Riesz class are linearly independent.

Exercise 5.4 (Commuting operators in the functional calculus)**.** Let $\varphi \in (\varphi_L, \pi)$ and suppose that $T \in \mathcal{L}(H)$ commutes with L, i.e., $TL \subseteq LT$.

(a) Prove $f(L)T = Tf(L)$ for all $f \in \mathcal{E}(\mathrm{S}_\varphi)$.

(b) Conclude that $Tf(L) \subseteq f(L)T$ holds for every $f \in \mathrm{M}_L(\mathrm{S}_\varphi)$.

Exercise 5.5. In this exercise we investigate the functional calculus on the Hilbert space $\ker(L)$, where L acts as the (sectorial, though not particularly interesting) zero operator. Let $\varphi \in (\varphi_L, \pi)$ and suppose that the holomorphic function $f : \mathrm{S}_\varphi \to \mathbb{C}$ has a polynomial limit c at zero and polynomial control at infinity.

(a) Make a guess, what $f(L)u$ should be whenever $u \in \ker(L)$ and don't read further.

(b) Prove your guess.[1]

Exercise 5.6. Prove consistency of the extended functional calculus for the negative Laplacian on $\mathbb{R}^n$ as stated in Example 5.14.

[1]That is, prove $f(L)u = cu$ for $u \in \ker(L)$.

6. First applications of functional calculus

In this lecture we will study concrete examples of operators that emerge from the functional calculus in Definition 5.13. In particular, we will see fractional powers of sectorial operators and we will confirm that through the functional calculus (and with a little care) we can 'manipulate operators as if they were functions in the complex plane'. The general setup is the same as in Lecture 5.

Notation 6.1. Throughout the entire lecture, L denotes a sectorial operator in a Hilbert space H with inner product $\langle \cdot\, , \cdot \rangle$ and norm $\| \cdot \|$.

6.1. Fractional powers

We fix the principal branch of the complex logarithm. Then, for all $\alpha \in \mathbb{C}$ with $\operatorname{Re}\alpha > 0$, the fractional power function $\mathbf{z}^\alpha = \mathrm{e}^{\alpha \log(\mathbf{z})}$ has a polynomial limit at zero and polynomial control at infinity in the sense of Proposition 5.18 (a) on S_φ for any $\varphi \in (0, \pi)$. More precisely, we can take $c = 0$ and $s = t = \operatorname{Re}\alpha$ in Proposition 5.18 (a). Hence, $(1 + \mathbf{z})^{-n}$ regularizes $\mathbf{z}^\alpha$ provided that $n > \operatorname{Re}\alpha$.

Definition 6.2. *For $\alpha \in \mathbb{C}$ with $\operatorname{Re}\alpha > 0$ define $L^\alpha := \mathbf{z}^\alpha(L)$, called* fractional powers *of L.*

According to Example 5.14 (see also Exercise 5.6), this definition agrees with our earlier one for the negative Laplacian on $\mathbb{R}^n$ (Definition 4.10) and according to Corollary 5.16, there is no ambiguity in the definition when $\alpha = 1$.

Proposition 6.3. *Let $\alpha, \beta \in \mathbb{C}$ with $\operatorname{Re}\alpha, \operatorname{Re}\beta > 0$. The fractional powers have the following properties:*

(a) $L^\alpha L^\beta = L^{\alpha+\beta}$ $(= L^\beta L^\alpha)$.

(b) If $k \in \mathbb{N}$ is such that $k > \operatorname{Re}\alpha$, then $\operatorname{dom}(L^k)$ is a core for $\operatorname{dom}(L^\alpha)$. If, in addition, L is injective, then also $\operatorname{dom}(L^k) \cap \operatorname{ran}(L^k)$ is a core.

Proof. We begin with (a). The general properties of the functional calculus in Theorem 5.15 yield $L^\alpha L^\beta \subseteq L^{\alpha+\beta}$ with $\operatorname{dom}(L^\alpha L^\beta) = \operatorname{dom}(L^{\alpha+\beta}) \cap \operatorname{dom}(L^\beta)$ and it remains to show that $\operatorname{dom}(L^{\alpha+\beta}) \subseteq \operatorname{dom}(L^\beta)$.

To this end, we fix an integer $k > \max(\operatorname{Re}\alpha, \operatorname{Re}\beta)$ and use $(1+\mathbf{z})^{-k}$ as a regularizer for $\mathbf{z}^\alpha$ and $\mathbf{z}^\beta$, noting that $(1+\mathbf{z})^{-2k}$ regularizes $\mathbf{z}^{\alpha+\beta}$. Now, let $u \in \operatorname{dom}(L^{\alpha+\beta})$. By construction, $u \in \operatorname{dom}(L^\beta)$ is equivalent to

$$v \coloneqq (\mathbf{z}^\beta(1+\mathbf{z})^{-k})(L)u \in \operatorname{ran}((1+L)^{-k}) = \operatorname{dom}(L^k) \tag{6.1}$$

and it is this property that we are going to check. In order to bring $\mathbf{z}^{\alpha+\beta}$ into play, we 'regularize' v as follows, using Proposition 5.9 in the first three steps:

$$\begin{aligned}(L(1+L)^{-1})^k(1+L)^{-k}v &= \left(\frac{\mathbf{z}^k}{(1+\mathbf{z})^{2k}}\right)(L)v \\ &= \left(\frac{\mathbf{z}^{\beta+k}}{(1+\mathbf{z})^{3k}}\right)(L)u \\ &= \left(\frac{\mathbf{z}^{k-\alpha}}{(1+\mathbf{z})^{k}}\right)(L)\left(\frac{\mathbf{z}^{\alpha+\beta}}{(1+\mathbf{z})^{2k}}\right)(L)u \\ &=: \left(\frac{\mathbf{z}^{k-\alpha}}{(1+\mathbf{z})^{k}}\right)(L)w,\end{aligned}$$

where, similar to (6.1), we have $w \in \operatorname{dom}(L^{2k})$ by assumption on u. Applying Corollary 5.17 repeatedly, we conclude that the whole expression above is contained in $\operatorname{dom}(L^{2k})$. Now, we have to remove the 'regularization' of v: Exercise 1.3 yields $(1+L)^{-k}v \in \operatorname{dom}(L^{2k})$ and (6.1) follows.

Assertion (b) is an example of a more general property of the functional calculus that we state and prove next. □

Lemma 6.4. *Let $\varphi \in (\varphi_L, \pi)$ and $f \in \mathrm{M}_L(\mathrm{S}_\varphi)$. The following spaces are cores for $f(L)$.*

(a) $\operatorname{dom}(L^k)$ *with $k \in \mathbb{N}$ provided that $(1+\mathbf{z})^{-k}$ regularizes f.*

(b) $\operatorname{dom}(L^k) \cap \operatorname{ran}(L^k)$ *with $k \in \mathbb{N}$ provided that L is injective and $(\mathbf{z}(1+\mathbf{z})^{-2})^k$ regularizes f.*

In particular, in these cases $f(L)$ is densely defined.

Proof. In order to prove (a), we first show $\operatorname{dom}(L^k) \subseteq \operatorname{dom}(f(L))$. We use the regularizer $e \coloneqq (1+\mathbf{z})^{-k}$ for f. The first observation is that $\operatorname{dom}(L^k) = \operatorname{ran}(e(L))$ is the range of this particular regularizer. Taking $u = e(L)w$ in this set, we obtain

$$(ef)(L)u = (ef)(L)e(L)w = e(L)(ef)(L)w \in \operatorname{ran}(e(L)),$$

so $u \in \operatorname{dom}(f(L))$.

For the claim that $\operatorname{dom}(L^k)$ is a core for $\operatorname{dom}(f(L))$, we use the functions $e_t \coloneqq (1+t\mathbf{z})^{-k}$ with $t > 0$. Under the functional calculus (Example 5.21) they correspond to the approximate identities $e_t(L) = (1+tL)^{-k}$ that already appeared in Proposition 2.4 (d). Now, take $u \in \operatorname{dom}(f(L))$. We have $e_t(L)u \in \operatorname{dom}(L^k) \subseteq \operatorname{dom}(f(L))$ and $e_t(L)u \to u$ in H in the limit as $t \searrow 0$. In the same manner, using Theorem 5.15, we get

$$f(L)e_t(L)u = (fe_t)(L)u = (e_t f)(L)u = e_t(L)f(L)u \to f(L)u$$

in H. This means that $e_t(L)u$ tends to u in $\operatorname{dom}(f(L))$.

The proof of (b) is the same upon using $e = (\mathbf{z}(1+\mathbf{z})^{-2})^k$ and

$$e_t := (1+t\mathbf{z})^{-k}\big(t^{-1}\mathbf{z}(1+t^{-1}\mathbf{z})^{-1}\big)^k = (1+t\mathbf{z})^{-k}\big(1-(1+t^{-1}\mathbf{z})^{-1}\big)^k,$$

once we have noticed that $\operatorname{ran}(e(L)) = \operatorname{dom}(L^k) \cap \operatorname{ran}(L^k)$ by Proposition 2.4 (e) and that $\operatorname{ran}(e_t(L)) = \operatorname{ran}(e(L))$.

Finally, $f(L)$ is densely defined since the respective cores are dense in H, see again Proposition 2.4 (e). □

It follows from Proposition 6.3 (a) that L^n is unambiguously defined when $n \in \mathbb{N}$. We do not know anything very specific about fractional power domains at this point, but Proposition 6.3 (b) provides some link with the domains of integer powers of L. We will give more explicit formulæ to compute fractional powers later in this lecture.

Remark 6.5. If L is injective, then $\mathbf{z}^\alpha \in \mathrm{M}_L(\mathrm{S}_\varphi)$ even for any $\alpha \in \mathbb{C}$, because these functions have polynomial control at zero and infinity, see Proposition 5.18. This gives rise to fractional powers $L^\alpha := \mathbf{z}^\alpha(L)$ for $\alpha \in \mathbb{C}$. Proposition 5.20 yields that all these operators are injective and that $(L^\alpha)^{-1} = L^{-\alpha}$ holds as expected.

6.2. The exponential function and semigroups

In your lectures on ordinary differential equations you have learned that the matrix exponential function can be used to solve the initial value problem for linear systems with constant coefficients. In retrospect, this was probably your first encounter with a functional calculus!

If $\varphi_L < \pi/2$, then also for L we can define an exponential function

$$\mathrm{e}^{-tL} := \mathrm{e}^{-t\mathbf{z}}(L) \qquad (t > 0)$$

via the elementary functional calculus, compare with Remark 5.3 (a), and, given $u_0 \in H$, the function $u(t) := \mathrm{e}^{-tL}u_0$ should be a solution to

$$\begin{cases} u'(t) + Lu(t) = 0 & (t > 0), \\ u(0) = u_0. \end{cases} \tag{6.2}$$

If, for instance, $L = (-\Delta)_{\mathrm{H}_0^1(\Omega)}$ is the negative Dirichlet Laplacian on an open set Ω, then (6.2) formally corresponds to the initial/boundary value problem for the heat equation,

$$\begin{cases} \partial_t u(t,x) - \Delta u(t,x) = 0 & (x \in \Omega,\ t > 0), \\ \qquad\qquad\quad u(t,x) = 0 & (x \in \partial\Omega,\ t > 0), \\ \qquad\qquad\quad u(0,x) = u_0(x) & (x \in \Omega). \end{cases} \tag{6.3}$$

Approaching partial differential equations such as (6.3) via an *abstract Cauchy problem* such as (6.2) works surprisingly well and quickly leads to the concept of *maximal* L^p *regularity* that connects operator theory, harmonic analysis and the geometry of Banach spaces. You will have the chance to learn more about it in the second and third phase of the Internet Seminar. Here, we confine ourselves to showing that we can indeed solve (6.2) using functional calculus and suggest the proof of the following result as Exercise 6.3.

Proposition 6.6. *Suppose that $\varphi_L < \pi/2$. Define the family $(\mathrm{e}^{-tL})_{t>0}$, called* holomorphic semigroup generated by $-L$,[1] *as above and let $u_0 \in H$. The following properties hold.*

(a) Semigroup property: $\mathrm{e}^{-sL}\mathrm{e}^{-tL} = \mathrm{e}^{-(s+t)L}$ *for all $s, t > 0$.*

(b) Strong continuity at 0*:* $\lim_{t\to 0} \mathrm{e}^{-tL}u_0 = u_0$.

(c) Long-time behavior: $\lim_{t\to\infty} \mathrm{e}^{-tL}u_0 = Pu_0$, *where P is the projection onto* $\ker(L)$ *along* $\overline{\operatorname{ran}(L)}$.

(d) Solution to the abstract Cauchy problem: $\mathrm{e}^{-\cdot L} : (0,\infty) \to \mathcal{L}(H)$ *is continuously differentiable and* $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{-tL} = -L\mathrm{e}^{-tL}$.

It may well be that semigroups and their relation to PDEs now ring a bell: In fact, you have seen a different example when solving the Dirichlet problem on the upper half-space in Exercise 4.5.

6.3. The adjoint calculus

The class of sectorial operators is stable under taking adjoints in the following sense.

Lemma 6.7. *L^* is sectorial of angle $\varphi_{L^*} = \varphi_L$ and if L is injective, then so is L^*.*

Proof. Since sectors are invariant under complex conjugation, it follows from Propositions 1.21 and 1.20 (h) that L^* is sectorial with angle $\varphi_{L^*} \le \varphi_L$ and that the resolvents

[1] In order to understand this terminology, you will also have to solve Exercise 6.4.

are related by the formula

$$(\overline{\lambda} - L^*)^{-1} = ((\lambda - L)^{-1})^*. \tag{6.4}$$

Applying this result to L^* with adjoint L gives $\varphi_L \leq \varphi_{L^*}$. Since $\overline{\operatorname{ran}(L)} = \ker(L^*)^\perp$, see Proposition 1.20, we obtain from Proposition 2.4 (b) that L^* is injective provided that L is injective. $\square$

In terms of functional calculus, (6.4) can be seen as an example of a rule

$$f^*(L^*) = f(L)^* \tag{6.5}$$

lurking somewhere in the background.

Definition 6.8. *Let $\varphi \in (0, \pi)$ and $f : \mathrm{S}_\varphi \to \mathbb{C}$ be holomorphic. The function $f^* : \mathrm{S}_\varphi \to \mathbb{C}$, $z \mapsto \overline{f(\overline{z})}$ is called* holomorphic conjugate *of f.*

The holomorphic conjugate is again holomorphic and the mapping $f \mapsto f^*$ preserves the Dunford–Riesz class $\mathcal{E}(\mathrm{S}_\varphi)$ and, more generally, the classes of holomorphic functions with polynomial limits/control considered in Proposition 5.18. If there is any justice, also (6.5) will continue to hold.

First, we put all subtle domain considerations aside and look at the elementary functional calculus.

Lemma 6.9. *Let $\varphi \in (\varphi_L, \pi)$. Then* (6.5) *holds for all $f \in \mathcal{E}(\mathrm{S}_\varphi)$.*

Proof. For $f = \mathbf{1}$ there is nothing to prove and above we have already dealt with $f = (1 + \mathbf{z})^{-1}$. It remains to treat the case $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$.

We fix $\psi \in (\varphi_L, \varphi)$ and represent $f(L)$ as in Definition 5.5. According to Proposition 1.20, the map $\mathcal{L}(H) \ni K \mapsto K^* \in \mathcal{L}(H)$ is bounded and anti-linear. We first use Proposition A.13 to get

$$\begin{aligned} f(L)^* &= \left(\frac{1}{2\pi \mathrm{i}} \int_{\mathbb{R}} f(\gamma_\psi(t)) \cdot \gamma_\psi'(t) \cdot (\gamma_\psi(t) - L)^{-1} \,\mathrm{d}t \right)^* \\ &= \frac{-1}{2\pi \mathrm{i}} \int_{\mathbb{R}} \left(f(\gamma_\psi(t)) \cdot \gamma_\psi'(t) \cdot (\gamma_\psi(t) - L)^{-1} \right)^* \mathrm{d}t. \end{aligned}$$

Now, anti-linearity lets us continue by

$$\begin{aligned} &= \frac{-1}{2\pi \mathrm{i}} \int_{\mathbb{R}} \overline{f(\gamma_\psi(t))} \cdot \overline{\gamma_\psi}'(t) \cdot (\overline{\gamma_\psi}(t) - L^*)^{-1} \,\mathrm{d}t \\ &= \frac{-1}{2\pi \mathrm{i}} \int_{\mathbb{R}} f^*(\overline{\gamma_\psi}(t)) \cdot \overline{\gamma_\psi}'(t) \cdot (\overline{\gamma_\psi}(t) - L^*)^{-1} \,\mathrm{d}t \\ &= f^*(L^*), \end{aligned}$$

where we have used $\overline{\gamma_\psi}(t) = \gamma_\psi(-t)$ in the final step. $\square$

Our general result is as follows. It will become apparent from the proof that the extension of (6.5) does not work by purely algebraic means and hence, we only treat the subclasses of $\mathrm{M}_L(\mathrm{S}_\varphi)$ that have been described in Proposition 5.18.

Proposition 6.10. *Let $\varphi \in (\varphi_L, \pi)$ and $f \colon \mathrm{S}_\varphi \to \mathbb{C}$ be holomorphic. The rule* (6.5) *is valid in the following cases.*

(a) f has a polynomial limit at zero and polynomial control at infinity.

(b) L is injective and f has polynomial control at zero and infinity.

Proof. Let e be a suitable regularizer for f in the calculus for L: $e := (1+\mathbf{z})^{-k}$ in case (a) and $e := (\mathbf{z}(1+\mathbf{z})^{-2})^k$ in case (b), both with k sufficiently large, see Proposition 5.18. Then e also regularizes f in the calculus for L^*, using Lemma 6.7 in case (b). We treat both cases simultaneously.

Starting from the definition of $f(L)$, we obtain from Lemma 6.9 and $e^* = e$ that

$$f^*(L^*) = e(L^*)^{-1}(ef^*)(L^*) = (e(L)^*)^{-1}(ef)(L)^*.$$

The calculation rules (e) and (d) in Proposition 1.20 yield

$$= (e(L)^{-1})^*(ef)(L)^* = \big((ef)(L)e(L)^{-1}\big)^*. \tag{6.6}$$

Comparing the outcome with the definition of $f(L)$, we see that $e(L)^{-1}$ is on the wrong side of $(ef)(L)$ and this is precisely what Lemma 6.4 will correct for us.

Indeed, $(ef)(L) = f(L)e(L)$ holds by Theorem 5.15 and therefore $(ef)(L)e(L)^{-1}u = f(L)u$ for $u \in \operatorname{ran}(e(L)) = \operatorname{dom}((ef)(L)e(L)^{-1})$. However, the lemma says that the latter is a core for $f(L)$, hence $\overline{(ef)(L)e(L)^{-1}} = f(L)$ by Exercise 1.2 (c). Going back to (6.6), we can now use calculation rule (f) from Proposition 1.20 to derive the desired identity

$$f^*(L^*) = \big(\overline{(ef)(L)e(L)^{-1}}\big)^* = f(L)^*. \qquad \square$$

Corollary 6.11. *If $\alpha > 0$, then $(L^\alpha)^* = (L^*)^\alpha$.*

Proof. Simply note that $\mathbf{z}^\alpha$ is its own holomorphic conjugate. $\square$

6.4. Kato's second representation theorem

In this short section we prove a remarkable theorem due to Kato. As in Lecture 2 we let V be another Hilbert space that is continuously and densely embedded into H.

If L is the self-adjoint operator associated with a bounded, accretive, elliptic and symmetric sesquilinear form on V, then we can determine the domain of $\sqrt{L}$ and this operator appears naturally in order to represent the form on V via the inner product on H, hence the name. Here is the result.

Theorem 6.12 (Kato's second representation theorem)**.** *Let a be a bounded, accretive, elliptic and symmetric sesquilinear form on V and let L be the associated self-adjoint operator in H. Then* $\operatorname{dom}(\sqrt{L}) = V$ *and*

$$a(u, v) = \langle \sqrt{L}u, \sqrt{L}v \rangle \qquad (u, v \in V).$$

Proof. First, we let $u, v \in \operatorname{dom}(L)$. By definition of L, Proposition 6.3 and Corollary 6.11 we get

$$a(u, v) = \langle Lu, v \rangle = \langle \sqrt{L}\sqrt{L}u, v \rangle = \langle \sqrt{L}u, \sqrt{L}v \rangle. \tag{6.7}$$

Since a is bounded and elliptic, we conclude

$$\|\sqrt{L}u\|^2 + \|u\|^2 = a(u) + \|u\|^2 \lesssim \|u\|_V^2 \lesssim \operatorname{Re}\big(a(u)\big) + \|u\|^2 = \|\sqrt{L}u\|^2 + \|u\|^2,$$

which means that the Hilbert space norms on V and $\operatorname{dom}(\sqrt{L})$ are equivalent on the common subspace $\operatorname{dom}(L)$. The latter is dense in both V and $\operatorname{dom}(\sqrt{L})$, see Propositions 2.23 and 6.3. Thus, $V = \operatorname{dom}(\sqrt{L})$ and (6.7) extends to all $u, v \in V$ by density. □

This provides us with a far-reaching generalization of the Kato property from Proposition 4.12 for elliptic operators in divergence form. Recall from Lecture 3 that such operators are self-adjoint if the coefficients are Hermitean and that $a(u, u) \simeq \|\nabla u\|_{\mathrm{L}^2(\Omega)}^2$ for all $u \in \mathrm{H}_0^1(\Omega)$ since the coefficients are bounded and elliptic.

Corollary 6.13 (Kato property for self-adjoint elliptic operators)**.** *Let $L = -\operatorname{div}(A\nabla\,\cdot)$ be an elliptic operator in divergence form with Dirichlet boundary conditions on an open set Ω. In addition, suppose that $A(x)$ is Hermitean for a.e. $x \in \Omega$. Then* $\operatorname{dom}(\sqrt{L}) = \mathrm{H}_0^1(\Omega)$ *and*

$$\|\sqrt{L}u\|_{\mathrm{L}^2(\Omega)} \simeq \|\nabla u\|_{\mathrm{L}^2(\Omega)} \qquad (u \in \mathrm{H}_0^1(\Omega)).$$

The question whether the condition $A = A^*$ can be dropped in Corollary 6.13 became known as the 'Kato conjecture'. Pascal Auscher remarks at the beginning of his beautiful essay on the mathematical œuvre of Yves Meyer [Aus23]:

> *"Tosio Kato's square root conjecture is one example of a question arising from one field, formulated in a second one and finding its solution in a third one. Namely, the question arising in the sixties from the work of T. Kato, motivated by partial differential equations in inhomogeneous media, was set using a functional analysis framework and it was finally methods from real harmonic analysis that put a final end to the problem as posed."*

In fact, while studying wave propagation in inhomogeneous media (the conductivity of which is related to the real and self-adjoint matrices $A(x)$), Kato wanted to understand the stability of (12.1) with respect to small L^∞-perturbations of A on $\mathbb{R}^n$: Could it be true that

$$\|\sqrt{L_0}u - \sqrt{L_1}u\|_{\mathrm{L}^2(\mathbb{R}^n)} \lesssim \|A_0 - A_1\|_{\mathrm{L}^\infty(\mathbb{R}^n)}\|\nabla u\|_{\mathrm{L}^2(\mathbb{R}^n)}, \tag{6.8}$$

provided that $\|A_0 - A_1\|_{\mathrm{L}^\infty(\mathbb{R}^n)}$ is sufficiently small? It was known that proving (12.1) for *all* complex, elliptic coefficients, using only L^∞-information on A, would automatically give holomorphic dependence in a suitable sense, and hence (6.8).

It took almost 40 years until eventually Auscher–Hofmann–Lacey–McIntosh–Tchamitchian [AHL⁺02] confirmed the Kato conjecture . . . but remarkable mathematical developments paved the way: The one-dimensional case turned out to be essentially equivalent to the L^2-boundedness of the Cauchy integral on a Lipschitz curve due to Coifman–McIntosh–Meyer [CMM82] and led to the study of anti-symmetric singular integrals and the famous $T(1)$- and $T(b)$-theorems of David–Journé–Semmes [DJ84, DJS85] in real harmonic analysis. In our personal story, we will need five more lectures until we have the necessary tools at hand and can present the proof of the Kato conjecture.

Functional calculus, as an ‘algebraic’ framework for Kato’s conjecture, evolved at the same time. Most notably, the development of H^∞-calculus and its close relation to harmonic analysis is one of Alan McIntosh’s mathematical legacies. We will touch upon these topics in the next two lectures.

6.5. The Calderón reproducing formula

As a first showcase how functional calculus can be seen as an L-adapted Fourier analysis, we extend the Calderón reproducing formula in its formulation (4.10) to general sectorial operators. The class of functions f changes as an artefact of using two very different methods of proof: In Lecture 4 we have used $\mathcal{S}(\mathbb{R}^n)$, which is well-adapted to the Fourier transform, whereas here and in the upcoming lectures we will use $\mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ related to functional calculus.

We begin with a simple lemma that shows that the reproducing formula somehow deals with a singular integral, that is, a uniformly bounded function on $(0,\infty)$ integrated against the measure ${}^{\mathrm{d}t}/_t$.

Lemma 6.14. *If $\varphi \in (\varphi_L, \pi)$ and $f \in \mathcal{E}(\mathrm{S}_\varphi)$, then the map*

$$(0,\infty) \to \mathcal{L}(H), \quad t \mapsto f(tL)$$

is continuous and (uniformly) bounded.

Proof. There is nothing to do for $f = \mathbf{1}$. For $f = (1+\mathbf{z})^{-1}$ or more generally, $f = (\lambda - \mathbf{z})^{-1}$ with $\lambda \in \mathbb{C} \setminus \overline{\mathrm{S}_{\varphi_L}}$, we have

$$(\lambda - t\mathbf{z})^{-1}(L) = (\lambda - tL)^{-1}. \tag{6.9}$$

Continuity in t follows by holomorphy of the resolvent and uniform boundedness by sectoriality of L.

Finally, if $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$, then also $f \circ t\mathbf{z} \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ and thanks to the bounds on f and the sectoriality of L, we control the Cauchy integral defining $f(tL)$ uniformly by

$$\|f(tL)\|_{\mathcal{L}(H)} \leq \int_0^\infty C \min(|t\tau|^s, |t\tau|^{-s}) \, \frac{\mathrm{d}\tau}{\tau} \overset{t\tau=\rho}{=} \int_0^\infty C \min(|\rho|^s, |\rho|^{-s}) \, \frac{\mathrm{d}\rho}{\rho} < \infty,$$

where $C, s > 0$ are independent of t. Continuity in t follows by dominated convergence, using the bound $|f(tz)| \leq C \max(t^s, t^{-s}) \min(|z|^s, |z|^{-s})$. □

Remark 6.15. Above, we have intuitively understood $f(tL) := (f \circ t\mathbf{z})(L)$ through the calculus for L. The identity (6.9) implies that $L_t := tL$ is sectorial with the same angle as L and it is just as natural to understand $f(tL) := f(L_t)$ through the calculus for L_t. Fortunately, both interpretations give the same operator: For $f = \mathbf{1}$ this is by definition, for $f = (1+\mathbf{z})^{-1}$ it follows directly from (6.9) and for $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ we use the same identity and change variables in the Cauchy integral that defines $f(L_t)$. This is a very simple example of a *composition rule* for functional calculi. We refer to [Haa06, Sect. 1.3.2] for more.

We are ready for the main result in this section.

Theorem 6.16 (Calderón reproducing formula)**.** *Let $\varphi \in (\varphi_L, \pi)$ and let $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ be normalized to*

$$\int_0^\infty f(t) \, \frac{\mathrm{d}t}{t} = 1. \tag{6.10}$$

Then we have for all $u \in \overline{\mathrm{ran}(L)}$ that

$$\int_0^\infty f(tL)u \, \frac{\mathrm{d}t}{t} = u$$

as an improper integral[2] in H. In particular, if L is injective, then the statement is valid for all $u \in H$.

Proof. The 'in particular' statement on injective operators follows from Proposition 2.4 (b). Throughout the proof we fix $C, s > 0$ such that $|f(z)| \leq C \min(|z|^s, |z|^{-s})$

[2] In other words, '$\int_0^\infty$' should be read '$\lim_{\substack{\varepsilon \to 0 \\ R \to \infty}} \int_\varepsilon^R$' as in Proposition 4.15.

for all $z \in \mathrm{S}_\varphi$. First of all, we define for $0 < a < b$ the function

$$F_{a,b} \colon \mathrm{S}_\varphi \to \mathbb{C}, \quad z \mapsto \int_a^b f(tz) \, \frac{\mathrm{d}t}{t}.$$

There are many ways to see that it is holomorphic – one is to use Morera's theorem and change the order of integration. The rest of the argument comes in three steps.

Step 1: Elementary properties of $F_{a,b}$.

For any $0 < a < b$ the estimate

$$|F_{a,b}(z)| \leq \int_a^b |f(tz)| \, \frac{\mathrm{d}t}{t} \leq C \int_a^b \min(|tz|^s, |tz|^{-s}) \, \frac{\mathrm{d}t}{t}$$

shows that $F_{a,b} \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$. Moreover, the substitution $\tau = |z|t$ gives

$$|F_{a,b}(z)| \leq C \int_0^\infty \min(\tau^s, \tau^{-s}) \, \frac{\mathrm{d}\tau}{\tau}$$

and, consequently, $(F_{a,b})_{0<a<b}$ is a bounded family in $\mathrm{H}^\infty(\mathrm{S}_\varphi)$. By dominated convergence, the limit

$$F(z) := \int_0^\infty f(tz) \, \frac{\mathrm{d}t}{t} = \lim_{\substack{a \to 0 \\ b \to \infty}} F_{a,b}(z)$$

exists uniformly for z in compact subsets of S_φ. In particular, F is holomorphic. For $z \in (0, \infty)$ we can change again variables $\tau = tz$ to see that $F(z) = 1$ and the identity theorem implies the same for all $z \in \mathrm{S}_\varphi$.

Step 2: Identification of $F_{a,b}(L)$.

We represent $F_{a,b}(L)$ for some $\psi \in (\varphi_L, \varphi)$ via a Cauchy integral and use Fubini's theorem to conclude that

$$F_{a,b}(L) = \frac{1}{2\pi \mathrm{i}} \int_a^b \int_{\gamma_\psi} f(tz)(z - L)^{-1} \, \mathrm{d}z \, \frac{\mathrm{d}t}{t} = \int_a^b f(tL) \, \frac{\mathrm{d}t}{t}.$$

Step 3: Convergence on $\mathrm{dom}(L) \cap \mathrm{ran}(L)$.

Let $u \in \mathrm{dom}(L) \cap \mathrm{ran}(L)$ and use Proposition 2.4 (e) to write $u = e(L)v$ for some $v \in H$, where $e := \mathbf{z}(1 + \mathbf{z})^{-2}$. Step 2 and the elementary properties of the functional calculus imply

$$\int_a^b f(tL)u \, \frac{\mathrm{d}t}{t} = F_{a,b}(L)e(L)v = (F_{a,b}e)(L)v = \frac{1}{2\pi \mathrm{i}} \int_{\gamma_\psi} F_{a,b}(z)e(z)(z - L)^{-1} v \, \mathrm{d}z.$$

The functions $F_{a,b}$ are uniformly bounded with respect to a, b and tend to 1 in the limit as $a \to 0, b \to \infty$, see Step 1. Thanks to the additional decay of the function e, we can use dominated convergence to get, as desired,

$$\lim_{\substack{a\to 0\\ b\to\infty}} \int_a^b f(tL)u \,\frac{\mathrm{d}t}{t} = \frac{1}{2\pi \mathrm{i}} \int_{\gamma_\psi} e(z)(z-L)^{-1}v \,\mathrm{d}z = e(L)v = u.$$

Step 4: Convergence on $\overline{\operatorname{ran}(L)}$.

We recall from Proposition 2.4 (e) that $\overline{\operatorname{ran}(L)} = \overline{\operatorname{dom}(L)\cap\operatorname{ran}(L)}$. Hence, and thanks to Step 2, it suffices to show that the family $(F_{a,b}(L))_{0<a<b}$ is bounded in $\mathcal{L}(H)$.[3] For this purpose, we write

$$F_{a,b}(z) = \int_0^1 f(tbz)\,\frac{\mathrm{d}t}{t} - \int_0^1 f(taz)\,\frac{\mathrm{d}t}{t} =: G(bz) - G(az).$$

We have

$$|G(z)| \leq \int_0^1 C|tz|^s \,\frac{\mathrm{d}t}{t} = \frac{C}{s}|z|^s,$$

but due to Step 1, we can also bound

$$|G(z)-1| = \left|\int_1^\infty f(tz)\,\frac{\mathrm{d}t}{t}\right| \leq \int_1^\infty C|tz|^{-s}\,\frac{\mathrm{d}t}{t} = \frac{C}{s}|z|^{-s}.$$

This proves that G has polynomial limits at zero and infinity in the sense of Exercise 5.3 and thus that $G \in \mathcal{E}(\mathrm{S}_\varphi)$. Uniform boundedness of $(F_{a,b}(L))_{0<a<b}$ is now a consequence of Lemma 6.14 and the proof is complete. □

A typical application for the Calderón reproducing formula will be as follows. We are given some $g \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)\setminus\{0\}$ and we know, for whatever reason, something good about $g(tL)$. Then we pick an $h \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ with

$$\int_0^\infty h(t)g(t)\,\frac{\mathrm{d}t}{t} = 1, \tag{6.11}$$

that is, $f := gh$ satisfies the assumption of Theorem 6.16, and we get the approximation of the identity that involves our favorable function g:

$$u = \int_0^\infty h(tL)g(tL)u\,\frac{\mathrm{d}t}{t} \qquad (u \in \overline{\operatorname{ran}(L)}).$$

One possible choice is $h(z) = c^{-1}g^*(z)$ with $c = \int_0^\infty |g(t)|^2 \frac{\mathrm{d}t}{t} > 0$. It is convenient to give such functions a name.

[3]This is the same 2ε-argument as in the proof of Proposition 2.4. We hope that you have convinced yourself in the meantime ;-)

Definition 6.17. *Let* $\varphi \in (0, \pi)$ *and* $g \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$. *A function* $h \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ *that satisfies* (6.11) *is called* Calderón sibling *of* g.

Theorem 6.16 can also be used to derive many classical representation formulæ, for example for fractional powers, with optimal domain of convergence. Here is an example, dating back to the work of Balakrishnan [Bal60] at the beginning of the 1960s.

Proposition 6.18 (Balakrishnan representation)**.** *Let* $0 < \operatorname{Re}\alpha < 1$. *Then we have for all* $u \in \operatorname{dom}(L^\alpha)$ *that*

$$L^\alpha u = \frac{\sin \alpha\pi}{\pi} \int_0^\infty t^{\alpha-1} L(t+L)^{-1} u \,\mathrm{d}t, \tag{6.12}$$

where the right-hand side is an improper integral in H.

Proof. Let us first give the proof when, in addition, L is injective. Let $u \in \operatorname{dom}(L^\alpha)$. A classical formula[4] says that

$$\int_0^\infty \frac{\tau^{1-\alpha}}{1+\tau} \frac{\mathrm{d}\tau}{\tau} = \frac{\pi}{\sin \alpha\pi}.$$

Thus, $f := \frac{\sin \alpha\pi}{\pi} \mathbf{z}^{1-\alpha}(1+\mathbf{z})^{-1}$ satisfies the assumption of Theorem 6.16 and we obtain as an improper integral,

$$\begin{aligned}
L^\alpha u &= \int_0^\infty f(\tau L) L^\alpha u \,\frac{\mathrm{d}\tau}{\tau} \\
&\overset{5.15}{=} \frac{\sin \alpha\pi}{\pi} \int_0^\infty (\tau L)^{1-\alpha} L^\alpha (1+\tau L)^{-1} u \,\frac{\mathrm{d}\tau}{\tau} \\
&\overset{6.3}{=} \frac{\sin \alpha\pi}{\pi} \int_0^\infty \tau^{1-\alpha} L(1+\tau L)^{-1} u \,\frac{\mathrm{d}\tau}{\tau} \\
&\overset{\tau = t^{-1}}{=} \frac{\sin \alpha\pi}{\pi} \int_0^\infty t^{\alpha-1} L(t+L)^{-1} u \,\mathrm{d}t.
\end{aligned}$$

If L is not injective, then we have to be more careful when applying the Calderón reproducing formula. Namely, we need to check beforehand that $L^\alpha u \in \overline{\operatorname{ran}(L)}$. To this end, we recall from Proposition 2.4 (b) and its proof that we have a topological decomposition $H = \ker(L) \oplus \overline{\operatorname{ran}(L)}$ in which the projection $P : H \to \ker(L)$ is given by $Pv = \lim_{j\to\infty}(1+jL)^{-1}v$. Now, we have

$$P(L^\alpha u) \overset{5.15}{=} \lim_{j\to\infty} j^{-\alpha}\big[(jL)^\alpha (1+jL)^{-1} u\big],$$

where the sequence in square brackets is uniformly bounded by Lemma 6.14. Thus, we get $P(L^\alpha u) = 0$ and, consequently, $L^\alpha u \in \overline{\operatorname{ran}(L)}$. □

[4] You may know this one from your complex analysis course. A direct proof is to substitute $\tau = \frac{1}{s} - 1$ and recognize the Beta/Gamma functions $B(\alpha, 1-\alpha) = \Gamma(\alpha)\Gamma(1-\alpha) = \frac{\pi}{\sin \pi\alpha}$. In any case, the precise constant will not matter in this course and you may simply write $C(\alpha)$.

6.6. Exercises

Exercise 6.1. Suppose that L is bounded. Prove that for any $\alpha \in \mathbb{C}$ with $\operatorname{Re}\alpha > 0$ also L^α is bounded.

Exercise 6.2 (Moment inequality)**.** Let $\alpha, \beta, \gamma \in \mathbb{C}$ with $0 < \operatorname{Re}\alpha < \operatorname{Re}\beta < \operatorname{Re}\gamma$. Prove that there is a constant C such that the *moment inequality* holds:

$$\|L^\beta u\| \le C\|L^\alpha u\|^{1-\theta}\|L^\gamma u\|^\theta \quad (u \in \operatorname{dom}(L^\gamma)),$$

where the interpolation parameter θ is determined by $\operatorname{Re}\beta = (1-\theta)\operatorname{Re}\alpha + \theta\operatorname{Re}\gamma$.

Hint: Start out with a representation $L^\beta u = \int_0^\infty f(\tau L)L^\beta u \,\frac{\mathrm{d}\tau}{\tau}$, where during the further course of the proof you will see how much decay of f you need. Split the integral at height $\tau = R$ to be chosen wisely and estimate both parts differently.

Exercise 6.3. In this exercise you are going to prove Proposition 6.6. You may proceed as follows.

(a) Indicate, where the semigroup property in Proposition 6.6 (a) comes from.

(b) Prove parts (b) and (c) of Proposition 6.6 for $u_0 \in \ker(L)$.

(c) Argue that in order to complete the proofs of Proposition 6.6 (b) and Proposition 6.6 (c), it suffices to treat the case $u_0 \in \operatorname{dom}(L) \cap \operatorname{ran}(L)$.

(d) Complete the proofs of parts (b) and (c) of Proposition 6.6.

(e) Prove Proposition 6.6 (d) by differentiation under the integral sign.

Exercise 6.4 (An extension of Proposition 5.18)**.** We suppose again $\varphi_L < \pi/2$ and let $\varphi \in (\varphi_L, \pi/2)$.

(a) Construct a holomorphic extension of $t \mapsto \mathrm{e}^{-tL}$ to a suitable sector S_θ. Which relation between θ and φ_L is required?

(b) Prove that e^{-tL} is injective for every $t > 0$.

(c) Conclude that $\mathrm{M}_L(\mathrm{S}_\varphi)$ contains more functions than those in Proposition 5.18.

(d) Could something similar be done for sectorial operators of angle $\varphi_L \ge \pi/2$?

Exercise 6.5 (Scaling properties of fractional powers)**.** Let $\varphi \in (\varphi_L, \pi)$ and $\alpha > 0$ be such that $\alpha\varphi < \pi$. We are going to show that L^α is sectorial of angle $\alpha\varphi$ by proceeding as follows.

(a) Let $\lambda \in \mathbb{C} \setminus \overline{\mathrm{S}_{\alpha\varphi}}$. Argue that

$$f_{\alpha,\lambda} := \frac{\lambda}{\lambda - \mathbf{z}^\alpha} - \frac{|\lambda|^{1/\alpha}}{|\lambda|^{1/\alpha} + \mathbf{z}}$$

is a function in $\mathrm{H}_0^\infty(\mathrm{S}_\varphi)$.

(b) Prove $\lambda \in \varrho(L^\alpha)$ and give a formula for $R(\lambda, L^\alpha)$ via the functional calculus for L.

(c) Conclude.

7. H^∞-calculus

In many situations it can be decisive to know whether, for some given sectorial operator L and some holomorphic function f, the operator $f(L)$ is bounded. For the negative Laplacian on $\mathbb{R}^n$, Theorem 4.14 provides us with a complete picture: $f(L)$ is bounded precisely when f is bounded and the latter's supremum norm controls the former's operator norm.

In this section we investigate such type of bounds for more general sectorial operators L and ask for given $\varphi \in (\varphi_L, \pi)$:

Is $f(L) \in \mathcal{L}(H)$ for all $f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)$?

In order to define $f(L)$ in the first place, we are led to assuming that L is injective, because only then we can use the universal regularizer $e = \mathbf{z}(1+\mathbf{z})^{-2}$ for f, compare with Remark 5.11. This assumption, however, is no particular restriction as long as we are not limited to one specific Hilbert space. More precisely, you will learn in Exercise 7.5 that every sectorial operator becomes injective and sectorial on the closed subspace $\overline{\operatorname{ran}(L)} \subseteq H$. Sadly — or should we say fortunately? — the answer to the question above will be 'no' in general and such operators get their own quality label. But we shall see that elliptic operators in divergence form, and, more generally, m-accretive operators, do have this beautiful property.

Notation 7.1. In the whole lecture, H is a Hilbert space with inner product $\langle \cdot , \cdot \rangle$ and norm $\| \cdot \|$. We denote by $e := \mathbf{z}(1+\mathbf{z})^{-2}$ the universal regularizer for bounded, holomorphic functions.

7.1. The notion of a bounded H^∞-calculus

To distinguish operators for which the answer to the question above is 'yes', we introduce the following notion.

Definition 7.2. *Let L be an injective sectorial operator in H. We say that L has a* bounded H^∞-calculus *of angle $\varphi \in (\varphi_L, \pi)$ if there is a constant C such that for all $f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)$ we have $f(L) \in \mathcal{L}(H)$ with norm bound*

$$\|f(L)\|_{\mathcal{L}(H)} \le C\|f\|_{\infty,\varphi}. \tag{7.1}$$

The infimum over all such φ is denoted by φ_L^∞ and called H^∞-angle *of L.*

Remark 7.3. (a) If L has a bounded H^∞-calculus of angle φ, then

$$\mathrm{H}^\infty(\mathrm{S}_\varphi) \to \mathcal{L}(H), \quad f \mapsto f(L)$$

is a bounded algebra homomorphism by Theorem 5.15.

(b) The *bound* C is useful, because it links the 'size' of the operator $f(L)$ with the 'size' of the function f. However, the sole property that $f(L) \in \mathcal{L}(H)$ for every $f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)$ furnishes some C by abstract nonsense, see Exercise 7.4.

(c) The notion of 'having a bounded H^∞-calculus' dates back to Alan McIntosh's seminal treatise [McI86], even though we actually require that 'the $\mathrm{H}^\infty(\mathrm{S}_\varphi)$-calculus' that we have constructed in Lecture 5 is bounded. For more on these uniqueness issues we refer interested readers to [Haa06, Sect. 5.3].

Usually, it is much harder to check $f(L) \in \mathcal{L}(H)$ for a genuine bounded holomorphic function than to prove the bound (7.1) for nice f. This is why McIntosh's convergence lemma below is so useful.

Lemma 7.4 (Convergence lemma)**.** *Let L be an injective sectorial operator in H. Let $\varphi \in (\varphi_L, \pi)$ and $(f_j) \subseteq \mathrm{H}^\infty(\mathrm{S}_\varphi)$ be a bounded sequence that converges uniformly on compact subsets of S_φ. Then the limit f is again in $\mathrm{H}^\infty(\mathrm{S}_\varphi)$ and*

$$f_j(L)u \to f(L)u \quad as \quad j \to \infty \quad for\ all \quad u \in \mathrm{dom}(L) \cap \mathrm{ran}(L).$$

Moreover, we have:

(a) If $f_j(L) \in \mathcal{L}(H)$ for all j and if there exists $T \in \mathcal{L}(H)$ such that $f_j(L) \to T$ strongly as $j \to \infty$, then $f(L) = T$.

(b) If $\sup_j \|f_j(L)\|_{\mathcal{L}(H)} < \infty$, then $f(L) \in \mathcal{L}(H)$ and $f_j(L) \to f(L)$ strongly as $j \to \infty$.

Before we come to the proof, we state explicitly the application we have in mind.

Corollary 7.5. *If* (7.1) *holds for all $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$, then it also holds for all $f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)$ and hence, L has a bounded H^∞-calculus of angle φ.*

Proof. Given $f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)$, we construct an approximation of f by

$$f_j := e^{1/j} f. \tag{7.2}$$

In order to see that f_j is holomorphic, it suffices to note that e maps $\mathrm{S}_\varphi \subseteq \mathbb{C} \setminus (-\infty, 0]$ into itself.[1] Moreover, we have $f_j \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ with $\|f_j\|_{\infty,\varphi} \leq \|f\|_{\infty,\varphi} \|e\|_{\infty,\varphi}^{1/j}$ and

[1]Without loss of generality, consider the case $0 < \arg(z) < \varphi$. Then $0 < \arg(1+z) < \arg(z)$, which implies $\arg(e(z)) = \arg(z) - 2\arg(1+z) \in (-\arg(z), \arg(z))$, and therefore $e(z) \in \mathrm{S}_\varphi$.

$f_j \to f$ in the limit as $j \to \infty$ uniformly on compact subsets of S_φ. Our assumption lets us apply Lemma 7.4 (b) to this sequence. Thus, we get for every $u \in H$ that

$$\|f(L)u\| = \lim_{j\to\infty} \|f_j(L)u\| \overset{(7.1)}{\leq} C \liminf_{j\to\infty} \|f\|_{\infty,\varphi} \|e\|_{\infty,\varphi}^{1/j} \|u\| = C\|f\|_{\infty,\varphi}\|u\|. \qquad \square$$

Proof of Lemma 7.4. The uniform convergence on compact subsets of S_φ together with the uniform boundedness of (f_j) implies that $f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)$. First, we let $u \in \mathrm{dom}(L) \cap \mathrm{ran}(L)$ and we recall from Proposition 2.4 (e) that this means $u = e(L)v$ for some $v \in H$. We also recall from Lemma 6.4 that u belongs to $\mathrm{dom}(f(L))$ and $\mathrm{dom}(f_j(L))$ for every j. By construction, we have

$$\begin{aligned} f_j(L)u &= e(L)^{-1}(ef_j)(L)u = (ef_j)(L)v, \\ f(L)u &= e(L)^{-1}(ef)(L)u \ = (ef)(L)v, \end{aligned}$$

and it remains to pass to the limit on the right-hand sides, using the pointwise bound $|ef_j| \leq (\sup_j \|f_j\|_{\infty,\varphi})|e|$ on S_φ and the dominated convergence theorem:

$$\begin{aligned} \lim_{j\to\infty} (ef_j)(L)v &= \lim_{j\to\infty} \frac{1}{2\pi \mathrm{i}} \int_{\gamma_\psi} e(z) f_j(z)(z-L)^{-1} v \, \mathrm{d}z \\ &= \frac{1}{2\pi\mathrm{i}} \int_{\gamma_\psi} e(z) f(z)(z-L)^{-1} v \, \mathrm{d}z = (ef)(L)v. \end{aligned}$$

This proves the general claim on strong convergence on $\mathrm{dom}(L) \cap \mathrm{ran}(L)$.

Under the additional assumption in (a), we know that

$$\big\{(u, Tu) \mid u \in \mathrm{dom}(L) \cap \mathrm{ran}(L)\big\} \subseteq f(L).$$

Here, $\mathrm{dom}(L) \cap \mathrm{ran}(L)$ is dense in H because L is injective, see Proposition 2.4 (e) and (b). Thus, the closure of the left-hand side is T. As $f(L)$ is closed, we first obtain $T \subseteq f(L)$ and then $T = f(L)$ since $\mathrm{dom}(T) = H$.

Finally, in the situation of (b) the sequence of bounded operators $(f_j(L))$ is bounded and converges strongly on a dense subset of H. Hence, it converges strongly everywhere on H to some operator $T \in \mathcal{L}(H)$ and according to (a) we must have $T = f(L)$. $\square$

7.2. Examples

Let us consider three examples of operators in the light of Definition 7.2.

Example 7.6 (The negative Laplacian on $\mathbb{R}^n$)**.** This operator has a bounded H^∞-calculus of any angle $\varphi \in (0, \pi)$ and we have the explicit norm bound

$$\|f(-\Delta)\|_{\mathcal{L}(\mathrm{L}^2(\mathbb{R}^n))} \leq \|f\|_{\infty,\varphi} \quad (f \in \mathrm{H}^\infty(\mathrm{S}_\varphi))$$

that has already appeared in Theorem 4.14. (We checked later on in Example 5.14 that the multiplier calculus from Lecture 4 is consistent with our general theory.)

Example 7.7 (Multiplication operators)**.** Let M_m be our usual multiplication operator in $\mathrm{L}^2(\mathbb{R}^n)$. In Exercise 7.3 you will be asked to provide all details for the following discussion. The conditions $m \neq 0$ almost everywhere and $\operatorname{essran}(m) \subseteq \overline{\mathrm{S}_\nu}$ are necessary and sufficient for M_m being injective and sectorial of angle $\nu \in (0, \pi)$. In this case, if $\varphi \in (\nu, \pi)$ and $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$, then similarly to Example 5.8 we obtain $f(M_m) = M_{f \circ m}$ and thus

$$\|f(M_m)\|_{\mathcal{L}(\mathrm{L}^2(\mathbb{R}^n))} \leq \|f \circ m\|_{\mathrm{L}^\infty(\mathbb{R}^n)} = \|f\|_{\infty,\varphi} \quad (f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)).$$

It follows from Corollary 7.5 that M_m has a bounded H^∞-calculus of angle φ and in particular, we have $\varphi_{M_m}^\infty = \varphi_{M_m}$.

Example 7.8 (Bounded operators)**.** A bounded and invertible sectorial operator L in H has a bounded H^∞-calculus and $\varphi_L^\infty = \varphi_L$. To this end, let $\varphi \in (\varphi_L, \pi)$ and $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$. Using the decay of f, we find

$$f(L) = \lim_{\substack{\varepsilon \to 0 \\ R \to \infty}} \frac{1}{2\pi \mathrm{i}} \int_{\gamma_{\psi,\varepsilon,R}} f(z)(z - L)^{-1} \, \mathrm{d}z, \tag{7.3}$$

where the paths $\gamma_{\psi,\varepsilon,R}$ with $0 < \varepsilon < 1 < R$ are as in the proof of Lemma 5.4. Now, keep an eye on Figure 7.1. Since L is bounded, the spectrum of L is compact. Since L is invertible, it does not contain 0. Hence, we can fix ε, R such that $\sigma(L)$ lies in the interior of $\gamma_{\psi,\varepsilon,R}$. Since $z \mapsto f(z)(z - L)^{-1}$ is holomorphic in $\mathrm{S}_\varphi \setminus \sigma(L)$, the Cauchy integral formula tells us that

$$\frac{1}{2\pi \mathrm{i}} \int_{\gamma_{\psi,\varepsilon,R}} f(z)(z - L)^{-1} \, \mathrm{d}z = \frac{1}{2\pi \mathrm{i}} \int_{\gamma_{\psi,\varepsilon',R'}} f(z)(z - L)^{-1} \, \mathrm{d}z,$$

whenever $0 < \varepsilon' < \varepsilon$ and $R' > R$. Thus, we can omit the limit in (7.3) and write

$$f(L) = \frac{1}{2\pi \mathrm{i}} \int_{\gamma_{\psi,\varepsilon,R}} f(z)(z - L)^{-1} \, \mathrm{d}z.$$

Since the length $\ell(\gamma_{\psi,\varepsilon,R})$ of $\gamma_{\psi,\varepsilon,R}$ is finite, we do no longer need the decay of f to control the integral on the right. This was the key point! Now, we obtain

$$\|f(L)\|_{\mathcal{L}(H)} \leq \left(\frac{\ell(\gamma_{\psi,\varepsilon,R})}{2\pi} \sup_{z \in \gamma_{\psi,\varepsilon,R}} \|(z - L)^{-1}\|_{\mathcal{L}(H)} \right) \|f\|_{\infty,\varphi} =: C \|f\|_{\infty,\varphi},$$

where C depends on L and φ but not on f. This calculation was for $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ and the claim follows again from Corollary 7.5.

All classes considered above are stable under taking adjoints. Still, it is worth pointing out the following general duality principle.

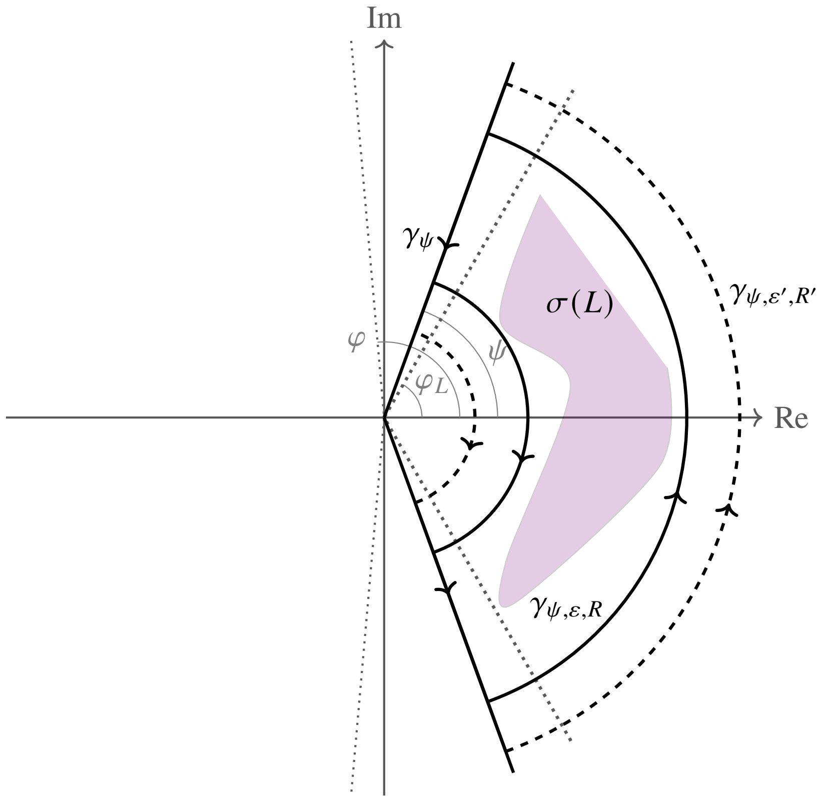

Figure 7.1.: The configuration of paths $\gamma_{\psi,\varepsilon,R}$ and $\gamma_{\psi,\varepsilon',R'}$ as in Example 7.8.

Lemma 7.9. *Let L be an injective sectorial operator in H and let $\varphi \in (\varphi_L, \pi)$. Then L has a bounded H^∞-calculus of angle φ if and only if L^* does. In particular, $\varphi_L^\infty = \varphi_{L^*}^\infty$.*

Proof. This is a straightforward application of the adjoint calculus. Indeed, if L has a bounded H^∞-calculus of angle $\varphi \in (\varphi_L, \pi)$ with bound C, then

$$\|f(L^*)\|_{\mathcal{L}(H)} \overset{6.10}{=} \|f^*(L)^*\|_{\mathcal{L}(H)} \overset{1.20\,(h)}{=} \|f^*(L)\|_{\mathcal{L}(H)} \leq C\|f^*\|_{\infty,\varphi} = C\|f\|_{\infty,\varphi}.$$

Applying the same reasoning to L^* with adjoint $(L^*)^* = L$, yields the claim. □

We have seen examples of sectorial operators with a bounded H^∞-calculus and experience tells us since many years that most sectorial operators that we encounter, in particular those arising from partial differential equations, have this property. However, it does not come for free:

Warning 7.10. There is a Hilbert space H and an invertible sectorial operator L in H that does not have a bounded H^∞-calculus of any angle. One of the first such examples due to McIntosh and Yagi [MY90] is an intricate tensor product of matrices on larger and larger finite dimensional spaces. A general method for constructing counterexamples by means of conditional Schauder bases is presented in [Haa06, Ch. 9]. All known counterexamples look somewhat pathological at first sight and if you know a simpler one, please let us know!

The examples above also had in common that in the end

$$\varphi_L^\infty = \varphi_L$$

is the best possible angle. Here, there is no caveat but precise results and proofs have to wait until the next lecture.

7.3. The H$^\infty$-calculus for m-accretive operators

We come to a significantly more involved example of sectorial operators that have a bounded H$^\infty$-calculus: m-accretive operators. Their last appearance on the show dates back to Episode 2 and we recall from Definition 2.22 that L is m-accretive if it satisfies the particular resolvent estimate

$$\|(\lambda + L)^{-1}\|_{\mathcal{L}(H)} \leq \frac{1}{\mathrm{Re}(\lambda)} \quad (\mathrm{Re}(\lambda) > 0), \tag{7.4}$$

which implies sectoriality of angle $\pi/2$. We have seen in Theorem 2.21 that the operators associated with bounded, elliptic and accretive sesquilinear forms are of this type. The following lemma provides an equivalent characterization of m-accretivity that is typically easier to verify for concrete examples. We leave the proof for you as Exercise 7.1.

Lemma 7.11. *For a linear operator L in H the following are equivalent:*

(a) *L is m-accretive.*

(b) $\mathrm{Re}\,\langle Lu, u\rangle \geq 0$ *for all* $u \in \mathrm{dom}(L)$ *and there exists some* $\lambda > 0$ *such that* $\mathrm{ran}(\lambda + L) = H$.

Moreover, the range condition *in (b) can be dropped if L is bounded.*

Here is the main result in this second part of the lecture:

Theorem 7.12 (Von Neumann's inequality)**.** *Let L be an m-accretive operator in H and let $\varphi \in (\pi/2, \pi)$. Then*

$$\|f(L)\|_{\mathcal{L}(H)} \leq \|f\|_{\infty,\pi/2} \qquad (f \in \mathcal{E}(\mathrm{S}_\varphi)). \tag{7.5}$$

If L is injective, then (7.5) *continues to hold for all $f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)$ and L has a bounded* H$^\infty$*-calculus of angle φ.*

Let us discuss the theorem in more detail before delving into the proof. Ours is not the original formulation due to von Neumann [vN51], which dates back long before

H^∞-calculus, but a very close relative of it [Haa06, Rem. 7.1.9]. The statement can be turned into an equivalence: Namely, if L is merely sectorial of angle $\pi/2$, then it is m-accretive if and only if (7.5) holds. For the converse direction, it suffices to note that (7.5) for $f_\lambda := (\lambda + \mathbf{z})^{-1}$ with $\mathrm{Re}(\lambda) > 0$ yields (7.4).

For the proof of von Neumann's inequality we will use an approximation of L by bounded m-accretive operators with additional spectral properties. This is the content of the following lemma.

Lemma 7.13. *Let L be an m-accretive operator in H. For $\varepsilon \in (0,1)$ define*

$$L_\varepsilon := (\varepsilon + L)(1 + \varepsilon L)^{-1}.$$

The following properties hold:

(a) L_ε is bounded and $L_\varepsilon - \varepsilon$ is m-accretive. In particular, L_ε is m-accretive.

(b) $\varphi_{L_\varepsilon} < \pi/2$.

(c) For all $z \in \mathbb{C}$ with $\mathrm{Re}(z) < 0$ and all $u \in H$ we have

$$(z - L_\varepsilon)^{-1}u \to (z - L)^{-1}u \quad \text{as} \quad \varepsilon \searrow 0.$$

(d) Let $\varphi \in (\pi/2, \pi)$. For all $f \in \mathcal{E}(\mathrm{S}_\varphi)$ and all $u \in H$ we have

$$f(L_\varepsilon)u \to f(L)u \quad \text{as} \quad \varepsilon \searrow 0.$$

Proof. (a) Boundedness follows on writing

$$\begin{aligned} L_\varepsilon &= \big(\varepsilon^{-1} + L - \varepsilon^{-1} + \varepsilon\big)\big(\varepsilon^{-1}(\varepsilon^{-1} + L)^{-1}\big) \\ &= \varepsilon^{-1} - (\varepsilon^{-2} - 1)(\varepsilon^{-1} + L)^{-1}. \end{aligned}$$

Using the Cauchy–Schwarz inequality and m-accretivity of L, we obtain for every $u \in H$ that

$$\begin{aligned} \mathrm{Re}\langle L_\varepsilon u, u\rangle &\geq \varepsilon^{-1}\|u\|^2 - (\varepsilon^{-2} - 1)\|(\varepsilon^{-1} + L)^{-1}u\|\|u\| \\ &\geq \varepsilon^{-1}\|u\|^2 - (\varepsilon^{-2} - 1)\varepsilon\|u\|^2 = \varepsilon\|u\|^2. \end{aligned}$$

Lemma 7.11 yields that $L_\varepsilon - \varepsilon$ (and L_ε) are m-accretive.

(b) Let us collect some spectral information on L_ε in Figure 7.2. According to (a), L_ε is bounded and $L_\varepsilon - \varepsilon$ is m-accretive. Hence, the spectrum of L_ε is a compact subset of the right half-plane $\{z \in \mathbb{C} \mid \mathrm{Re}(z) \geq \varepsilon\}$ and therefore contained in $\overline{\mathrm{S}_\nu}$ for some angle $\nu \in (0, \pi/2)$ depending on ε.

In order to show the resolvent estimate required for sectoriality of angle ν, we fix $\psi \in (\nu, \pi)$ and set $R_\varepsilon := \|L_\varepsilon\|_{\mathcal{L}(H)}$. On the compact set $\{z \in \mathbb{C} \mid z \notin \mathrm{S}_\psi$ and $|z| \leq$

$2R_\varepsilon\}$ the resolvent $R(\,\cdot\,, L_\varepsilon)$ is even uniformly bounded by holomorphy, whereas for $|z| > 2R_\varepsilon$ we can use a Neumann series to bound

$$\begin{aligned}\big\|R(z, L_\varepsilon)\big\|_{\mathcal{L}(H)} &= \big\|z^{-1}(1 - z^{-1}L_\varepsilon)^{-1}\big\|_{\mathcal{L}(H)} \\ &= \Big\|z^{-1}\sum_{k=0}^{\infty}(z^{-1}L_\varepsilon)^k\Big\|_{\mathcal{L}(H)} \le 2|z|^{-1}.\end{aligned}$$

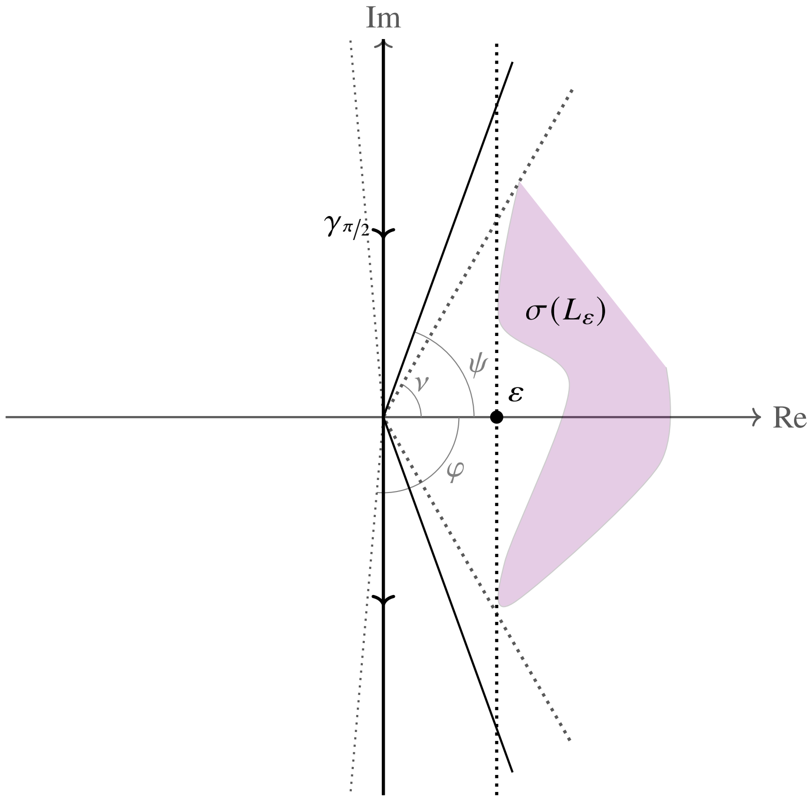

Figure 7.2.: The spectrum of the bounded operator L_ε is a compact subset of the half-plane $\{z \in \mathbb{C} \mid \mathrm{Re}(z) \ge \varepsilon\}$, hence contained in the closure of some sector S_ν with $\nu < \pi/2$. The resolvent bound $\|R(z, L_\varepsilon)\|_{\mathcal{L}(H)} \le 2|z|^{-1}$ for $|z| > 2\|L_\varepsilon\|_{\mathcal{L}(H)}$, coming from a Neumann series, yields sectoriality of L_ε with angle ν. In Theorem 7.12 the function f is holomorphic on S_φ.

(c) We set $v := (z - L)^{-1}u \in \mathrm{dom}(L)$ and write the difference of resolvents as

$$\begin{aligned}(z - L)^{-1}u - (z - L_\varepsilon)^{-1}u &= (z - L_\varepsilon)^{-1}(L - L_\varepsilon)(z - L)^{-1}u \\ &= (z - L_\varepsilon)^{-1}\big(L - (\varepsilon + L)(1 + \varepsilon L)^{-1}\big)v \\ &= (z - L_\varepsilon)^{-1}\big(Lv - (1 + \varepsilon L)^{-1}Lv - \varepsilon(1 + \varepsilon L)^{-1}v\big).\end{aligned}$$

Since L_ε is m-accretive, we can bound its resolvent independently of ε to obtain

$$\begin{aligned}&\left\|(z-L)^{-1}u-(z-L_\varepsilon)^{-1}u\right\|\\ &\leq \frac{1}{|\operatorname{Re}(z)|}\Big(\left\|Lv-(1+\varepsilon L)^{-1}Lv\right\|+\varepsilon\left\|(1+\varepsilon L)^{-1}v\right\|\Big).\end{aligned}$$

In the limit as $\varepsilon \to 0$, both norms on the right tend to zero in virtue of Proposition 2.4 (d).

(d) There is nothing to do for $f = \mathbf{1}$ and in (c) we have already dealt with $f = (1+\mathbf{z})^{-1}$. Let now $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$. We fix $\psi \in ({}^{\pi}/{}_2, \varphi)$ in order to write

$$f(L)u - f(L_\varepsilon)u = \frac{1}{2\pi\mathrm{i}}\int_{\gamma_\psi} f(z)\big((z-L)^{-1}u-(z-L_\varepsilon)^{-1}u\big)\,\mathrm{d}z.$$

Since both L and L_ε are m-accretive, we obtain an integrable majorant for the integrands on the right by estimating

$$\left\|f(z)\big((z-L)^{-1}u-(z-L_\varepsilon)^{-1}u\big)\right\| \leq \frac{2|f(z)|}{|z|\sin(\psi-{}^{\pi}/{}_2)}\|u\| \quad (z\in\gamma_\psi).$$

The claim follows by dominated convergence and (c). □

We are ready to prove von Neumann's inequality.

Proof of Theorem 7.12. Let us first prove the theorem for L_ε in place of L.

We begin by deriving a particular formula for $f(L_\varepsilon)$ when $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$. Since L_ε is sectorial of angle smaller than ${}^{\pi}/{}_2$, we can calculate

$$f(L_\varepsilon) = \frac{1}{2\pi\mathrm{i}}\int_{\gamma_{\pi/2}} f(z)(z-L_\varepsilon)^{-1}\,\mathrm{d}z$$

via a Cauchy integral along the imaginary axis, see also Figure 7.2. The possibility to choose this path is crucial for the argument and in general it is not permitted for the operator L itself. Keeping Figure 7.2 and Proposition 1.21 in mind, we see that $(\mathbf{z}+L_\varepsilon^*)^{-1}$ is holomorphic on the right half-plane $\{z\in\mathbb{C}\colon \operatorname{Re}(z) > -\varepsilon\}$ and bounded by ${}^{C}/{}_{|\mathbf{z}|}$. Hence, we can use Cauchy's integral formula to compute

$$0 = \int_{\gamma_{\pi/2}} f(z)(z+L_\varepsilon^*)^{-1}\,\mathrm{d}z.$$

The two preceding identities yield

$$\begin{aligned}f(L_\varepsilon) &= \frac{1}{2\pi\mathrm{i}}\int_{\gamma_{\pi/2}} f(z)\big((z-L_\varepsilon)^{-1}-(z+L_\varepsilon^*)^{-1}\big)\,\mathrm{d}z\\ &= \frac{1}{2\pi\mathrm{i}}\int_{\gamma_{\pi/2}} f(z)(z+L_\varepsilon^*)^{-1}(L_\varepsilon+L_\varepsilon^*)(z-L_\varepsilon)^{-1}\,\mathrm{d}z.\end{aligned}$$

Now, we use the symmetry property $\overline{z} = -z$, which holds precisely when $z \in \gamma_{\pi/2}$, in order to obtain for any $u, v \in H$ that

$$\begin{aligned} \langle f(L_\varepsilon)u, v\rangle &= \frac{1}{2\pi \mathrm{i}} \int_{\gamma_{\pi/2}} f(z) \big\langle (L_\varepsilon + L_\varepsilon^*)(z - L_\varepsilon)^{-1}u, (\overline{z} + L_\varepsilon)^{-1}v \big\rangle \,\mathrm{d}z \\ &= -\frac{1}{2\pi \mathrm{i}} \int_{\gamma_{\pi/2}} f(z) \big\langle (L_\varepsilon + L_\varepsilon^*)(z - L_\varepsilon)^{-1}u, (z - L_\varepsilon)^{-1}v \big\rangle \,\mathrm{d}z. \end{aligned} \tag{7.6}$$

At this point, we introduce the bounded auxiliary operator

$$T \coloneqq L_\varepsilon + L_\varepsilon^*.$$

Clearly, T is self-adjoint and since L_ε is m-accretive, we obtain from Lemma 7.11 that

$$\operatorname{Re}\langle Tu, u\rangle = \operatorname{Re}\big(\langle L_\varepsilon u, u\rangle + \langle u, L_\varepsilon u\rangle\big) = 2\operatorname{Re}\langle L_\varepsilon u, u\rangle \geq 0 \qquad (u \in H)$$

and hence that T is m-accretive. The key point is that the square root $\sqrt{T}$ is self-adjoint (Corollary 6.11) and bounded (Exercise 6.1). Going back to (7.6), we obtain the representation

$$\begin{aligned} \langle f(L_\varepsilon)u, v\rangle &= -\frac{1}{2\pi \mathrm{i}} \int_{\gamma_{\pi/2}} f(z) \big\langle \sqrt{T}(z - L_\varepsilon)^{-1}u, \sqrt{T}(z - L_\varepsilon)^{-1}v \big\rangle \,\mathrm{d}z \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\mathrm{i}t) \big\langle \sqrt{T}(\mathrm{i}t - L_\varepsilon)^{-1}u, \sqrt{T}(\mathrm{i}t - L_\varepsilon)^{-1}v \big\rangle \,\mathrm{d}t, \end{aligned} \tag{7.7}$$

where in the last step we have kept in mind that $\gamma_{\pi/2}$ runs from $\mathrm{i} \cdot \infty$ to $-\mathrm{i} \cdot \infty$.

We claim that this formula is in fact valid for all $f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)$, because, similarly to Example 7.8, the decay of f is no longer needed to make sense of the integral on the right-hand side. Indeed, let $f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)$ and consider a bounded sequence $(f_j) \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ that converges to f uniformly on compact subsets of S_φ.[2] Then (7.7) yields

$$\langle f_j(L_\varepsilon)u, v\rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} f_j(\mathrm{i}t) \big\langle \sqrt{T}(\mathrm{i}t - L_\varepsilon)^{-1}u, \sqrt{T}(\mathrm{i}t - L_\varepsilon)^{-1}v \big\rangle \,\mathrm{d}t$$

and we want to pass to the limit as $j \to \infty$. On the left, the convergence lemma yields $f_j(L_\varepsilon)u \to f(L_\varepsilon)u$ since we have $\operatorname{dom}(L_\varepsilon) = H = \operatorname{ran}(L_\varepsilon)$. On the right, the integrand is continuous and for t large it is bounded by $Ct^{-2}\|\sqrt{T}\|_{\mathcal{L}(H)}^2 \|u\|\|v\|$ due to the resolvent bounds for L_ε, see again Figure 7.2, where C that does not depend on j and t. Thus, dominated convergence applies. Altogether, we have shown (7.7) for $f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)$. In particular, taking $f = \mathbf{1}$ and $u = v$, gives a formula for the norm:

$$\|u\|^2 = \frac{1}{2\pi} \int_{\mathbb{R}} \big\|\sqrt{T}(\mathrm{i}t - L_\varepsilon)^{-1}u\big\|^2 \,\mathrm{d}t \qquad (u \in H). \tag{7.8}$$

[2]One example is $f_j = (\mathbf{z}(1 + \mathbf{z})^{-2})^{1/j} f$, see the proof of Corollary 7.5.

Now, we let $f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)$, use the Cauchy–Schwarz inequality in (7.7) and then apply (7.8) in order to remove $\sqrt{T}$ entirely from the game:

$$\begin{aligned}\left|\langle f(L_\varepsilon)u, v\rangle\right| &\le \frac{\|f\|_{\infty,\pi/2}}{2\pi}\bigg(\int_{\mathbb{R}}\big\|\sqrt{T}(\mathrm{i}t - L_\varepsilon)^{-1}u\big\|^2\,\mathrm{d}t\bigg)^{1/2}\bigg(\int_{\mathbb{R}}\big\|\sqrt{T}(\mathrm{i}t - L_\varepsilon)^{-1}v\big\|^2\,\mathrm{d}t\bigg)^{1/2}\\ &= \|f\|_{\infty,\pi/2}\|u\|\|v\|.\end{aligned}$$

The choice $v = f(L_\varepsilon)u$ yields

$$\|f(L_\varepsilon)\|_{\mathcal{L}(H)} \le \|f\|_{\infty,\pi/2} \qquad (f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)), \tag{7.9}$$

which is von Neumann's inequality for L_ε.

We complete the proof by approximating L by L_ε. For $f \in \mathcal{E}(\mathrm{S}_\varphi)$ we know by virtue of Lemma 7.13 (d) and (7.9) that

$$\|f(L)u\| = \lim_{\varepsilon\to 0}\|f(L_\varepsilon)u\| \le \|f\|_{\infty,\pi/2}\|u\| \qquad (u \in H),$$

which establishes (7.5). If, in addition, L is injective, then (7.5) follows for all $f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)$ by Corollary 7.5 — or, more precisely, its proof, since we want to pass to the limit with the norm on $\mathrm{S}_{\pi/2}$ rather than S_φ. □

7.4. Exercises

Exercise 7.1 (Characterizations of m-accretivity)**.** Provide a proof of Lemma 7.11.

Hint: You will find inspiration from the proof of Theorem 2.21 . . .

Exercise 7.2 (m-accretive fractional powers)**.** Let L be an m-accretive operator in H and let $\alpha \in (0,1)$. Prove that L^α is m-accretive.

Exercise 7.3. Fill in the details left out in the discussion of the H^∞-calculus for multiplication operators in Example 7.7.

Hint: Compute $\langle f(M_m)u, v\rangle_{\mathrm{L}^2(\mathbb{R}^n)}$ for $u, v \in \mathrm{L}^2(\mathbb{R}^n)$ instead of trying to pull the variable x directly into the Bochner integral as in Example 5.8.

Exercise 7.4 (An automatic bound for the H^∞-calculus)**.** Let L be an injective sectorial operator in H. Let $\varphi \in (\varphi_L, \pi)$ and suppose that we have $f(L) \in \mathcal{L}(H)$ for every $f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)$. Prove that there is a constant $C \ge 0$ such that

$$\|f(L)\|_{\mathcal{L}(H)} \le C\|f\|_{\infty,\varphi}$$

holds for every such f.

Exercise 7.5 (The injective part)**.** In this exercise we describe a possibility to define an H^∞-calculus for a non-injective sectorial operator L in H, using the maximal restriction of L to an operator in the closed subspace $\overline{\operatorname{ran}(L)}$ of H:

$$L|_{\overline{\operatorname{ran}(L)}} := L \cap \big(\overline{\operatorname{ran}(L)} \times \overline{\operatorname{ran}(L)}\big).$$

(a) Let $\lambda \in \varrho(L)$. Prove that $(\lambda - L)^{-1}$ maps $\overline{\operatorname{ran}(L)}$ into itself.

(b) Conclude that $L|_{\overline{\operatorname{ran}(L)}}$ is a sectorial operator in $\overline{\operatorname{ran}(L)}$ and that $\varphi_{L|_{\overline{\operatorname{ran}(L)}}} = \varphi_L$.

(c) Prove that $L|_{\overline{\operatorname{ran}(L)}}$ is injective.

Let $\varphi \in (\varphi_L, \pi)$. Given $f \in \mathrm{H}^\infty(\mathrm{S}_\varphi)$, the properties above allow us to define $f(L)$ as a closed operator in $\overline{\operatorname{ran}(L)}$ via

$$f(L) := f(L|_{\overline{\operatorname{ran}(L)}}).$$

Of course, this definition is only meaningful, if it is compatible with $f(L)$ whenever $f \in \mathrm{M}_L(\mathrm{S}_\varphi)$. In order to see that this is indeed the case, we proceed as follows:

(d) Let $f \in \mathcal{E}(\mathrm{S}_\varphi)$. Prove that $f(L)u = f(L|_{\overline{\operatorname{ran}(L)}})u$ for all $u \in \overline{\operatorname{ran}(L)}$.

Hint: Exercise 5.5.

(e) Let $f \in \mathrm{M}_L(\mathrm{S}_\varphi)$. Conclude that $f(L|_{\overline{\operatorname{ran}(L)}})$ is the maximal restriction of $f(L)$ to an operator in $\overline{\operatorname{ran}(L)}$.

8. Quadratic estimates vs. functional calculus

In Corollary 4.17 we have used the Fourier transform to prove a fundamental inequality that compares the $\mathrm{L}^2(\mathbb{R}^n)$-norm with a quadratic integral involving functional calculus for the negative Laplacian:

$$\int_0^\infty \|f(-t\Delta)u\|_{\mathrm{L}^2(\mathbb{R}^n)}^2 \,\frac{\mathrm{d}t}{t} \simeq \|u\|_{\mathrm{L}^2(\mathbb{R}^n)}^2, \tag{8.1}$$

where f was a suitable non-zero function on $(0, \infty)$. This is the most basic example of a continuous Littlewood–Paley inequality. In our writing of (8.1) the specific properties of the Fourier transform seem to have disappeared and it makes sense to ask whether the same type of estimates hold for a general sectorial operator L in place of $-\Delta$. In return, if this works out, it should provide us with some sort of 'Fourier analysis adapted to the operator L' and we could try to generalize other special properties of the negative Laplacian from Lecture 4. In this lecture, we will explore these ideas further and we will learn that 'L-adapted Fourier analysis' is closely related to 'L having a bounded H^∞-calculus'.

Notation 8.1. As in the previous lecture, L denotes an injective sectorial operator in a Hilbert space H with inner product $\langle \cdot\,, \cdot \rangle$ and norm $\| \cdot \|$.

8.1. McIntosh's theorem

We begin by modelling a notion of quadratic estimates based on (8.1) that can hold for the given sectorial operator L and a given non-zero function f — or not.

Definition 8.2. *Let $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi) \setminus \{0\}$ for some $\varphi \in (\varphi_L, \pi)$. We say that L satisfies* quadratic estimates *(with auxiliary function f) if there is $C > 0$ such that*

$$C^{-1}\|u\|^2 \leq \int_0^\infty \|f(tL)u\|^2 \,\frac{\mathrm{d}t}{t} \leq C\|u\|^2 \qquad (u \in H).$$

It satisfies lower (upper) quadratic estimates *if only the first (the second) estimate holds.*

We find it instructive to showcase that quadratic estimates always hold when L is self-adjoint and where the specific argument breaks down if it is not. This also gives a non-Fourier related proof for the negative Laplacian (using a different class of auxiliary functions).

Lemma 8.3. *Let $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi) \setminus \{0\}$ for some $\varphi \in (\varphi_L, \pi)$ and suppose that L is self-adjoint. Set $c := \int_0^\infty |f(t)|^2 \frac{\mathrm{d}t}{t}$. Then L satisfies the quadratic 'estimates'*

$$\int_0^\infty \|f(tL)u\|^2 \frac{\mathrm{d}t}{t} = c\|u\|^2 \qquad (u \in H).$$

Proof. Let $u \in H$. The key observation is that we have $f(tL)^* = f^*(tL)$ for all $t > 0$, see Lemma 6.9. Using the monotone convergence theorem, we can write

$$\int_0^\infty \|f(tL)u\|^2 \frac{\mathrm{d}t}{t} = \lim_{R\to\infty} \int_{1/R}^R \langle f(tL)u, f(tL)u\rangle \frac{\mathrm{d}t}{t} = \left\langle \lim_{R\to\infty} \int_{1/R}^R (f^*f)(tL)u \frac{\mathrm{d}t}{t}, u \right\rangle.$$

In other words: Self-adjointness of L allows us to write the quadratic integral as a sesquilinear quantity amenable to 'cancellation properties' of f, because there are no absolute values anymore. Since $c = \int_0^\infty f^*(t)f(t) \frac{\mathrm{d}t}{t}$, the Calderón reproducing formula (Theorem 6.16) implies that the limit on the right-hand side equals cu and the claim follows. □

The principal result in this section, due to McIntosh [McI86], shows that quadratic estimates are independent of the auxiliary function and characterize the boundedness of the H^∞-calculus. In particular, and most certainly surprisingly, the H^∞-angle in Hilbert spaces is automatically the best possible one: We have $\varphi_L^\infty = \varphi_L$!

Theorem 8.4 (McIntosh)**.** *The following statements are equivalent.*

(a) L satisfies quadratic estimates for some *auxiliary function $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi) \setminus \{0\}$, where $\varphi \in (\varphi_L, \pi)$ is* some *angle.*

(b) L satisfies quadratic estimates for all *auxiliary functions $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi) \setminus \{0\}$, where $\varphi \in (\varphi_L, \pi)$ is* any *angle.*

(c) L has a bounded H^∞-calculus of some *angle $\varphi \in (\varphi_L, \pi)$.*

(d) L has a bounded H^∞-calculus of any angle $\varphi \in (\varphi_L, \pi)$.

We prepare the proof of Theorem 8.4 by a succession of four short lemmas.

Lemma 8.5. *Let $\varphi \in (\varphi_L, \pi)$. For any two $f, g \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ there exists a function $\zeta \in \mathrm{L}^1((0,\infty); \frac{\mathrm{d}t}{t})$ such that*

$$\|f(rL)g(sL)\|_{\mathcal{L}(H)} \leq \zeta(rs^{-1}) \qquad (r, s > 0).$$

Proof. Fix $\psi \in (\varphi_L, \varphi)$. By sectoriality of L and the decay of f, g, we can bound the Cauchy integral defining $f(rL)g(sL)$ by

$$\begin{aligned} \|f(rL)g(sL)\|_{\mathcal{L}(H)} &= \left\| \frac{1}{2\pi \mathrm{i}} \int_{\gamma_\psi} f(rz)g(sz)(z-L)^{-1} \,\mathrm{d}z \right\|_{\mathcal{L}(H)} \\ &\leq C \int_0^\infty \min\big((rt)^{s_f}, (rt)^{-s_f}\big) \min\big((st)^{s_g}, (st)^{-s_g}\big) \,\frac{\mathrm{d}t}{t} \end{aligned} \tag{8.2}$$

for some constants $C, s_f, s_g > 0$. We substitute $\tau := st$ and obtain

$$\begin{aligned} \|f(rL)g(sL)\|_{\mathcal{L}(H)} &\leq C \int_0^\infty \min\big((rs^{-1}\tau)^{s_f}, (rs^{-1}\tau)^{-s_f}\big) \min\big(\tau^{s_g}, \tau^{-s_g}\big) \,\frac{\mathrm{d}\tau}{\tau} \\ &=: \zeta(rs^{-1}) \end{aligned}$$

and, by Tonelli's theorem and the substitution $\rho := t\tau$, we obtain as required

$$\begin{aligned} \int_0^\infty \zeta(t) \,\frac{\mathrm{d}t}{t} &= C \int_0^\infty \min\big(\tau^{s_g}, \tau^{-s_g}\big) \int_0^\infty \min\big((t\tau)^{s_f}, (t\tau)^{-s_f}\big) \,\frac{\mathrm{d}t}{t} \,\frac{\mathrm{d}\tau}{\tau} \\ &= C \bigg(\int_0^\infty \min\big(\rho^{s_f}, \rho^{-s_f}\big) \,\frac{\mathrm{d}\rho}{\rho} \bigg) \bigg(\int_0^\infty \min\big(\tau^{s_g}, \tau^{-s_g}\big) \,\frac{\mathrm{d}\tau}{\tau} \bigg) < \infty. \qquad \square \end{aligned}$$

Remark 8.6. The lemma can be generalized to the following 'regularized H^∞-bound': For all $h \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ we have

$$\|f(rL)h(L)g(sL)\|_{\mathcal{L}(H)} \leq \zeta(rs^{-1})\|h\|_{\infty,\varphi} \qquad (r, s > 0).$$

Indeed, it suffices to estimate h by its supremum norm in the second line of (8.2). This version will be important when proving '$(b) \Longrightarrow (d)$' in Theorem 8.4.

The second lemma shows that different (non identically zero) auxiliary functions f yield equivalent quadratic norms.

Lemma 8.7. *Let $\varphi \in (\varphi_L, \pi)$. Given $f, g \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ with g not identically zero, there is a constant C such that*

$$\int_0^\infty \|f(sL)u\|^2 \,\frac{\mathrm{d}s}{s} \leq C \int_0^\infty \|g(tL)u\|^2 \,\frac{\mathrm{d}t}{t} \qquad (u \in H).$$

Proof. We begin by fixing a Calderón sibling $g^\sharp \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ of g as in Definition 6.17. Let $s > 0$. From the Calderón reproducing formula and since $f(sL)$ is bounded, we obtain

$$f(sL)u = \int_0^\infty f(sL)g^\sharp(tL)g(tL)u \,\frac{\mathrm{d}t}{t}$$

as an improper integral. Thus, we have

$$\int_0^\infty \|f(sL)u\|^2 \frac{\mathrm{d}s}{s} \leq \int_0^\infty \left(\int_0^\infty \big\|f(sL)g^\sharp(tL)g(tL)u\big\| \frac{\mathrm{d}t}{t}\right)^2 \frac{\mathrm{d}s}{s}.$$

Now, we use Lemma 8.5 with the functions $f, g^\sharp$ and the Cauchy–Schwarz inequality to perform what is often called a 'Schur type bound' in reminiscence of the Schur test for integral operators (Exercise 8.1):

$$\begin{aligned}
\int_0^\infty \|f(sL)u\|^2 \frac{\mathrm{d}s}{s} &\leq \int_0^\infty \left(\int_0^\infty \zeta(st^{-1}) \|g(tL)u\| \frac{\mathrm{d}t}{t}\right)^2 \frac{\mathrm{d}s}{s} \\
&\leq \int_0^\infty \left(\int_0^\infty \zeta(st^{-1}) \frac{\mathrm{d}t}{t}\right)\left(\int_0^\infty \zeta(st^{-1}) \|g(tL)u\|^2 \frac{\mathrm{d}t}{t}\right) \frac{\mathrm{d}s}{s} \\
&\overset{t = s\tau^{-1}}{=} \left(\int_0^\infty \zeta(\tau) \frac{\mathrm{d}\tau}{\tau}\right)\left(\int_0^\infty \int_0^\infty \zeta(st^{-1}) \|g(tL)u\|^2 \frac{\mathrm{d}s}{s}\frac{\mathrm{d}t}{t}\right) \\
&\overset{s = t\tau}{=} \left(\int_0^\infty \zeta(\tau) \frac{\mathrm{d}\tau}{\tau}\right)^2\left(\int_0^\infty \|g(tL)u\|^2 \frac{\mathrm{d}t}{t}\right). \qquad \square
\end{aligned}$$

The third ingredient is a duality principle that links upper quadratic estimates with lower quadratic estimates for the adjoint operator. Recall from Lemma 6.7 that the adjoint of an injective sectorial operator is of the same type and that $\varphi_L = \varphi_{L^*}$.

Lemma 8.8. *Let $\varphi \in (\varphi_L, \pi)$. If L satisfies upper quadratic estimates for all $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi) \setminus \{0\}$, then L^* satisfies lower quadratic estimates for all such f.*

Proof. We fix $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi) \setminus \{0\}$ and let $f^\sharp$ be a Calderón sibling, allowing us to write any given $u \in H$ via the reproducing formula

$$u = \int_0^\infty f^\sharp(tL^*)f(tL^*)u \frac{\mathrm{d}t}{t}.$$

Testing this equation with u yields

$$\|u\|^2 = \int_0^\infty \big\langle f^\sharp(tL^*)f(tL^*)u, u\big\rangle \frac{\mathrm{d}t}{t},$$

where we control the right-hand side using the Cauchy–Schwarz inequality and upper quadratic estimates for L with auxiliary function $(f^\sharp)^*$: Indeed, taking the duality relation in Lemma 6.9 into account, we get

$$\begin{aligned}
&= \int_0^\infty \big\langle f(tL^*)u, (f^\sharp)^*(tL)u\big\rangle \frac{\mathrm{d}t}{t} \\
&\leq \left(\int_0^\infty \big\|f(tL^*)u\big\|^2 \frac{\mathrm{d}t}{t}\right)^{1/2}\left(\int_0^\infty \big\|(f^\sharp)^*(tL)u\big\|^2 \frac{\mathrm{d}t}{t}\right)^{1/2}
\end{aligned}$$

$$\lesssim \left(\int_0^\infty \|f(tL^*)u\|^2 \, \frac{\mathrm{d}t}{t} \right)^{1/2} \|u\|.$$

It remains to divide both sides by $\|u\|$ and then take squares. □

While these three lemmas will help us proving boundedness of the H^∞-calculus from quadratic estimates, the converse will rely on the so-called 'unconditionality lemma'. The name stems from the orthogonality estimates that are typically used in bringing this lemma into play. (Stay patient for a moment, please.)

Lemma 8.9 (Unconditionality lemma)**.** *Suppose that the $\mathrm{H}^\infty(\mathrm{S}_\varphi)$-calculus for L is bounded with bound C_L. Given $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$, there exists a constant $C_{f,\varphi}$ that depends only on f and φ, such that*

$$\Big\| \sum_{j \in \mathbb{Z}} a_j f(t2^j L) \Big\|_{\mathcal{L}(H)} \leq C_L C_{f,\varphi} \|a\|_{\ell^\infty(\mathbb{Z})}$$

for all $t > 0$ and all sequences $(a_j) \in \ell^\infty(\mathbb{Z})$ with only finitely many non-zero elements.

Proof. We pick $C, s > 0$ such that $|f(z)| \leq C \min(|z|^s, |z|^{-s})$ for all $z \in \mathrm{S}_\varphi$ and obtain

$$\begin{aligned} \Big\| \sum_{j \in \mathbb{Z}} a_j f(t2^j L) \Big\|_{\mathcal{L}(H)} &\leq C_L \Big\| \sum_{j \in \mathbb{Z}} a_j f(t2^j \,\cdot) \Big\|_{\infty,\varphi} \\ &\leq C_L C \|a\|_{\ell^\infty(\mathbb{Z})} \sup_{z \in \mathrm{S}_\varphi} \sum_{j \in \mathbb{Z}} \min\big(|2^j tz|^s, |2^j tz|^{-s}\big). \end{aligned}$$

For fixed $z \in \mathrm{S}_\varphi$, we let $j(z)$ be the unique integer that satisfies $1 \leq |2^{j(z)} tz| < 2$. Since $2^{j-j(z)} \leq |2^j tz| < 2^{j-j(z)+1}$, we can complete the proof noting that

$$\sum_{j \in \mathbb{Z}} \min(|2^j tz|^s, |2^j tz|^{-s}) \leq \sum_{j < j(z)} 2^{s(j-j(z)+1)} + \sum_{j \geq j(z)} 2^{-s(j-j(z))} = \frac{2}{1 - 2^{-s}}.$$ □

We come to the proof of McIntosh's theorem.

Proof of Theorem 8.4. We have '$(a) \iff (b)$' by Lemma 8.7 — if f, g are defined on sectors of different angle, then we apply the lemma on the smaller one. Since clearly '$(d) \implies (c)$', we complete the proof by showing '$(b) \implies (d)$' and '$(c) \implies (a)$'.

'$(b) \implies (d)$':

We fix any angle $\varphi \in (\varphi_L, \pi)$. Due to Corollary 7.5, it suffices to prove for all $h \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ the norm bound $\|h(L)\|_{\mathcal{L}(H)} \leq c\|h\|_{\infty,\varphi}$ with c independent of h. We fix any auxiliary function $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi) \setminus \{0\}$. The quadratic estimates for L yield

$$\|h(L)u\|^2 \leq C \int_0^\infty \|f(sL)h(L)u\|^2 \, \frac{\mathrm{d}s}{s} \qquad (u \in H).$$

In order to bound the right-hand side, we repeat word-by-word the proof of Lemma 8.7 with $g = f$ except that we bound $f(sL)h(L)f^\sharp(tL)f(tL)u$ in norm via Remark 8.6 instead of Lemma 8.5. This procedure gives us

$$\int_0^\infty \|f(sL)h(L)u\|^2 \,\frac{\mathrm{d}s}{s} \leq \|h\|_{\infty,\varphi}^2 \|\zeta\|_{\mathrm{L}^1((0,\infty);\mathrm{d}t/t)}^2 \bigg(\int_0^\infty \|f(tL)u\|^2 \,\frac{\mathrm{d}t}{t}\bigg).$$

On the right-hand side, we can use the quadratic estimates again. Putting it all together, we have shown the desired bound

$$\|h(L)u\|^2 \leq C^2 \|h\|_{\infty,\varphi}^2 \|\zeta\|_{\mathrm{L}^1((0,\infty);\mathrm{d}t/t)}^2 \|u\|^2.$$

'(c) $\Longrightarrow$ (a)':

We assume that L has a bounded H^∞-calculus of angle $\varphi \in (\varphi_L, \pi)$ with bound C_L and take any $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi) \setminus \{0\}$. Given $u \in H$, we break the quadratic integral into dyadic pieces and perform a change of variables to obtain

$$\begin{aligned} \int_0^\infty \|f(tL)u\|^2 \,\frac{\mathrm{d}t}{t} &= \lim_{N\to\infty} \sum_{j=-N}^N \int_{2^j}^{2^{j+1}} \|f(tL)u\|^2 \,\frac{\mathrm{d}t}{t} \\ &= \lim_{N\to\infty} \int_1^2 \sum_{j=-N}^N \|f(t2^jL)u\|^2 \,\frac{\mathrm{d}t}{t}. \end{aligned} \tag{8.3}$$

For fixed j and t, we expand the norm of the integrand as follows, using the orthonormal system $e_j := \pi^{-1/2}\mathrm{e}^{\mathrm{i}j\cdot}$, $j \in \mathbb{Z}$, of $\mathrm{L}^2(0,\pi)$ and the unconditionality lemma:

$$\begin{aligned} \sum_{j=-N}^N \|f(t2^jL)u\|^2 &= \sum_{j,k=-N}^N \int_0^\pi e_j(s)\overline{e_k(s)} \langle f(t2^jL)u, f(t2^kL)u\rangle \,\mathrm{d}s \\ &= \int_0^\pi \bigg\| \sum_{j=-N}^N e_j(s) f(t2^jL)u \bigg\|^2 \,\mathrm{d}s \\ &\leq \int_0^\pi \big(C_L C_{f,\varphi} \pi^{-1/2}\|u\|\big)^2 \,\mathrm{d}s = (C_L C_{f,\varphi})^2 \|u\|^2. \end{aligned} \tag{8.4}$$

Together, (8.3) and (8.4) yield upper quadratic estimates for L with auxiliary function f. The same is true for L^* since this operator has a bounded H^∞-calculus with the same bound and angle according to Lemma 7.9. Now, lower quadratic estimates for $L = (L^*)^*$ with auxiliary function f follow from Lemma 8.8 and the proof is complete. □

Combining Theorem 8.4 with Theorem 7.12 yields the following:

Corollary 8.10. *Injective m-accretive operators L in H satisfy quadratic estimates and have a bounded H^∞-calculus with $\varphi_L^\infty = \varphi_L$. In particular, elliptic operators in divergence form with Dirichlet boundary conditions have these properties.*

As far as the H^∞-calculus is concerned, we have now improved the angle from $\varphi_L^\infty \leq \pi/2$ in Theorem 7.12 to the best possible $\varphi_L^\infty = \varphi_L$, at the expense of losing the universal constant 1 in the corresponding estimate. It is instructive to recapitulate how we got there by a two-step procedure: First, we have used algebraic identities for the resolvent on the imaginary axis to prove von Neumann's inequality for holomorphic functions on large sectors. Through the equivalence with the Fourier-inspired quadratic estimates, we have then removed the assumption on the angle so that in the end our holomorphic functions need not even be defined on the imaginary axis.

Remark 8.11. Occasionally, it is useful to know how the equivalence constant for quadratic estimates depends on the operator and its bounded H^∞-calculus and vice versa. As an exercise, we advise the reader to go through the proof and work out this dependence qualitatively, see Exercise 8.2.

8.2. Kato's theorem on subcritical fractional powers

As in Lecture 2, let V be another Hilbert space that is continuously and densely embedded into H. At the beginning of the 1960s, Kato investigated the abstract, non-autonomous evolution equation

$$\frac{\mathrm{d}}{\mathrm{d}t}u(t) + L(t)u(t) = f(t), \qquad (0 \leq t < T),$$

in which the unknown is $u = u(t)$ and each 'coefficient' $L(t)$ is an m-accretive operator associated with a sesquilinear form $a(t)$ on V. In the exercises to Lecture 4 we have seen that the domain of $L(t)$ can, potentially, undergo drastic changes. In order to obtain a unique solution for given initial data $u(0)$ in a suitable sense, Kato was lead to making the *stability assumption* that for some $\alpha = 1/m$, $m \in \mathbb{N}$, the domains of fractional powers $L(t)^\alpha$ are independent of t. Subsequently, in 1961, he proved the remarkable result [Kat61] that for $\alpha \in (0, 1/2)$ said independence is a mere consequence of comparability of the quadratic forms in the following sense.

Theorem 8.12 (Kato)**.** *Let $a_1, a_2 \colon V \times V \to \mathbb{C}$ be bounded, elliptic and sectorial sesquilinear forms and let $\alpha \in (0, 1/2)$. Suppose that*

$$\operatorname{Re} a_2(u) \lesssim \operatorname{Re} a_1(u) \qquad (u \in V). \tag{8.5}$$

If L_1, L_2 are the associated operators in H, then $\operatorname{dom}(L_1^\alpha) \subseteq \operatorname{dom}(L_2^\alpha)$ *and*

$$\|L_2^\alpha u\| \lesssim \|L_1^\alpha u\| \qquad (u \in \operatorname{dom}(L_1^\alpha)).$$

Remark 8.13. Reversing the roles of a_1 and a_2, we see that the two-sided estimate $\operatorname{Re} a_1(u) \simeq \operatorname{Re} a_2(u)$ for all $u \in V$ implies $\operatorname{dom}(L_1^\alpha) = \operatorname{dom}(L_2^\alpha)$ with a two-sided estimate for fractional powers. This always happens for $a_2 = a_1^*$, corresponding to $L_2 = L_1^*$ by Proposition 2.24, and we obtain for any operator L_1 as above that $\operatorname{dom}(L_1^\alpha) = \operatorname{dom}((L_1^*)^\alpha)$, whenever $\alpha \in (0, 1/2)$.

Our formulation of Theorem 8.12 slightly differes from Kato's original work and for the proof we follow the idea of Carbonaro and Dragičević [CD23] to use quadratic estimates, which back in 1961 were pure science-fiction. Assumption (8.5) will be brought into play through the following elementary lemma.

Lemma 8.14. *Let* $a_1, a_2 \colon V \times V \to \mathbb{C}$ *be bounded sesquilinear forms that satisfy* (8.5) *and suppose that* a_2 *is sectorial. Then for all* $w \in \operatorname{dom}(L_1)$ *and* $z \in \operatorname{dom}(L_2)$ *the associated operators satisfy*

$$|\langle w, L_2 z\rangle| \lesssim \|L_1 w\|^{1/2} \|w\|^{1/2} \|L_2 z\|^{1/2} \|z\|^{1/2}.$$

Proof. The definition of L_2 and the Cauchy–Schwarz inequality for sectorial forms (Exercise 2.1) yield

$$|\langle w, L_2 z\rangle| = |a_2(z, w)| \lesssim \big(\mathrm{Re}(a_2(w))\big)^{1/2} \big(\mathrm{Re}(a_2(z))\big)^{1/2}.$$

On the right-hand side, (8.5) allows us to replace $a_2(w)$ by $a_1(w)$. Thus, rewriting everything in terms of operators, we complete the estimate by

$$|\langle w, L_2 z\rangle| \lesssim |\langle L_1 w, w\rangle|^{1/2} |\langle L_2 z, z\rangle|^{1/2} \leq \|L_1 w\|^{1/2} \|w\|^{1/2} \|L_2 z\|^{1/2} \|z\|^{1/2}. \qquad \square$$

Proof of Theorem 8.12. For starters, we recall that L_1, L_2 are m-accretive (Theorem 2.21) and so are their adjoints that are associated with the adjoint sesquilinear forms a_1^*, a_2^* (Proposition 2.24).

We present the proof in the case that L_1 and L_2 are both injective. Then also their adjoints are injective (Lemma 6.7). In Exercise 8.3 you will learn how to eliminate this additional assumption by a minor refinement of our reasoning below. The proof comes in four steps. Throughout, we work with arbitrary elements $u \in \operatorname{dom}(L_1^\alpha)$, $v \in \operatorname{dom}((L_2^*)^\alpha)$ and implicit constants are independent of u, v.

Step 1: It suffices to show the 'bilinear' estimate $|\langle u, (L_2^*)^\alpha v\rangle| \lesssim \|L_1^\alpha u\| \|v\|$.

Admitting this 'bilinear' estimate, $v \mapsto \langle u, (L_2^*)^\alpha v\rangle$ can be extended by density to an anti-linear functional on H since $(L_2^*)^\alpha$ is densely defined by Lemma 6.4. The Riesz representation theorem presents us with some $w \in H$ of norm $\|w\| \lesssim \|L_1^\alpha u\|$ such that $\langle u, (L_2^*)^\alpha v\rangle = \langle w, v\rangle$ for all $v \in \operatorname{dom}((L_2^*)^\alpha)$. This means $(u, w) \in ((L_2^*)^\alpha)^*$ and since we have $((L_2^*)^\alpha)^* = L_2^\alpha$ by Corollary 6.11, we find $\|L_2^\alpha u\| = \|w\| \lesssim \|L_1^\alpha u\|$ as required.

Step 2: A reproducing formula for the bilinear term.

Our goal is to bound $\langle u, (L_2^*)^\alpha v\rangle$ by an integral that allows us to use quadratic estimates for L_1 on u and for L_2^* on v. To this end, we introduce the function

$$\Psi \colon (0, \infty) \to \mathbb{C}, \quad t \mapsto \big\langle (1 + tL_1)^{-1} u, (1 + tL_2^*)^{-1} (L_2^*)^\alpha v \big\rangle.$$

According to Proposition 1.15, it is continuously differentiable with

$$\begin{aligned}\Psi'(t) = &-\big\langle L_1(1+tL_1)^{-2}u, (1+tL_2^*)^{-1}(L_2^*)^\alpha v\big\rangle \\ &- \big\langle (1+tL_1)^{-1}u, L_2^*(1+tL_2^*)^{-2}(L_2^*)^\alpha v\big\rangle\end{aligned}$$

and thanks to the sectoriality of L_1 and L_2 and Proposition 2.4 it is bounded with limits

$$\lim_{t\to 0}\Psi(t) = \langle u, (L_2^*)^\alpha v\rangle \quad \text{and} \quad \lim_{t\to\infty}\Psi(t) = 0.$$

By the fundamental theorem of calculus, we get

$$|\langle u, (L_2^*)^\alpha v\rangle| = \left|\int_0^\infty \Psi'(t)\,\mathrm{d}t\right| \leq \int_0^\infty |\Psi'(t)|\,\mathrm{d}t \leq \mathrm{I} + \mathrm{II},$$

where

$$\begin{aligned}\mathrm{I} &:= \int_0^\infty \big|\big\langle L_1(1+tL_1)^{-2}u, (1+tL_2^*)^{-1}(L_2^*)^\alpha v\big\rangle\big|\,\mathrm{d}t, \\ \mathrm{II} &:= \int_0^\infty \big|\big\langle (1+tL_1)^{-1}u, L_2^*(1+tL_2^*)^{-2}(L_2^*)^\alpha v\big\rangle\big|\,\mathrm{d}t.\end{aligned}$$

Our goal is to gain control by $\|L_1^\alpha u\|\|v\|$ on both terms.

Step 3: Estimate of the first term.

This one is easier, because there are auxiliary functions f, g of class $\mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ on any sector for which it takes the form

$$\mathrm{I} = \int_0^\infty \big|\big\langle f(tL_1)L_1^\alpha u, g(tL_2^*)v\big\rangle\big|\,\frac{\mathrm{d}t}{t}.$$

Indeed, we take $f := \mathbf{z}^{1-\alpha}(1+\mathbf{z})^{-2}$, $g := \mathbf{z}^\alpha(1+\mathbf{z})^{-1}$ and sort out the powers of t. Thus, we can use the Cauchy–Schwarz inequality in order to bring quadratic estimates for injective m-accretive operators (Corollary 8.10) into play and conclude that

$$\mathrm{I} \leq \bigg(\int_0^\infty \|f(tL_1)L_1^\alpha u\|^2\,\frac{\mathrm{d}t}{t}\bigg)^{1/2}\bigg(\int_0^\infty \|g(tL_2^*)v\|^2\,\frac{\mathrm{d}t}{t}\bigg)^{1/2} \lesssim \|L_1^\alpha u\|\|v\|.$$

Step 4: Estimate of the second term.

The same strategy does not apply directly to II, because we are missing powers of L_1 on the left. In order to fix this, we use another reproducing formula, keeping $t > 0$ fixed and introducing

$$\Upsilon\colon (0,\infty) \to \mathbb{C}, \quad s \mapsto \big\langle (1+sL_1)^{-1}u, L_2^*(1+tL_2^*)^{-2}(L_2^*)^\alpha v\big\rangle.$$

By the same argument as in Step 2, this function vanishes in the limit as $s \to \infty$ and is differentiable with

$$\Upsilon'(s) = -\big\langle L_1(1+sL_1)^{-2}u, L_2^*(1+tL_2^*)^{-2}(L_2^*)^\alpha v\big\rangle.$$

By means of the regularly decaying holomorphic functions $f = \mathbf{z}^{1-\alpha}(1+\mathbf{z})^{-2}$ and $h = \mathbf{z}^\alpha(1+\mathbf{z})^{-2}$ we can write

$$\Upsilon'(s) = -s^{\alpha-1}t^{-\alpha}\big\langle f(sL_1)L_1^\alpha u, L_2^* h(tL_2^*)v\big\rangle,$$

and it will be useful to keep in mind that $f(sL_1)L_1^\alpha u$ belongs to the domain of L_1: This is due to Theorem 5.15, because $\mathbf{z}f = \mathbf{z}^{2-\alpha}(1+\mathbf{z})^{-2}$ is regularly decaying and thus $sL_1 f(sL_1) = (\mathbf{z}f)(sL_1)$ is bounded. Using the fundamental theorem of calculus again, we obtain

$$\begin{aligned}
\mathrm{II} &= \int_0^\infty \big|\Upsilon(t)\big| \,\mathrm{d}t = \int_0^\infty \bigg|\int_t^\infty \Upsilon'(s)\,\mathrm{d}s\bigg| \,\mathrm{d}t \\
&\leq \int_0^\infty \int_t^\infty \Big(\frac{s}{t}\Big)^\alpha \big|\big\langle f(sL_1)L_1^\alpha u, L_2^* h(tL_2^*)v\big\rangle\big| \,\frac{\mathrm{d}s}{s}\,\mathrm{d}t \qquad (8.6)\\
&\overset{s=tr}{=} \int_1^\infty r^\alpha \int_0^\infty \big|\big\langle f(rtL_1)L_1^\alpha u, L_2^* h(tL_2^*)v\big\rangle\big| \,\mathrm{d}t\,\frac{\mathrm{d}r}{r}.
\end{aligned}$$

For the moment, let us focus on the inner integral. It is tempting to simply write it as

$$\int_0^\infty \big|\big\langle f(rtL_1)L_1^\alpha u, (\mathbf{z}h)(tL_2^*)v\big\rangle\big| \,\frac{\mathrm{d}t}{t},$$

apply the Cauchy–Schwarz inequality and then quadratic estimates as in Step 3, but this would give a uniform bound with respect to r leading nowhere, since we still have to integrate with respect to r in (8.6). It is at this point that the assumption (8.5) will allow us to produce additional decay in r. Indeed, Lemma 8.14 applied to the bounded, sectorial forms a_1 and a_2^* lets us gain control of the t-integral by

$$\begin{aligned}
&\int_0^\infty \big|\big\langle f(rtL_1)L_1^\alpha u, L_2^* h(tL_2^*)v\big\rangle\big| \,\mathrm{d}t \\
&\lesssim \int_0^\infty \big\|L_1 f(rtL_1)L_1^\alpha u\big\|^{1/2}\big\|f(rtL_1)L_1^\alpha u\big\|^{1/2}\big\|L_2^* h(tL_2^*)v\big\|^{1/2}\big\|h(tL_2^*)v\big\|^{1/2}\,\mathrm{d}t \\
&\leq r^{-1/2}\int_0^\infty \big\|(\mathbf{z}f)(rtL_1)L_1^\alpha u\big\|^{1/2}\big\|f(rtL_1)L_1^\alpha u\big\|^{1/2}\big\|(\mathbf{z}h)(tL_2^*)v\big\|^{1/2}\big\|h(tL_2^*)v\big\|^{1/2}\,\frac{\mathrm{d}t}{t} \\
&\leq r^{-1/2}\bigg(\int_0^\infty \big\|(\mathbf{z}f)(rtL_1)L_1^\alpha u\big\|^2\,\frac{\mathrm{d}t}{t}\bigg)^{1/4}\bigg(\int_0^\infty \big\|f(rtL_1)L_1^\alpha u\big\|^2\,\frac{\mathrm{d}t}{t}\bigg)^{1/4} \\
&\qquad\cdot\bigg(\int_0^\infty \big\|(\mathbf{z}h)(tL_2^*)v\big\|^2\,\frac{\mathrm{d}t}{t}\bigg)^{1/4}\bigg(\int_0^\infty \big\|h(tL_2^*)v\big\|^2\,\frac{\mathrm{d}t}{t}\bigg)^{1/4} \\
&\lesssim r^{-1/2}\|L_1^\alpha u\|\|v\|,
\end{aligned}$$

where we have used quadratic estimates (and the change of variables $rt = s$) in the final step. Using this bound back in (8.6), we find

$$\mathrm{II} \lesssim \int_1^\infty r^{\alpha-1/2} \|L_1^\alpha u\| \|v\| \, \frac{\mathrm{d}r}{r} = \frac{1}{1/2 - \alpha} \|L_1^\alpha u\| \|v\| \tag{8.7}$$

precisely because of the assumption $\alpha < 1/2$, which is used for the one and only time here, whereas all other steps would have worked for $\alpha \in (0, 1)$. □

Let us discover some consequences of Theorem 8.12 when applied to elliptic operators in divergence form subject to Dirichlet boundary conditions on a non-empty open set $\Omega \subseteq \mathbb{R}^n$ as in Lecture 3. Let $L_1 = (-\Delta)_{\mathrm{H}_0^1(\Omega)}$ be the negative Dirichlet Laplacian and let $L_2 = L = -\operatorname{div}(A\nabla\,\cdot)$ be any such operator. The common domain for the sesquilinear forms a_1, a_2 is $V = \mathrm{H}_0^1(\Omega)$ and, by ellipticity, we have for all $u, v \in V$ that

$$\lambda \operatorname{Re} a_1(u) = \lambda \int_\Omega |\nabla u|^2 \, \mathrm{d}x \leq \operatorname{Re} a_2(u) \leq \Lambda \int_\Omega |\nabla u|^2 \, \mathrm{d}x = \Lambda \operatorname{Re} a_1(u).$$

Hence, Theorem 8.12 applies and yields the following:

Corollary 8.15. *Let $\Omega \subseteq \mathbb{R}^n$ be a non-empty open set, $L = -\operatorname{div}(A\nabla\,\cdot)$ be an elliptic operator in divergence form with Dirichlet boundary conditions on Ω. Then for all $\alpha \in (0, 1/2)$ we have $\operatorname{dom}(L^\alpha) = \operatorname{dom}((-\Delta)^\alpha_{\mathrm{H}_0^1(\Omega)})$ along with the homogeneous estimate*

$$\|L^\alpha u\|_{\mathrm{L}^2(\Omega)} \simeq \|(-\Delta)^\alpha_{\mathrm{H}_0^1(\Omega)} u\|_{\mathrm{L}^2(\Omega)} \qquad (u \in \operatorname{dom}(L^\alpha)).$$

In particular, $\operatorname{dom}(L^\alpha)$ *is independent of A.*

The corollary is all the more surprising since $\operatorname{dom}(L)$ heavily depends on A and is in general different from $\operatorname{dom}((-\Delta)_{\mathrm{H}_0^1(\Omega)})$, see Exercises 4.6 and 4.3. A point left open is if (and how) we can actually determine the domains of fractional powers of the negative Dirichlet Laplacian from the 'data' $\mathrm{H}_0^1(\Omega)$ and α. On a general open set Ω, this question can be answered through the theory of complex interpolation spaces that you will encounter in the project phase: Vaguely speaking, $\operatorname{dom}((-\Delta)^\alpha_{\mathrm{H}_0^1(\Omega)})$ sits in between $\mathrm{L}^2(\Omega)$ and $\operatorname{dom}((-\Delta)^{1/2}_{\mathrm{H}_0^1(\Omega)})$ and, by Theorem 6.12, the latter coincides with $\mathrm{H}_0^1(\Omega)$.

In the special case $\Omega = \mathbb{R}^n$ we have determined $\operatorname{dom}((-\Delta)^\alpha) =: \mathrm{H}^{2\alpha}(\mathbb{R}^n)$ in Proposition 4.13. The mathematical technology that we developed for L over the last lectures has enabled us to generalize this result to general elliptic operators:

Corollary 8.16. *Let $L = -\operatorname{div}(A\nabla\,\cdot)$ be an elliptic operator in divergence form on $\mathbb{R}^n$ and let $\alpha \in (0, 1/2)$. Then $\operatorname{dom}(L^\alpha) = \mathrm{H}^{2\alpha}(\mathbb{R}^n)$ along with the homogeneous estimate*

$$\|L^\alpha u\|^2_{\mathrm{L}^2(\mathbb{R}^n)} \simeq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|u(x) - u(y)|^2}{|x - y|^{4\alpha}} \, \frac{\mathrm{d}x \, \mathrm{d}y}{|x - y|^n} \qquad (u \in \operatorname{dom}(L^\alpha)).$$

You have seen that for $\alpha = 1/2$, related to the Kato property discussed in Section 6.4, the technique of proof breaks down brutally in the very final estimate (8.7). There can be no way of saving Theorem 8.12 for $\alpha = 1/2$ in this generality, because McIntosh [McI72] constructed an operator L_1 for which $\mathrm{dom}(\sqrt{L_1}) \neq \mathrm{dom}(\sqrt{L_1^*})$. Concerning the Kato property we have gone as far as we possibly can with functional calculus alone and now, methods from real harmonic analysis have to enter the game . . .

8.3. Exercises

Exercise 8.1 (Schur test)**.** For a fair comparison with the proof of Lemma 8.5, we present the classical Schur test for integral operators. Let $(X, \mu), (Y, \nu)$ be σ-finite measure spaces and let $K\colon X \times Y \to \mathbb{C}$ be a measurable kernel for which there exist measurable functions $p\colon X \to (0, \infty)$, $q\colon Y \to (0, \infty)$ and constants $\alpha, \beta \geq 0$ such that

$$\int_Y |K(x, y)| q(y) \, \mathrm{d}\nu(y) \leq \alpha p(x) \qquad (\text{a.e. } x \in X),$$

$$\int_X |K(x, y)| p(x) \, \mathrm{d}\mu(x) \leq \beta q(y) \qquad (\text{a.e. } y \in Y).$$

Prove that the operator defined by

$$Tf(x) = \int_Y K(x, y) f(y) \, \mathrm{d}\nu(y)$$

is bounded $\mathrm{L}^2(Y, \nu) \to \mathrm{L}^2(X, \mu)$ with norm not exceeding $\sqrt{\alpha\beta}$.

Exercise 8.2 (Quantitative McIntosh theorem)**.** Let L be an injective sectorial operator in a Hilbert space H and let $\varphi \in (\varphi_L, \pi)$.

(a) Suppose that L satisfies for some constant $C \geq 1$ the quadratic estimate

$$C^{-1}\|u\|^2 \leq \int_0^\infty \|tL(1 + tL)^{-2}u\|^2 \, \frac{\mathrm{d}t}{t} \leq C\|u\|^2 \qquad (u \in H).$$

Which further 'data' of L do you need to give a bound for its H^∞-calculus of angle φ? Write down such a bound explicitly.

(b) Conversely, suppose that L has a bounded H^∞-calculus of angle φ with bound C_L. Can you determine a constant C such that the quadratic estimate in (a) holds?

Exercise 8.3 (The general version of Theorem 8.12)**.** So far, we have only proved Theorem 8.12 in the case that L_1 and L_2 are both injective. In this exercise you will refine the proof in order to cover the general case.

(a) Carefully go through the proof of Theorem 8.12 in the injective case and single out the four spots where this assumptions has been used. Don't read further until you found them.

(b) Eliminate the use of Proposition 2.4 when determining the limit $\lim_{t\to\infty} \Psi(t)$ in Step 2.

Hint: Make use of L_2^α on the right-hand side.

(c) Use Exercise 7.5 to get the upper quadratic estimates in Steps 3 and 4 in the general case.

(d) Prove that $\ker(L_1) \subseteq \ker(L_2)$.

(e) Conclude that also $\lim_{s\to\infty} \Upsilon(s) = 0$ in Step 4 continues to hold.

Exercise 8.4 (Kato's definition of the square root)**.** Let L be an operator in H. A linear operator R in H with $R^2 = L$ is called *a square root* of L. In general, just as for numbers and matrices, there can be many meaningful square roots.

In this exercise we are going to prove that an m-accretive operator has a unique square root in the class of m-accretive operators. This brings us back to Kato's original definition of the square root [Kat95, Ch. V.3.11]. From now on, let L be m-accretive. We proceed as follows.

(a) Recall that $\sqrt{L}$ defined by our functional calculus is m-accretive and even sectorial of angle $\pi/4$.

For uniqueness, we let R be an m-accretive square root of L. The subtlety in the statement we are trying to prove is that we cannot compute $\sqrt{R^2}$ in the functional calculus for R, because $\sqrt{\mathbf{z}^2}$ is not holomorphic on sectors of angle larger than $\pi/2$.

(b) Argue that $\operatorname{dom}(L^2) = \operatorname{dom}(R^4)$ is a core for both L and R.

(c) Conclude that it suffices to prove $\sqrt{L}u = Ru$ for $u \in \operatorname{dom}(L^2)$.

(d) Prove that $\sqrt{L}Ru = R\sqrt{L}u$.

Hint: Use Exercise 5.4 with $T = (1 + R)^{-1}$.

(e) Conclude that $\big(\sqrt{L} + R\big)\big(\sqrt{L} - R\big)u = 0$.

For u fixed as above, we introduce $v := (\sqrt{L} - R)u$ and our goal is to show $v = 0$.

(f) Argue that $\operatorname{Re}\big(\big\langle \sqrt{L}v, v\big\rangle\big) = 0$.

(g) Use this information to prove that for every $t > 0$ we have

$$\operatorname{Re}\big(\langle v - (1 + tL)^{-1}v, v\rangle\big) = 0.$$

(h) Complete the proof under the additional assumption that L is injective.

In order to also complete the proof for non-injective L, we split $u = u_{\ker} + u_{\overline{\operatorname{ran}}}$ according to the topological splitting $H = \ker(L) \oplus \overline{\operatorname{ran}(L)}$ in Proposition 2.4 (b).

8. *Quadratic estimates vs. functional calculus*

(i) Carefully check that the argument above can be used to prove $\sqrt{L}u_{\overline{\text{ran}}} = Ru_{\overline{\text{ran}}}$.

(j) Prove that $\ker(L) \subseteq \ker(\sqrt{L}) \cap \ker(R)$.

(k) Conclude.

9. The Hardy–Littlewood maximal operator

Within the realm of harmonic analysis, a fundamental principle asserts that the study of functions can be effectively pursued by examining their averages on multiple scales. Approximation by convolution is just one incarnation of the general principle. One operator lies at the heart of the theory — the Hardy–Littlewood maximal operator. In this lecture, we study its boundedness, which in return has a variety of further consequences of which we can only present some showcases: Lebesgue differentiation, almost everywhere differentiability of Sobolev functions on the real line and Carleson's lemma, establishing a link between averaging operators and quadratic estimates.

Notation 9.1. For an integrable function u on a set E of positive measure, we denote the average on E by

$$(u)_E := \fint_E u \,\mathrm{d}y := \frac{1}{|E|}\int_E u \,\mathrm{d}y.$$

Balls are always open and given a ball B and a number $c > 0$, we write cB for the concentric ball with c-times the radius. For convenience, we let $\omega_n := |B(0,1)|$ be the Lebesgue measure of the n-dimensional unit ball. Finally, $\mathbb{R}^{n+1}_+ := \mathbb{R}^n \times (0,\infty)$ is the upper half space in dimension $(n+1)$.

9.1. Averages and the maximal operator

Let $u\colon \mathbb{R}^n \to \mathbb{C}$ be any locally integrable function. Replacing u by its averages at scale $t > 0$ gives a slightly better behaved function

$$x \mapsto \fint_{B(x,t)} u(y)\,\mathrm{d}y$$

that is continuous as a consequence of the dominated convergence theorem. However, sending $t \to 0$, we do not yet know whether averages stay bounded for fixed x, let alone if they converge to $u(x)$, unless u is continuous itself. We sometimes call these 'rough averages', because they can be written as

$$\fint_{B(x,t)} u(y)\,\mathrm{d}y = \frac{1}{\omega_n t^n}\int_{\mathbb{R}^n} \mathbf{1}_{B(0,1)}\Big(\frac{x-y}{t}\Big)u(y)\,\mathrm{d}y = (\chi_t * u)(x)$$

by means of the discontinuous function $\chi := \omega_n^{-1}\mathbf{1}_{B(0,1)}$ and its associated mollifier. 'Smooth averages' $(\theta_t * u)(x)$, using a function $\theta \in \mathrm{C}_\mathrm{c}^\infty(B(0,1))$ with $\int_{\mathbb{R}^n} \theta \,\mathrm{d}y = 1$, have even better properties at fixed scale, see Proposition 3.3 (b), but the same questions arise when trying to pass to the limit. Both types of averages are controlled by the following maximal function.

Definition 9.2. *Let $u \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n)$. We define the* maximal function *of u by*

$$(\mathcal{M}u)(x) := \sup_{B \ni x} \fint_B |u| \,\mathrm{d}y \qquad (x \in \mathbb{R}^n),$$

where the possibly infinite supremum is taken over all balls B that contain x, and call $\mathcal{M}$ the Hardy–Littlewood maximal operator.

The maximal operator is sublinear, i.e., for $u, v \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $\lambda \in \mathbb{C}$ we have

$$|\mathcal{M}(u+v)| \leq |\mathcal{M}u| + |\mathcal{M}v| \quad \text{and} \quad |\mathcal{M}(\lambda u)| = |\lambda||\mathcal{M}u|$$

everywhere on $\mathbb{R}^n$. You might have expected that we only consider balls centered in x. Since every ball containing x is contained in a ball of doubled radius centered in x, this would have led to a comparable operator. Our uncentered version sometimes allows for simpler arguments, for example in the following lemma.

Lemma 9.3. *If $u \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n)$, then the maximal function $\mathcal{M}u \colon \mathbb{R}^n \to [0,\infty]$ is lower semicontinuous. In particular, it is measurable.*

Proof. Lower semi-continuity means that for all $\lambda \geq 0$ the superlevel sets

$$O_\lambda := \{x \in \mathbb{R}^n \mid (\mathcal{M}u)(x) > \lambda\}$$

are open. But if $x \in O_\lambda$, then x is contained in a ball B with $\fint_B |u| \,\mathrm{d}y > \lambda$ and, again by definition of the maximal operator, we conclude $B \subseteq O_\lambda$. □

When it comes to bounding the Hardy–Littlewood maximal operator on Lebesgue spaces, one case is particularly easy: Since averages of functions are always bounded by their essential supremum, we obtain for all $u \in \mathrm{L}^\infty(\mathbb{R}^n)$ that

$$\|\mathcal{M}u\|_{\mathrm{L}^\infty(\mathbb{R}^n)} \leq \|u\|_{\mathrm{L}^\infty(\mathbb{R}^n)}. \tag{9.1}$$

The situation is dramatically different on the other side of the Lebesgue scale.

Lemma 9.4. *If $u \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $\mathcal{M}u \in \mathrm{L}^1(\mathbb{R}^n)$, then already $u = 0$.*

Proof. We fix $r > 0$ and show $|u| = 0$ almost everywhere on $B(0,r)$. To this end, we bound the maximal function for $|x| \geq r$ from below by

$$(\mathcal{M}u)(x) \geq \fint_{B(0,2|x|)} |u| \,\mathrm{d}y \geq \frac{1}{\omega_n (2|x|)^n} \int_{B(0,r)} |u| \,\mathrm{d}y.$$

Since $\int_{\mathbb{R}^n \setminus B(0,r)} \frac{1}{|x|^n} \, \mathrm{d}x = \infty$, the left-hand side can only be integrable if we have $|u| = 0$ almost everywhere on $B(0, r)$. ◻

The proof of the lemma suggests that global integrability is the main issue, but in fact $\mathcal{M}u$ for $u \in \mathrm{L}^1(\mathbb{R}^n)$ does not even have to be locally integrable. For an explicit counterexample, we refer to Exercise 9.2. It turns out, however, that a weaker type of L^1-bound does hold true for the maximal operator.

Theorem 9.5 (Hardy–Littlewood I)**.** *If* $u \in \mathrm{L}^1(\mathbb{R}^n)$*, then for every* $\lambda > 0$ *we have the weak-type bound*

$$\left|\left\{x \in \mathbb{R}^n \mid (\mathcal{M}u)(x) > \lambda\right\}\right| \leq \frac{3^n}{\lambda} \|u\|_{\mathrm{L}^1(\mathbb{R}^n)}.$$

In particular, $\mathcal{M}u$ *is finite almost everywhere.*

The terminology 'weak-type bound' is justified by *Markov's inequality*: If $v \in \mathrm{L}^1(\mathbb{R}^n)$ is an integrable function and $\lambda > 0$, then

$$\lambda\left|\left\{x \in \mathbb{R}^n \mid |v(x)| > \lambda\right\}\right| \leq \|v\|_{\mathrm{L}^1(\mathbb{R}^n)}. \tag{9.2}$$

Theorem 9.5 and the considerations above show that the left-hand side above can be bounded uniformly with respect to $\lambda > 0$ even if v is not integrable.

We need two lemmas to prepare the proof of Theorem 9.5 and the subsequent results.

Lemma 9.6 (Layer cake formula)**.** *If* $p \in (0, \infty)$ *and* u *is a measurable function on a* σ*-finite measure space* (X, μ)*, then*

$$\int_X |u(x)|^p \, \mathrm{d}\mu(x) = \int_0^\infty p\lambda^{p-1} \mu\big(\left\{x \in X : |u(x)| > \lambda\right\}\big) \, \mathrm{d}\lambda.$$

Proof. Simply apply Tonelli's theorem after having written

$$\int_X |u(x)|^p \, \mathrm{d}\mu(x) = \int_X \int_0^{|u(x)|} p\lambda^{p-1} \, \mathrm{d}\lambda \, \mathrm{d}\mu(x).$$ ◻

Lemma 9.7 (Vitali covering lemma)**.** *Every finite collection* $\mathcal{B}$ *of balls in* $\mathbb{R}^n$ *admits a sub-collection* $\mathcal{B}'$ *of pairwise disjoint balls such that*

$$\bigcup_{B \in \mathcal{B}} B \subseteq \bigcup_{B' \in \mathcal{B}'} 3B'.$$

Proof. Since $\mathcal{B}$ is a finite collection, we can write it in decreasing order of radius as $B_1, \ldots, B_N$. We select the balls for $\mathcal{B}'$ by a simple greedy algorithm:

```
Select B_1.
For each j = 2, ..., N:
Select B_j if it is disjoint to all previously selected balls.
Otherwise discard it.
```

By construction, the algorithm returns a collection $\mathcal{B}'$ of pairwise disjoint balls. In order to check the '3-covering property', let $x \in B_j$ for some j. There is nothing more to do if $B_j \in \mathcal{B}'$, so assume B_j has been discarded by the algorithm. This means that there was some index $k < j$ such that $B_k \in \mathcal{B}'$ and $B_j \cap B_k$ contains some element y. Writing $B_j = B(x_j, r_j)$ and $B_k = B(x_k, r_k)$, we obtain

$$|x - x_k| \leq |x - x_j| + |x_j - y| + |y - x_k| < r_j + r_j + r_k \leq 3r_k,$$

where in the final step we have used that $r_j \leq r_k$ by the ordering of the balls in $\mathcal{B}$. Thus, we have $x \in 3B_k$ and the proof is complete. □

We are ready to give the proof of the weak-type bound for the maximal operator.

Proof of Theorem 9.5. We set $O_\lambda \coloneqq \{x \in \mathbb{R}^n \mid (\mathcal{M}u)(x) > \lambda\}$. By inner regularity of the Lebesgue measure, it is enough to prove

$$|K| \leq \frac{3^n}{\lambda} \|u\|_{\mathrm{L}^1(\mathbb{R}^n)}$$

for every compact $K \subseteq O_\lambda$. Let us fix K. We can cover K by balls B with the property

$$⨍_B |u| \,\mathrm{d}y > \lambda, \tag{9.3}$$

because the definition of the maximal function implies that every $x \in O_\lambda$ is contained in such a ball. By compactness and the Vitali covering lemma, there is a finite collection $\mathcal{B}$ of balls with property (9.3) and a pairwise-disjoint sub-collection $\mathcal{B}'$ such that

$$K \subseteq \bigcup_{B \in \mathcal{B}} B \subseteq \bigcup_{B' \in \mathcal{B}'} 3B'.$$

Thus, we have

$$|K| \leq \sum_{B' \in \mathcal{B}'} |3B'| = \sum_{B' \in \mathcal{B}'} 3^n |B'| \overset{(9.3)}{\leq} \frac{3^n}{\lambda} \sum_{B' \in \mathcal{B}'} \int_{B'} |u| \,\mathrm{d}y \leq \frac{3^n}{\lambda} \int_{\mathbb{R}^n} |u| \,\mathrm{d}y = \frac{3^n}{\lambda} \|u\|_{\mathrm{L}^1(\mathbb{R}^n)},$$

where in the final estimate we have used that the balls B' are pairwise disjoint. □

The norm bound (9.1) on $\mathrm{L}^\infty(\mathbb{R}^n)$ and the weak-type bound on $\mathrm{L}^1(\mathbb{R}^n)$ can be combined to cover all other Lebesgue exponents in between. The strategy of proof is not limited to the maximal operator and you will learn about the underlying interpolation principle in Exercise 9.5.

Theorem 9.8 (Hardy–Littlewood II). *Let* $1 < p < \infty$. *If* $u \in \mathrm{L}^p(\mathbb{R}^n)$, *then* $\mathcal{M}u \in \mathrm{L}^p(\mathbb{R}^n)$ *and*

$$\|\mathcal{M}u\|_{\mathrm{L}^p(\mathbb{R}^d)} \le C\|u\|_{\mathrm{L}^p(\mathbb{R}^d)}, \quad \textit{where} \quad C := 2 \cdot \Big(\frac{p3^n}{p-1}\Big)^{1/p}.$$

Proof. We start with the layer cake formula

$$\|\mathcal{M}u\|^p_{\mathrm{L}^p(\mathbb{R}^n)} = \int_0^\infty p\lambda^{p-1}\big|\{x \in \mathbb{R}^n \mid (\mathcal{M}u)(x) > \lambda\}\big| \,\mathrm{d}\lambda. \tag{9.4}$$

The trick is to estimate the superlevel set of the maximal function by splitting u according to its own superlevel sets $U_\lambda := \{x \in \mathbb{R}^n \mid |u(x)| > \lambda\}$ at comparable height:

$$u = \mathbf{1}_{U_{\lambda/2}}u + \mathbf{1}_{\mathbb{R}^n \setminus U_{\lambda/2}}u =: u_{1,\lambda/2} + u_{\infty,\lambda/2}.$$

Since $u_{\infty,\lambda/2}$ is essentially bounded by $\lambda/2$, we find $\mathcal{M}u \le \mathcal{M}u_{1,\lambda/2} + \lambda/2$ by sublinearity. This leads us to

$$\Big\{x \in \mathbb{R}^n \mid (\mathcal{M}u)(x) > \lambda\Big\} \subseteq \Big\{x \in \mathbb{R}^n \mid (\mathcal{M}u_{1,\lambda/2})(x) > \frac{\lambda}{2}\Big\},$$

whereupon Theorem 9.5 yields

$$\Big|\Big\{x \in \mathbb{R}^n \mid (\mathcal{M}u)(x) > \lambda\Big\}\Big| \le 3^n\Big(\frac{\lambda}{2}\Big)^{-1}\|u_{1,\lambda/2}\|_{\mathrm{L}^1(\mathbb{R}^n)}.$$

(We could show that $\|u_{1,\lambda/2}\|_{\mathrm{L}^1(\mathbb{R}^n)}$ is finite but of course the estimate continues to hold if it is infinite.) Now, we put all information back into (9.4) and are left with

$$\begin{aligned}
\|\mathcal{M}u\|^p_{\mathrm{L}^p(\mathbb{R}^n)} &\le 3^n \int_0^\infty 2p\lambda^{p-2}\|u_{1,\lambda/2}\|_{\mathrm{L}^1(\mathbb{R}^n)} \,\mathrm{d}\lambda \\
&= 3^n \int_0^\infty 2p\lambda^{p-2} \int_{U_{\lambda/2}} |u(y)| \,\mathrm{d}y \,\mathrm{d}\lambda \\
&\overset{\text{Tonelli}}{=} 3^n \int_{\mathbb{R}^n} |u(y)| \int_0^{2|u(y)|} 2p\lambda^{p-2} \,\mathrm{d}\lambda \,\mathrm{d}y \\
&= 3^n \int_{\mathbb{R}^n} |u(y)| \frac{2p(2|u(y)|)^{p-1}}{p-1} \,\mathrm{d}y \\
&= \Big(\frac{p2^p3^n}{p-1}\Big)\|u\|^p_{\mathrm{L}^p(\mathbb{R}^n)}.
\end{aligned}$$

□

9.2. Lebesgue differentiation

We come back to the initial question about almost everywhere convergence of averages. For locally integrable $u\colon \mathbb{R}^n \to \mathbb{C}$, $x \in \mathbb{R}^n$ and $t > 0$ we can bound

$$\Big|⨍_{B(x,t)} u(y) \,\mathrm{d}y - u(x)\Big| = \Big|⨍_{B(x,t)} \big(u(y) - u(x)\big) \,\mathrm{d}y\Big| \le ⨍_{B(x,t)} |u(y) - u(x)| \,\mathrm{d}y,$$

and if u is continuous at x, then the right-hand side can be further controlled by $\sup_{|y-x|<t} |u(y) - u(x)|$, which tends to 0 in the limit as $t \to 0$. Since $\mathrm{C}_c^\infty(\mathbb{R}^n)$ is dense in $\mathrm{L}^1(\mathbb{R}^n)$ by Proposition 3.3 (d), there is hope to extend pointwise convergence of averages to locally integrable functions u. Of course, we can only hope for convergence almost everywhere — averages do not change upon altering u on a nullset. The weak-type bound for the Hardy–Littlewood maximal operator is precisely what is needed to master the appearing two-parameter limits, as the proof of the following theorem will show.

Theorem 9.9 (Lebesgue differentiation theorem)**.** *If $u\colon \mathbb{R}^n \to \mathbb{C}$ is locally integrable, then for a.e. $x \in \mathbb{R}^n$ we have*

$$\lim_{t\to 0} \fint_{B(x,t)} |u(y) - u(x)| \,\mathrm{d}y = 0. \tag{9.5}$$

Proof. The claim is local in the sense that it suffices to prove it for u replaced by $\mathbf{1}_{B(0,k)}u$ for all $k \in \mathbb{N}$. Hence, we will assume without loss of generality that u is integrable. Now, we set

$$u^*(x) := \limsup_{t\to 0} \fint_{B(x,t)} |u(y) - u(x)| \,\mathrm{d}y \qquad (x \in \mathbb{R}^n) \tag{9.6}$$

and $O_\lambda := \{x \in \mathbb{R}^n \mid u^*(x) > \lambda\}$. Our task is to prove $|O_\lambda| = 0$ for every $\lambda > 0$. Indeed, this would imply that for a.e. $x \in \mathbb{R}^n$ we have $u^*(x) = 0$ and thus also (9.5).

To this end, we fix $\lambda > 0$ and let $\varepsilon > 0$. Proposition 3.3 (d) lets us pick $v \in \mathrm{C}_c^\infty(\mathbb{R}^n)$ with $\|u - v\|_{\mathrm{L}^1(\mathbb{R}^n)} \leq \varepsilon$. For $x \in \mathbb{R}^n$ we have

$$u^*(x) \leq \limsup_{t\to 0} \fint_{B(x,t)} \big(|u(y) - v(y)| + |v(y) - v(x)| + |u(x) - v(x)|\big) \,\mathrm{d}y$$

and hence, the definition of the maximal operator and continuity of v lead us to

$$u^*(x) \leq (\mathcal{M}(u - v))(x) + |u(x) - v(x)|. \tag{9.7}$$

If the left-hand side exceeds λ, then at least one of the terms on the right has to exceed $\lambda/2$. This means that

$$O_\lambda \subseteq \Big\{x \in \mathbb{R}^n \mid (\mathcal{M}(u-v))(x) > \frac{\lambda}{2}\Big\} \cup \Big\{x \in \mathbb{R}^n \mid |(u-v)(x)| > \frac{\lambda}{2}\Big\}.$$

The measure of the first set on the right is under control thanks to Theorem 9.5, whereas for the second one we can simply use Markov's inequality from (9.2). Thereby, we obtain

$$|O_\lambda| \leq \frac{3^n}{\lambda/2}\|u - v\|_{\mathrm{L}^1(\mathbb{R}^n)} + \frac{1}{\lambda/2}\|u - v\|_{\mathrm{L}^1(\mathbb{R}^n)} \leq \frac{2(3^n + 1)}{\lambda}\varepsilon$$

and since $\varepsilon > 0$ was arbitrary, $|O_\lambda| = 0$ follows. □

Remark 9.10. (a) Points x at which (9.5) holds, are called *Lebesgue points* of u.

(b) One instructive interpretation of the proof of Lebesgue's differentiation theorem is that u^* in (9.6) is an auxiliary maximal function that is modeled after the particular claim of the theorem. The argument hinges on the bound by the genuine maximal operator and the property $v^* = 0$ for $v \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$, see (9.7).

Corollary 9.11 (Pointwise convergence of smooth averages)**.** *Let $\theta \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$ be such that $\int_{\mathbb{R}^n} \theta \, \mathrm{d}y = 1$. If $u\colon \mathbb{R}^n \to \mathbb{C}$ is locally integrable, then for a.e. $x \in \mathbb{R}^n$ we have*

$$\lim_{t\to 0}(\theta_t * u)(x) = u(x).$$

Proof. With $B(0, R)$ containing the support of θ, it suffices to note that the functions θ_t have integral 1 because this allows us to bound

$$\begin{aligned}|(\theta_t * u)(x) - u(x)| &= \left|\int_{\mathbb{R}^n} \theta_t(x-y)(u(y) - u(x)) \, \mathrm{d}y\right| \\ &\leq \frac{\|\theta\|_\infty}{t^n} \int_{B(x,Rt)} |u(y) - u(x)| \, \mathrm{d}y \\ &= \omega_n R^n \|\theta\|_\infty \fint_{B(x,Rt)} |u(y) - u(x)| \, \mathrm{d}y.\end{aligned}$$

Now, Theorem 9.9 applies. □

In dimension $n = 1$ we can also study one-sided averages $\frac{1}{h}\int_x^{x+h} u \, \mathrm{d}y$, where h can be negative and the integral is oriented as in Theorem 3.12.

Corollary 9.12 (One-sided averages)**.** *If $u\colon \mathbb{R} \to \mathbb{C}$ is locally integrable, then for a.e. $x \in \mathbb{R}$ we have*

$$\lim_{h\to 0} \frac{1}{h} \int_x^{x+h} u(y) \, \mathrm{d}y = u(x).$$

Proof. We can directly apply Theorem 9.9, upon noting that

$$\left|\frac{1}{h}\int_x^{x+h} u(y) \, \mathrm{d}y - u(x)\right| \leq 2 \fint_{B(x,|h|)} |u(y) - u(x)| \, \mathrm{d}y. \qquad \square$$

In conjunction with Theorem 3.12, convergence of one-sided averages has implications on classical differentiability of H^1-functions on the real line. More precisely, if $u \in \mathrm{L}^1((a, b))$ is weakly differentiable with $u' \in \mathrm{L}^1((a, b))$ on a bounded interval (a, b), then, according to said theorem, u has a continuous representative given by

$$u(x) = C + \int_a^x u'(y) \, \mathrm{d}y \qquad (x \in [a, b]).$$

Corollary 9.12 applies to (the extension by 0 to $\mathbb{R}^n$ of) u' and yields the following.

Corollary 9.13 (Absolutely continuous functions)**.** *Let (a, b) be a bounded interval and let $u \in \mathrm{L}^1((a, b))$ be weakly differentiable with $u' \in \mathrm{L}^1((a, b))$. Then the continuous representative of u is differentiable a.e. on (a, b) with*

$$\lim_{h\to 0} \frac{u(x + h) - u(x)}{h} = u'(x).$$

9.3. Quadratic estimates and Carleson's lemma

Once again, let us come back to the quadratic estimates of Proposition 4.16. Given a radial function $\theta \in \mathcal{S}(\mathbb{R}^n)$, we write out the norm on $\mathrm{L}^2(\mathbb{R}^n)$ and change variables ($\sqrt{t}$ to t) in order to bring them into the form

$$\iint_{\mathbb{R}^{n+1}_+} |(\theta_t * u)(x)|^2 \,\frac{\mathrm{d}x \,\mathrm{d}t}{t} \leq C\|u\|^2_{\mathrm{L}^2(\mathbb{R}^n)} \qquad (u \in \mathrm{L}^2(\mathbb{R}^n)). \tag{9.8}$$

Upon initial examination, this may appear to be connected to the smooth averages of u — however, it is, in fact, nearly the opposite since our assumption on θ was the cancellation property $\int_{\mathbb{R}^n} \theta \,\mathrm{d}x = 0$ rather than the averaging property $\int_{\mathbb{R}^n} \theta \,\mathrm{d}x = 1$. In the latter case we have

$$\|\theta_t * u\|_{\mathrm{L}^2(\mathbb{R}^n)} \geq \frac{1}{2}\|u\|_{\mathrm{L}^2(\mathbb{R}^n)}$$

for sufficiently small $t > 0$ by Proposition 3.3 (c) and, consequently, the left-hand side in (9.8) cannot be integrable with respect to the measure $\frac{\mathrm{d}t}{t}$ unless $u = 0$. The question that we want to raise here is the following:

> *Can one replace $\frac{\mathrm{d}x\,\mathrm{d}t}{t}$ in* (9.8) *by integration with respect to some other Borel measure $\mathrm{d}\nu(x, t)$, such that the estimate holds for averaging operators?*

The answer will lead us to the Carleson measures that have been introduced by Lennart Carleson in his famous solution of the corona problem in complex analysis [Car62]. It will be simpler, and also more convenient for the later lectures, to study Carleson measures in a dyadic setting rather than a continuous one. This means that the average scales will be $t = 2^j$ for $j \in \mathbb{Z}$ and instead of arbitrary balls we use a fixed grid of dyadic cubes that we introduce next.

Definition 9.14. *(a) We denote by $\square := \bigcup_{j\in\mathbb{Z}} \square_{2^j}$ the collection of (half-open) dyadic cubes, where*

$$\square_{2^j} := \left\{2^j x + [0, 2^j)^n \mid x \in \mathbb{Z}^n\right\}$$

are called dyadic cubes *of generation 2^j. By construction, the* sidelength *of each $Q \in \square_{2^j}$ is $\ell(Q) = 2^j$.*

(b) *If $t > 0$, then we set $\square_t := \square_{2^j}$ for the unique integer with $2^{j-1} < t \le 2^j$ and call the cubes in $\square_t$ dyadic cubes of generation t.*

For each $t > 0$ the dyadic cubes of generation t form a partition of $\mathbb{R}^n$. Hence, the following definition makes sense.

Definition 9.15. *Let $u \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$.*

(a) *For $t > 0$ we write*

$$(\mathcal{A}_t u)(x) := \fint_Q u(y) \, \mathrm{d}y,$$

where $Q \in \square_t$ is the unique dyadic cube of generation t that contains x. The operator $\mathcal{A}_t$ is called dyadic averaging operator *at scale t.*

(b) *The* dyadic maximal function *of u is given by*

$$(\mathcal{A}u)(x) := \sup_{t>0} (\mathcal{A}_t |u|)(x)$$

and $\mathcal{A}$ is called dyadic maximal operator.

The dyadic maximal operator is pointwisely controlled by the Hardy–Littlewood maximal operator and many further properties follow from the previous section as we shall see next.

Proposition 9.16. *For all $u \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n)$ and all $x \in \mathbb{R}^n$ we have*

$$(\mathcal{A}u)(x) \le \omega_n n^{n/2} (\mathcal{M}u)(x).$$

Moreover, we have $(\mathcal{A}_t u)(x) \to u(x)$ for a.e. $x \in \mathbb{R}^n$ in the limit as $t \to 0$.

Proof. Let $x \in \mathbb{R}^n$ and $t > 0$. For the unique cube Q used in the definition of $(\mathcal{A}_t|u|)(x)$ we have $Q \subseteq B(x, \sqrt{n}\ell(Q))$ and consequently

$$(\mathcal{A}_t|u|)(x) \le \fint_Q |u(y)| \, \mathrm{d}y \le \omega_n n^{n/2} \fint_{B(x,\sqrt{n}\ell(Q))} |u(y)| \, \mathrm{d}y \le \omega_n n^{n/2} (\mathcal{M}u)(x).$$

Taking the supremum in t gives the required bound. Likewise, for the almost everywhere convergence we bound

$$|(\mathcal{A}_t u)(x) - u(x)| \le \fint_Q |u(y) - u(x)| \, \mathrm{d}y \le \omega_n n^{n/2} \fint_{B(x,\sqrt{n}\ell(Q))} |u(y) - u(x)| \, \mathrm{d}y$$

and apply Lebesgue's differentiation theorem. □

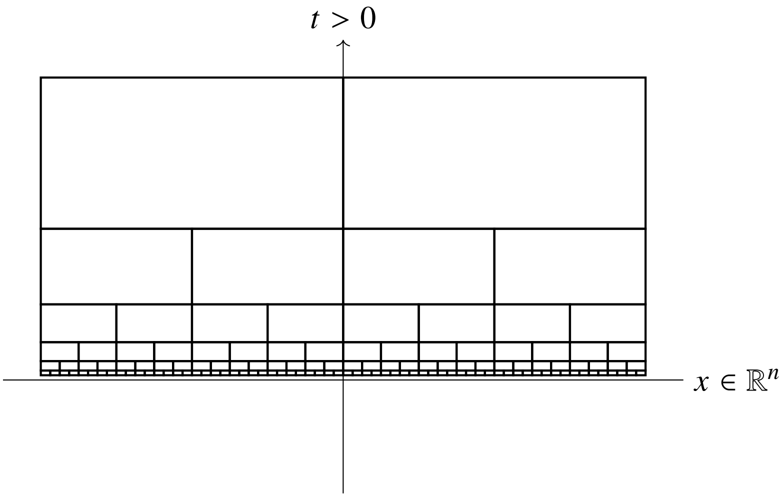

Figure 9.1.: Schematic representation of the decomposition of $\mathbb{R}^{n+1}_+$ into the Whitney boxes (not cubes!) $Q \times (2^{j-1}, 2^j]$ with $Q \in \square_{2^j}$, $j \in \mathbb{Z}$.

Let us come back to our motivational question, but in the dyadic setting, and suppose that ν was a Borel measure on $\mathbb{R}^{n+1}_+$ such that for some constant C we had the estimate

$$\iint_{\mathbb{R}^{n+1}_+} |\mathcal{A}_t u(x)|^2 \, \mathrm{d}\nu(x,t) \leq C \|u\|^2_{\mathrm{L}^2(\mathbb{R}^n)} \qquad (u \in \mathrm{L}^2(\mathbb{R}^n)). \tag{9.9}$$

Along the way, let us note that $(x, t) \mapsto \mathcal{A}_t u(x)$ is constant on a partition of $\mathbb{R}^{n+1}_+$ into so-called *Whitney boxes*, see Figure 9.1. In particular, it is ν-measurable. Given an arbitrary dyadic cube $Q \in \square_{2^j}$, we test the hypothetical estimate (9.9) with $u = \mathbf{1}_Q$ and claim

$$(\mathcal{A}_t \mathbf{1}_Q)(x) = 1 \quad \text{for all} \quad (x, t) \in Q \times (0, \ell(Q)]. \tag{9.10}$$

Indeed, let $x \in Q$, $t \leq \ell(Q) =: 2^j$ and Q' be the unique dyadic cube of generation t that contains x. Since $t \leq 2^j$, we have $Q' \subseteq Q$ and thus $(\mathcal{A}_t \mathbf{1}_Q)(x) = ⨍_{Q'} \mathbf{1}_Q \, \mathrm{d}y = 1$ as claimed. Using (9.10) in (9.9) yields

$$\nu\big(Q \times (0, \ell(Q)]\big) \leq C|Q|.$$

This means that (9.9) can only hold for ν in the following class of Borel measures.

Definition 9.17. *A Borel measure ν on $\mathbb{R}^{n+1}_+$ is called* Carleson measure *if it satisfies*

$$\|\nu\|_C := \sup_{Q \in \square} \frac{\nu\big(Q \times (0, \ell(Q)]\big)}{|Q|} < \infty.$$

For $Q \in \square$, the set $Q \times (0, \ell(Q)]$ is called Carleson box *over Q.*

Example 9.18. We consider the cone $\Gamma := \{(x,t) \in \mathbb{R}^{n+1}_+ \mid |x| < t\}$ in $\mathbb{R}^{n+1}_+$ with tip at the origin in $\mathbb{R}^n$. Then $\mathrm{d}\nu(x,t) := \mathbf{1}_\Gamma(x,t)\, \frac{\mathrm{d}x\,\mathrm{d}t}{t}$ defines a Carleson measure. To get a feel for the notation, we leave the calculation for you as Exercise 9.1.

It turns out that the Carleson condition is not only necessary but also sufficient for (9.9). We learned the following dyadic proof from Andrew Morris [Mor12].

Theorem 9.19 (Dyadic Carleson's lemma). *There is a constant C depending only on dimension such that if ν is a Carleson measure, then*

$$\iint_{\mathbb{R}^{n+1}_+} |\mathcal{A}_t u(x)|^2 \,\mathrm{d}\nu(x,t) \leq C\|\nu\|_C \|u\|^2_{\mathrm{L}^2(\mathbb{R}^n)} \qquad (u \in \mathrm{L}^2(\mathbb{R}^n)).$$

Proof. For $j \in \mathbb{Z}$ we let $\square_{2^j} = (Q_j^k)_k$ be an enumeration of the dyadic cubes of generation 2^j and write $\mathbb{R}^{n+1}_+$ as the union of the pairwise disjoint boxes $Q_j^k \times (2^{j-1}, 2^j]$, see Figure 9.1. Note that for all $(x,t) \in Q_j^k \times (2^{j-1}, 2^j]$ the unique dyadic cube used to define $\mathcal{A}_t u(x)$ is precisely Q_j^k. This allows us to write

$$\begin{aligned}\iint_{\mathbb{R}^{n+1}_+} |\mathcal{A}_t u(x)|^2 \,\mathrm{d}\nu(x,t) &= \sum_{j=-\infty}^{\infty} \sum_k \bigg| \fint_{Q_j^k} u \,\mathrm{d}y \bigg|^2 \nu\big(Q_j^k \times (2^{j-1}, 2^j]\big) \\ &=: \sum_{j=-\infty}^{\infty} \sum_k |u_j^k|^2 \nu_j^k,\end{aligned} \tag{9.11}$$

where we have introduced the numbers

$$\nu_j^k := \nu\big(Q_j^k \times (2^{j-1}, 2^j]\big) \quad \text{and} \quad u_j^k := \fint_{Q_j^k} u \,\mathrm{d}y.$$

Before we can use the Carleson condition on ν without creating too much uncontrollable overlap in the cubes Q_j^k, we need to arrange the boxes $Q_j^k \times (2^{j-1}, 2^j]$ in groups that cover suitable Carleson boxes. We shall do this according to the size of $|u_j^k|$ as follows.

Given a fixed threshold $\lambda > 0$, it may or may not happen that $|u_j^k| > \lambda$. If it happens, then by definition of the dyadic maximal function and Proposition 9.16 we have

$$Q_j^k \subseteq \Big\{x \in \mathbb{R}^n \mid (\mathcal{A}u)(x) > \lambda\Big\} \subseteq \Big\{x \in \mathbb{R}^n \mid (\mathcal{M}u)(x) > \frac{\lambda}{C}\Big\}, \tag{9.12}$$

where $C = \omega_n n^{n/2}$. Now, $u \in \mathrm{L}^2(\mathbb{R}^n)$ implies $\mathcal{M}u \in \mathrm{L}^2(\mathbb{R}^n)$ by Theorem 9.8 and hence the set on the right-hand side has finite measure (by Markov's inequality applied to $(\mathcal{M}u)^2$). Consequently, there is an upper bound on the generation 2^j of cubes with $|u_j^k| > \lambda$. It follows, that each such cube Q_j^k is contained in a maximal dyadic cube with said property, where 'maximal' means 'maximal sidelength'. We let $(Q_m(\lambda))_m$ be an enumeration of the maximal dyadic cubes with the property $|u_j^k| > \lambda$. Maximal dyadic cubes are pairwise disjoint because any two dyadic cubes are either disjoint or one contains the other.

Now, we go back to (9.11) and, similar to the proof of the layer cake formula, we write

$$\begin{aligned}\iint_{\mathbb{R}^{n+1}_+} |\mathcal{A}_t u(x)|^2 \,\mathrm{d}\nu(x,t) &= \sum_{j=-\infty}^{\infty} \sum_k |u_j^k|^2 \nu_j^k \\ &= \sum_{j=-\infty}^{\infty} \sum_k \int_0^{|u_j^k|} 2\lambda \nu_j^k \,\mathrm{d}\lambda \\ &= \int_0^\infty 2\lambda \sum_{j=-\infty}^{\infty} \sum_k \mathbf{1}_{(0,|u_j^k|)}(\lambda) \nu_j^k \,\mathrm{d}\lambda,\end{aligned}$$

where the last step follows by monotone convergence. For fixed λ we are summing over all dyadic cubes Q_j^k with $|u_j^k| > \lambda$. We rearrange the sum according to the maximal cube to which Q_j^k belongs and recall $\nu_j^k = \nu(Q_j^k \times (\ell(Q_j^k)/2, \ell(Q_j^k)])$ in order to obtain

$$\leq \int_0^\infty 2\lambda \sum_m \sum_{Q \subseteq Q_m(\lambda)} \nu\Big(Q \times (\ell(Q)/2, \ell(Q)]\Big) \,\mathrm{d}\lambda.$$

We have been generous by putting in fact every dyadic sub-cube Q of $Q_m(\lambda)$ into the sum. For fixed $Q_m(\lambda)$, the boxes $Q \times (\ell(Q)/2, \ell(Q)]$ with $Q \subseteq Q_m(\lambda)$ form a partition of the Carleson box $Q_m(\lambda) \times (0, \ell(Q_m(\lambda))]$, compare with Figure 9.1. Thus, we get

$$= \int_0^\infty 2\lambda \sum_m \nu\Big(Q_m(\lambda) \times (0, \ell(Q_m(\lambda))]\Big) \,\mathrm{d}\lambda.$$

At this point, we use that ν is a Carleson measure and that the maximal cubes are pairwise disjoint and contained in a superlevel set of the dyadic maximal function, see (9.12), to conclude

$$\begin{aligned}&\leq \|\nu\|_C \int_0^\infty 2\lambda \sum_m |Q_m(\lambda)| \,\mathrm{d}\lambda \\ &\leq \|\nu\|_C \int_0^\infty 2\lambda \Big|\big\{x \in \mathbb{R}^n \mid (\mathcal{A}u)(x) > \lambda\big\}\Big| \,\mathrm{d}\lambda \\ &= \|\nu\|_C \|\mathcal{A}u\|^2_{\mathrm{L}^2(\mathbb{R}^n)},\end{aligned}$$

where we have used the layer cake formula in the final step. The dyadic maximal operator is bounded on $\mathrm{L}^2(\mathbb{R}^n)$ by Proposition 9.16 and Theorem 9.8, which completes the proof (taking $C = 8\omega_n^2(3n)^n$). □

9.4. Exercises

Exercise 9.1 (Example of a Carleson measure)**.** Check that the measure defined in Example 9.18 is indeed a Carleson measure.

Exercise 9.2 (Local integrability of the maximal function)**.** In dimension $n = 1$ consider the function $u(x) = \frac{\mathbf{1}_{(0,1/2)}(x)}{x \log(x)^2}$. Verify that $u \in \mathrm{L}^1(\mathbb{R})$ while $\mathcal{M}u$ is not even locally integrable.

Exercise 9.3 (Kolmogorov's inequality)**.** The following inequality gives a local bound for the maximal function in Lebesgue spaces with exponent $p \in (0, 1)$. Let $p \in (0, 1)$ and $E \subseteq \mathbb{R}^n$ be measurable with finite measure. Show that there exists a constant $C > 0$ that depends only on p and n such that

$$\int_E |\mathcal{M}u(x)|^p \,\mathrm{d}x \leq C|E|^{1-p}\|u\|^p_{\mathrm{L}^1(\mathbb{R}^n)} \qquad (u \in \mathrm{L}^1(\mathbb{R}^n)).$$

Exercise 9.4 (Rectangular maximal operator)**.** In dimension $n \geq 2$ we define a rectangular maximal function of $u \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n)$ by

$$(\mathcal{R}u)(x) := \sup_{R \ni x} \fint_R |u(y)| \,\mathrm{d}y,$$

where, instead of balls, the supremum is taken over all rectangles $R \subseteq \mathbb{R}^n$ with sides parallel to the coordinate axes that contain x.

(a) Prove that $\mathcal{R}$ does *not* satisfy a weak type estimate as in Theorem 9.5.

Hint: Work in dimension $n = 2$ first and estimate the size of $\mathcal{R}u$ for $u = \mathbf{1}_{[-1,1]^2}$.

(b) Prove that $\mathcal{R}$ is still bounded on $\mathrm{L}^p(\mathbb{R}^n)$ for every $p \in (1, \infty]$.

Exercise 9.5 (An interpolation theorem)**.** The following theorem renders more clearly the interpolation idea in the proof of Theorem 9.8.

Theorem 9.20 (Marcinkiewicz interpolation theorem)**.** *Let $(X, \mu), (Y, \nu)$ be σ-finite measure spaces, let $1 \leq p_0 < p_1 < \infty$ and let T be a sublinear operator that is defined on all simple functions on X. Suppose that there are constants C_j, $j = 0, 1$, such that we have for all simple functions u on X and all $\lambda > 0$ that*

$$\nu\big(\{y \in Y \mid |(Tu)(y)| > \lambda\}\big)^{1/p_j} \leq C_j \frac{\|u\|_{\mathrm{L}^{p_j}(X)}}{\lambda}. \tag{9.13}$$

Let $\theta \in (0, 1)$ and define $p \in (p_0, p_1)$ by $1/p = 1-\theta/p_0 + \theta/p_1$. Then there is a constant C, depending on p_0, p_1, θ, such that we have for all u as above the strong-type bound

$$\|Tu\|_{\mathrm{L}^p(Y)} \leq C \cdot C_0^{1-\theta} C_1^{\theta} \|u\|_{\mathrm{L}^p(X)}.$$

(a) Convince yourself that the strong-type assumption

$$\|Tu\|_{\mathrm{L}^{p_j}(Y)} \leq C_j \|u\|_{\mathrm{L}^{p_j}(X)} \tag{9.14}$$

implies (9.13). This explains the name.

(b) Prove Theorem 9.20!

Hint: Proceed as in the proof of Theorem 9.8 but split u at height $\delta\lambda$, where δ is an additional degree of freedom that has to be chosen appropriately at the end. In the proof of Theorem 9.8 we have simply picked $\delta = 1/2$.

10. Sobolev embeddings

In Lecture 3 you have learned that an H^1-function on the real line is $1/2$-Hölder continuous, whereas H^1-functions on $\mathbb{R}^n$ with $n \geq 3$ do not even have to be locally bounded. In this lecture, we are going to investigate said phenomenon further and prove that in every dimension there are natural spaces of Lebesgue- and Hölder-type into which $H^1(\mathbb{R}^n)$ embeds continuously. The general principle to keep in mind is the following relation between the size of the gradient of a weakly differentiable function $u \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and properties of u itself:

$$\begin{tikzcd} & & \boxed{p < n\text{: higher integrability for } u} \\ \boxed{\nabla u \in L^p(\mathbb{R}^n)} \arrow[urr] \arrow[rr] \arrow[drr] & & \boxed{p = n\text{: critical case}} \\ & & \boxed{p > n\text{: Hölder continuity for } u} \end{tikzcd}$$

Figure 10.1.: A rule of thumb for Sobolev functions. For further background we refer to [Eva10, Ch. 5].

We will not go into full details during this first encounter of such *Sobolev embeddings*, but stick to our Sobolev spaces $H^1(\mathbb{R}^n)$. They already cover all three cases nonetheless: Hölder continuity when $n = 1$, criticality (a notion, to be explained in the lecture) when $n = 2$, and higher integrability when $n \geq 3$.

Notation 10.1. Given $p \in [1, \infty]$, its Hölder conjugate exponent $p' \in [1, \infty]$ is defined by $1/p + 1/p' = 1$. Integration with respect to the surface measure on the unit sphere in $\mathbb{R}^n$ will be denoted by $d\sigma$ and σ_{n-1} is the surface measure of the unit sphere.

10.1. Riesz potentials

Let us start by recalling that on the real line you have proved the following result in Exercise 3.1:

Lemma 10.2 (Sobolev embedding, $n = 1$)**.** *Every $u \in \mathrm{H}^1(\mathbb{R})$ has a Hölder continuous representative that satisfies*

$$\sup_{x \neq y} \frac{|u(x) - u(y)|}{|x - y|^{1/2}} \leq \|u'\|_{\mathrm{L}^2(\mathbb{R})}.$$

The embedding was an immediate consequence of the fundamental theorem of calculus — or to be more precise — a representation of u as an integral of its derivative. Unlike the fundamental theorem itself, representations of this type are also available in higher dimensions:

Lemma 10.3. *Let $n \geq 2$. For all $u \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$ we have*

$$u(x) = \frac{1}{\sigma_{n-1}} \int_{\mathbb{R}^n} \nabla u(y) \cdot \frac{x - y}{|x - y|^n} \, \mathrm{d}y \qquad (x \in \mathbb{R}^n).$$

Proof. For $x \in \mathbb{R}^n$ and $e \in \partial B(0, 1)$ we write

$$u(x) = -\int_0^\infty \frac{\mathrm{d}}{\mathrm{d}t} u(x + te) \, \mathrm{d}t = -\int_0^\infty (\nabla u)(x + te) \cdot e \, \mathrm{d}t. \tag{10.1}$$

We average over $e \in \partial B(0, 1)$ to obtain

$$\begin{aligned} u(x) \quad &= \quad -\frac{1}{\sigma_{n-1}} \int_{\partial B(0,1)} \int_0^\infty (\nabla u)(x + te) \cdot e \, \mathrm{d}t \, \mathrm{d}\sigma(e) \\ &\overset{\text{Fubini}}{=} \quad -\frac{1}{\sigma_{n-1}} \int_0^\infty \int_{\partial B(0,1)} (\nabla u)(x + te) \cdot \frac{te}{t^n} \, t^{n-1} \mathrm{d}\sigma(e) \, \mathrm{d}t \end{aligned}$$

and changing from polar to Cartesian coordinates via $z = te$ yields the claim

$$\begin{aligned} &= \quad -\frac{1}{\sigma_{n-1}} \int_{\mathbb{R}^n} (\nabla u)(x + z) \cdot \frac{z}{|z|^n} \, \mathrm{d}z \\ &\overset{y = x + z}{=} \quad \frac{1}{\sigma_{n-1}} \int_{\mathbb{R}^n} \nabla u(y) \cdot \frac{x - y}{|x - y|^n} \, \mathrm{d}y. \end{aligned}$$

□

For $u \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$ the previous lemma gives a pointwise bound

$$|u(x)| \leq \frac{1}{\sigma_{n-1}} \int_{\mathbb{R}^n} \frac{|\nabla u(y)|}{|x - y|^{n-1}} \, \mathrm{d}y \qquad (x \in \mathbb{R}^n). \tag{10.2}$$

We interpret this bound as saying that u is controlled by an integral operator $\mathcal{I}$ acting on $|\nabla u|$ and, consequently, Sobolev embeddings follow from mapping properties of $\mathcal{I}$ in Lebesgue spaces.

Definition 10.4. *For measurable* $f : \mathbb{R}^n \to \mathbb{C}$ *we define the* Riesz potential

$$(\mathcal{I} f)(x) := \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-1}} \, \mathrm{d}y$$

at all $x \in \mathbb{R}^n$ *for which the integral exists.*

Although we are only interested in embeddings for $\mathrm{H}^1(\mathbb{R}^n)$, it will turn out advantageous to examine $\mathcal{I}$ on Lebesgue spaces other than $\mathrm{L}^2(\mathbb{R}^n)$. A uniform bound of the form

$$\|\mathcal{I} f\|_{\mathrm{L}^q(\mathbb{R}^n)} \lesssim \|f\|_{\mathrm{L}^p(\mathbb{R}^n)} \qquad (f \in \mathrm{L}^p(\mathbb{R}^n)) \tag{10.3}$$

can only be true for specific pairs of exponents $p, q \in [1, \infty]$. Indeed, pick any non-negative $\phi \in \mathrm{C}_c^\infty(\mathbb{R}^n) \setminus \{0\}$ and consider the dilations $\delta_t \phi := \phi(t\,\cdot)$ for $t > 0$. By the transformation rule we can check that $\mathcal{I}(\delta_t \phi) = t^{-1} \delta_t(\mathcal{I}\phi)$ and hence the hypothetical bound (10.3) yields

$$t^{-1-n/q} \|\mathcal{I}\phi\|_{\mathrm{L}^q(\mathbb{R}^n)} = \|\mathcal{I}(\delta_t\phi)\|_{\mathrm{L}^q(\mathbb{R}^n)} \lesssim \|\delta_t \phi\|_{\mathrm{L}^p(\mathbb{R}^n)} = t^{-n/p} \|\phi\|_{\mathrm{L}^p(\mathbb{R}^n)}.$$

This bound can only hold for every $t > 0$ if the power of t on the left and right is the same, that is, if $1/q = 1/p - 1/n$ and hence also $1 \leq p \leq n$. We give a name to the exponent q related to p in these considerations.

Definition 10.5. *If* $p \in [1, n)$*, we call the number*

$$p^* := \frac{np}{n-p} \quad \textit{satisfying} \quad \frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}$$

the Sobolev conjugate *of* p*. In the critical case* $p = n$ *we set* $p^* := \infty$*.*

In controlling $\mathcal{I} f$, we will distinguish between the singular (or 'diagonal') part for y close to x and the remaining regular (or 'off-diagonal') part. The maximal operator helps us controlling the diagonal part as follows.

Lemma 10.6. *Let* $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n)$*. For any* $x \in \mathbb{R}^n$ *and* $r > 0$ *we have*

$$\int_{B(x,r)} \frac{|f(y)|}{|x-y|^{n-1}} \, \mathrm{d}y \leq 2^n \omega_n r (\mathcal{M} f)(x).$$

Proof. We split $B(x,r) \setminus \{0\} = \bigcup_{j=0}^\infty B_j \setminus B_{j+1}$, where $B_j := B(x, 2^{-j} r)$, and use $|x - y| \geq 2^{-j}r/2$ for $y \in B_j \setminus B_{j+1}$ to estimate

$$\begin{aligned}
\int_{B(x,r)} \frac{|f(y)|}{|x-y|^{n-1}} \, \mathrm{d}y &= \sum_{j=0}^\infty \int_{B_j \setminus B_{j+1}} \frac{|f(y)|}{|x-y|^{n-1}} \, \mathrm{d}y \\
&\leq \sum_{j=0}^\infty \frac{2^{n-1}}{(2^{-j}r)^{n-1}} \int_{B_j \setminus B_{j+1}} |f(y)| \, \mathrm{d}y \\
&\leq 2^{n-1} \omega_n r \sum_{j=0}^\infty 2^{-j} ⨍_{B_j} |f(y)| \, \mathrm{d}y.
\end{aligned}$$

Since each of the average integrals can be estimated by $(\mathcal{M}f)(x)$, we obtain the claim

$$\begin{aligned}&\le 2^{n-1}\omega_n r\sum_{j=0}^{\infty}2^{-j}(\mathcal{M}f)(x)\\&=2^n\omega_n r(\mathcal{M}f)(x).\end{aligned}\qquad\square$$

Lemma 10.6 is the key ingredient for the following theorem that confirms that the mapping properties for $\mathcal{I}$ suggested below (10.3) are indeed true, except for the endpoint cases $p = 1, n$, in which the bound fails in dimension $n \geq 2$, see Exercise 10.3.

Theorem 10.7 (Hardy–Littlewood–Sobolev inequality)**.** *Let $p \in (1, n)$ and $f \in \mathrm{L}^p(\mathbb{R}^n)$. The integral $(\mathcal{I}f)(x)$ converges absolutely for a.e. $x \in \mathbb{R}^n$ and there is a constant $C = C(n, p)$ such that*

$$\|\mathcal{I}f\|_{\mathrm{L}^{p^*}(\mathbb{R}^n)} \leq C\|f\|_{\mathrm{L}^p(\mathbb{R}^n)}.$$

Proof. Since $|\mathcal{I}(f)| \leq \mathcal{I}|f|$, it suffices to prove the estimate when f is a non-negative function. Moreover, there is nothing to prove if $f = 0$. Henceforth, we assume $f > 0$ on a set of positive measure and note for later that this implies $(\mathcal{M}f)(x) > 0$ for every $x \in \mathbb{R}^n$.

We begin by splitting off the diagonal part that is under control by Lemma 10.6:

$$(\mathcal{I}f)(x) \leq 2^n\omega_n r(\mathcal{M}f)(x) + \int_{\mathbb{R}^n\setminus B(x,r)} \frac{|f(y)|}{|x-y|^{n-1}}\,\mathrm{d}y, \tag{10.4}$$

where the parameter $r > 0$ needs to be chosen wisely in the following. For the off-diagonal part we use Hölder's inequality and calculate

$$\begin{aligned}\int_{\mathbb{R}^n\setminus B(x,r)} \frac{|f(y)|}{|x-y|^{n-1}}\,\mathrm{d}y &\leq \Big(\int_{\mathbb{R}^n\setminus B(x,r)} |f(y)|^p\,\mathrm{d}y\Big)^{1/p}\Big(\int_{\mathbb{R}^n\setminus B(x,r)} |x-y|^{-(n-1)p'}\,\mathrm{d}y\Big)^{1/p'}\\&\leq \|f\|_{\mathrm{L}^p(\mathbb{R}^n)}\Big(\sigma_{n-1}\int_r^{\infty} t^{n-(n-1)p/(p-1)}\,\frac{\mathrm{d}t}{t}\Big)^{(p-1)/p},\end{aligned}$$

where due to $n - (n-1)\frac{p}{p-1} = \frac{p-n}{p-1} < 0$ the integral in t is equal to $\frac{p-1}{n-p}r^{\frac{p-n}{p-1}}$. Going back to (10.4), we have shown the bound

$$(\mathcal{I}f)(x) \lesssim r(\mathcal{M}f)(x) + r^{1-n/p}\|f\|_{\mathrm{L}^p(\mathbb{R}^n)}. \tag{10.5}$$

In order to turn the additive bound into a multiplicative one, we pick r such that the two terms on the right coincide, that is, $r = \|f\|_{\mathrm{L}^p(\mathbb{R}^n)}^{p/n}(\mathcal{M}f)(x)^{-p/n}$. With this choice (10.5) becomes

$$(\mathcal{I}f)(x) \lesssim \|f\|_{\mathrm{L}^p(\mathbb{R}^n)}^{p/n}(\mathcal{M}f)(x)^{1-p/n} = \|f\|_{\mathrm{L}^p(\mathbb{R}^n)}^{p/n}(\mathcal{M}f)(x)^{p/p^*}$$

and integrating the p^*-th powers of both sides with respect to x leads to the desired bound

$$\left(\int_{\mathbb{R}^n} |(\mathcal{I}f)(x)|^{p^*} \, \mathrm{d}x\right)^{1/p^*} \lesssim \|f\|_{\mathrm{L}^p(\mathbb{R}^n)}^{p/n} \left(\int_{\mathbb{R}^n} |(\mathcal{M}f)(x)|^{p} \, \mathrm{d}x\right)^{(n-p)/np}$$
$$\overset{\text{Thm. 9.8}}{\lesssim} \|f\|_{\mathrm{L}^p(\mathbb{R}^n)}^{p/n} \|f\|_{\mathrm{L}^p(\mathbb{R}^n)}^{1-p/n} = \|f\|_{\mathrm{L}^p(\mathbb{R}^n)}. \qquad \square$$

As a corollary, we can extend the potential type representation in Lemma 10.3 to general functions in $\mathrm{H}^1(\mathbb{R}^n)$.

Corollary 10.8. *Let $n \geq 3$ and $u \in \mathrm{H}^1(\mathbb{R}^n)$. For a.e. $x \in \mathbb{R}^n$ we have*

$$u(x) = \frac{1}{\sigma_{n-1}} \int_{\mathbb{R}^n} \nabla u(y) \cdot \frac{x-y}{|x-y|^n} \, \mathrm{d}y.$$

Proof. By Lemma 10.3 the representation holds for $u \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$ and $\mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$ is dense in $\mathrm{H}^1(\mathbb{R}^n)$ by Proposition 3.18. In order to see that the representation persists through approximation in $\mathrm{H}^1(\mathbb{R}^n)$, it suffices to note two things: First, and thanks to the Hardy–Littlewood–Sobolev inequality for $n \geq 3$, the right-hand side above is an absolutely convergent integral for a.e. $x \in \mathbb{R}^n$ that defines a bounded linear operator $\mathrm{H}^1(\mathbb{R}^n) \to \mathrm{L}^{2^*}(\mathbb{R}^n)$ and second, convergence in Lebesgue spaces implies almost everywhere convergence of a subsequence. $\square$

Taking absolute values in the previous corollary immediately gives the following:

Corollary 10.9. *Let $n \geq 3$ and $u \in \mathrm{H}^1(\mathbb{R}^n)$. For a.e. $x \in \mathbb{R}^n$ we have*

$$|u(x)| \leq \frac{1}{\sigma_{n-1}} (\mathcal{I}|\nabla u|)(x).$$

The connection between the function u and the Riesz potential of its gradient from the previous corollary, combined with the Hardy–Littlewood–Sobolev inequality, culminates in the following Sobolev embedding.

Corollary 10.10 (Sobolev embedding, $n \geq 3$)**.** *Let $n \geq 3$. There exists a constant C such that*

$$\|u\|_{\mathrm{L}^{2^*}(\mathbb{R}^n)} \leq C \|\nabla u\|_{\mathrm{L}^2(\mathbb{R}^n)} \qquad (u \in \mathrm{H}^1(\mathbb{R}^n)).$$

So far, the case $n = 2$ is left open, because the Hardy–Littlewood–Sobolev inequality fails for $p = n$ corresponding to $p^* = \infty$. There cannot be a different strategy since a smooth function on $\mathbb{R}^2 \setminus \{0\}$ that coincides with $\log(\log(1 + 1/|x|))$ for $|x| \leq 1/2$ serves as a counterexample to the embedding $\mathrm{H}^1(\mathbb{R}^2) \subseteq \mathrm{L}^\infty(\mathbb{R}^2)$ by calculations that are entirely analogous to Exercise 3.4. However, all Lebesgue spaces 'in between' work and the proof of the following 'subcritical' Sobolev embeddings will be developed in Exercise 10.4.

Theorem 10.11. *If $n \geq 2$ and $2 \leq q < 2^*$, then there is a constant $C = C(n, q)$ such that*

$$\|u\|_{L^q(\mathbb{R}^n)} \leq C\|u\|_{L^2(\mathbb{R}^n)}^{1+\frac{n}{q}-\frac{n}{2}} \|\nabla u\|_{L^2(\mathbb{R}^n)}^{\frac{n}{2}-\frac{n}{q}} \qquad (u \in H^1(\mathbb{R}^n)).$$

If $n = 1$, then the inequality holds for $2 \leq q \leq \infty$. In particular, $H^1(\mathbb{R}^n) \subseteq L^q(\mathbb{R}^n)$ *for all admissible q.*

10.2. Sobolev–Poincaré inequalities

The proof of Lemma 10.3 suggests that potential-type representations, relating u and its gradient, should hold locally for smooth functions defined on an open set Ω, but we still need a corresponding density result to treat more general functions by approximation. Here it is:

Theorem 10.12 (Meyers–Serrin)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be non-empty and open and $k \in \mathbb{N}$. Then* $C^\infty(\Omega) \cap H^k(\Omega)$ *is dense in* $H^k(\Omega)$.

Proof. We write Ω as the union of open sets (V_j) such that $\overline{V_j} \subseteq \Omega$ for $j \in \mathbb{N}$ and every compact subset of Ω intersects only finitely many of the sets V_j. Then we let (η_j) be a partition of unity subordinate to (V_j), that is,

$$\eta_j \in C_c^\infty(V_j), \quad 0 \leq \eta_j \leq 1 \quad \text{and} \quad \sum_{j=1}^{\infty} \eta_j = 1 \text{ in } \Omega.$$

For convenience, we give one possible construction. As in the proof of Lemma 3.4, it starts from the increasingly ordered sets

$$\Omega_j := \{x \in \Omega \mid |x| < j \text{ and } \operatorname{dist}(x, \partial\Omega) > j^{-1}\}.$$

We define the overlapping annuli $V_1 := \Omega_5$ and $V_j := \Omega_{j+4} \setminus \overline{\Omega_{j+1}}$ for $j \geq 2$. Then we pick $\phi_1 \in C_c^\infty(V_1)$ with $\phi_1 = 1$ on Ω_4 and for $j \geq 2$ we take $\phi_j \in C_c^\infty(V_j)$ with $\phi_j = 1$ on $\Omega_{j+3} \setminus \Omega_{j+2}$, so that $\eta_j := \phi_j (\sum_{k=1}^\infty \phi_k)^{-1}$ does the job.

The properties of η_j together with Lemma 3.9 ensure that $\eta_j u \in H^k(\Omega)$ and that $\eta_j u$ vanishes outside of a compact subset of $V_j \subseteq \Omega$. Hence, we may identify it with its extension to $\mathbb{R}^n$ by zero and this will allow us to smooth out $\eta_j u$ by mollification.

To this end, let $\theta \in C_c^\infty(\mathbb{R}^n)$ be supported in the unit ball and normalized to $\int_{\mathbb{R}^n} \theta \, dx = 1$. For $t > 0$, let θ_t denote the mollifier associated to θ. Proposition 3.3, the smoothness of θ and the above-mentioned properties of the supports guarantee that $(\eta_j u) * \theta_{t_j} \in C_c^\infty(V_j)$ whenever $t_j > 0$ is small enough. Given $\varepsilon > 0$, Lemma 3.11 and yet another application of Proposition 3.3 allow us to choose t_j even smaller such that

$$\big\|(\eta_j u) * \theta_{t_j} - \eta_j u\big\|_{H^k(\Omega)} \leq \sum_{|\alpha| \leq k} \big\|[\partial^\alpha(\eta_j u)] * \theta_{t_j} - \partial^\alpha(\eta_j u)\big\|_{L^2(\Omega)} < 2^{-j}\varepsilon.$$

Finally, we define

$$u_\varepsilon := \sum_{j=1}^{\infty} (\eta_j u) * \theta_{t_j} \in \mathrm{C}^\infty(\Omega) \cap \mathrm{H}^k(\Omega),$$

where on every compact subset of Ω all but finitely many functions are identically zero, and estimate

$$\|u_\varepsilon - u\|_{\mathrm{H}^k(\Omega)} = \Big\| \sum_{j=1}^{\infty} \big((\eta_j u) * \theta_{t_j} - \eta_j u\big) \Big\|_{\mathrm{H}^k(\Omega)} \leq \sum_{j=1}^{\infty} 2^{-j}\varepsilon = \varepsilon. \qquad \square$$

The potential type estimate in (10.2) has made crucial use of the fact that on every line in $\mathbb{R}^n$ the compactly supported function u vanishes somewhere, see (10.1). Of course, this is no longer true when we work with general smooth functions on an open subset Ω and we need a geometric condition that helps us controlling the 'anchor points' when representing $u(x)$ via the fundamental theorem of calculus on line segments passing through x.

Definition 10.13. *A set $\Omega \subseteq \mathbb{R}^n$ is called* star-shaped with respect to a non-empty subset $E \subseteq \Omega$ *if for all $y \in E$ and $x \in \Omega$ the segment connecting x and y is contained in Ω.*

Example 10.14. Non-empty convex sets are star-shaped with respect to every non-empty subset. Hungry Pac-Man[1] is not convex but star-shaped with respect to a ball.

For our potential-type estimates on domains we recall the notation $(u)_E$ for averages, see Notation 9.1.

Lemma 10.15. *Let $n \geq 2$, let $\Omega \subseteq \mathbb{R}^n$ be open, bounded and star-shaped with respect to a measurable subset $E \subseteq \Omega$ with $|E| > 0$, and let $u \in \mathrm{H}^1(\Omega)$. Then for a.e. $x \in \Omega$,*

$$|u(x) - (u)_E| \leq \frac{\operatorname{diam}(\Omega)^n}{n|E|} \int_\Omega \frac{|\nabla u(y)|}{|x-y|^{n-1}} \,\mathrm{d}y.$$

Proof. We proceed in two steps.

Step 1: We assume $u \in \mathrm{C}^\infty(\Omega)$.

For $x \in \Omega$ and $y \in E$ we can use the fundamental theorem of calculus along the connecting line segment to write

$$u(x) - u(y) = -\int_0^{|x-y|} \frac{\mathrm{d}}{\mathrm{d}t} u(x + t\omega_{x,y}) \,\mathrm{d}t = -\int_0^{|x-y|} (\nabla u)(x + t\omega_{x,y}) \cdot \omega_{x,y} \,\mathrm{d}t,$$

[1] https://de.wikipedia.org/wiki/Pac-Man

where $\omega_{x,y} := \frac{y-x}{|y-x|}$ is a unit vector. We integrate both sides with respect to $y \in E$ in order to obtain

$$|E||u(x) - (u)_E| \leq \int_{\mathbb{R}^n} \mathbf{1}_E(y) \int_0^{|x-y|} \big|(\nabla u)(x + t\omega_{x,y})\big| \,\mathrm{d}t \,\mathrm{d}y.$$

Changing to polar coordinates $y = x + r\nu$ yields

$$\begin{aligned} &= \int_0^\infty \int_{\partial B(0,1)} \int_0^r \mathbf{1}_E(x + r\nu)\big|(\nabla u)(x + t\nu)\big| r^{n-1} \,\mathrm{d}t \,\mathrm{d}\sigma(\nu) \,\mathrm{d}r \\ &= \int_0^\infty \int_{\partial B(0,1)} \big|(\nabla u)(x + t\nu)\big| \bigg(\int_t^\infty r^{n-1} \mathbf{1}_E(x + r\nu) \,\mathrm{d}r \bigg) \,\mathrm{d}\sigma(\nu) \,\mathrm{d}t. \end{aligned}$$

The result will follow by regarding t as the radial variable for the polar coordinates instead of r. For this purpose, we observe that $\mathbf{1}_E(x + r\nu) = 0$ whenever $r > \mathrm{diam}(\Omega)$ and that for all $0 < t < r$ we have $x + r\nu \in E \Rightarrow x + t\nu \in \Omega$, since Ω is star-shaped with respect to E. Thus, $\mathbf{1}_E(x + r\nu) = \mathbf{1}_E(x + r\nu)\mathbf{1}_\Omega(x + t\nu)$ and we can continue by

$$\leq \int_0^\infty \int_{\partial B(0,1)} \big|(\mathbf{1}_\Omega \nabla u)(x + t\nu)\big| \bigg(\int_0^{\mathrm{diam}(\Omega)} r^{n-1} \,\mathrm{d}r \bigg) \,\mathrm{d}\sigma(\nu) \,\mathrm{d}t.$$

Returning to Cartesian coordinates via $y = x + t\nu$, we are left with the required estimate

$$= \frac{\mathrm{diam}(\Omega)^n}{n} \int_\Omega \frac{|\nabla u(y)|}{|x - y|^{n-1}} \,\mathrm{d}y.$$

Step 2: Extension by density.

Since $\mathrm{C}^\infty(\Omega) \cap \mathrm{H}^1(\Omega)$ is dense in $\mathrm{H}^1(\Omega)$ by Theorem 10.12 and as Ω is bounded, the general case can be obtained by approximation as in the proof of Corollary 10.8, even when $n = 2$. Indeed, since on the bounded set Ω we have a continuous inclusion $\mathrm{L}^2(\Omega) \subseteq \mathrm{L}^p(\Omega)$ for any $p \in (1, 2)$, the integral operator $\mathcal{I}(\mathbf{1}_\Omega \cdot)$ is bounded $\mathrm{H}^1(\Omega) \to \mathrm{L}^{p^*}(\mathbb{R}^n)$ thanks to Theorem 10.7. □

Remark 10.16. Of course $p = 2$ works in the final step if we are in dimension $n \geq 3$ but the result in dimension $n = 2$ has to use the Hardy–Littlewood–Sobolev inequality with $p \neq 2$.

We use the potential-type representation from the previous lemma to prove the following inequality.

Theorem 10.17 (Sobolev–Poincaré inequality)**.** *Let $n \geq 1$ and $\Omega \subseteq \mathbb{R}^n$ be non-empty, open, bounded and star-shaped with respect to some measurable subset $E \subseteq \Omega$ with $|E| > 0$. Furthermore, let*

- $1 \leq q \leq 2^*$ *if $n \geq 3$,*

- $1 \leq q < \infty$ *if* $n = 2$,
- $1 \leq q \leq \infty$ *if* $n = 1$.

Then there exists $C = C(n, q) > 0$ *such that for all* $u \in \mathrm{H}^1(\Omega)$ *we have*

$$\|u - (u)_E\|_{\mathrm{L}^q(\Omega)} \leq C \frac{\operatorname{diam}(\Omega)^n}{|E|} |\Omega|^{1/q - (1/2 - 1/n)} \|\nabla u\|_{\mathrm{L}^2(\Omega)}.$$

Proof. Observe that once we have the claim for some q_0, then we get the claim for all $q \in [1, q_0)$ from Hölder's inequality:

$$\|u - (u)_E\|_{\mathrm{L}^q(\Omega)} \leq |\Omega|^{1/q - 1/q_0} \|u - (u)_E\|_{\mathrm{L}^{q_0}(\Omega)} \lesssim \frac{\operatorname{diam}(\Omega)^n}{|E|} |\Omega|^{1/q - (1/2 - 1/n)} \|\nabla u\|_{\mathrm{L}^2(\Omega)}.$$

Moreover, by definition of the Riesz potential, Lemma 10.15 can be read as

$$|u(x) - (u)_E| \leq \frac{\operatorname{diam}(\Omega)^n}{n|E|} \big(\mathcal{I}(\mathbf{1}_\Omega |\nabla u|)\big)(x) \qquad (\text{a.e. } x \in \Omega). \tag{10.6}$$

Case 1: We assume $n \geq 3$ *and* $q = 2^*$.

Using (10.6) together with Theorem 10.7, we find

$$\|u - (u)_E\|_{\mathrm{L}^q(\Omega)} \leq \frac{\operatorname{diam}(\Omega)^n}{n|E|} \|\mathcal{I}(\mathbf{1}_\Omega |\nabla u|)\|_{\mathrm{L}^q(\Omega)} \lesssim \frac{\operatorname{diam}(\Omega)^n}{|E|} \|\nabla u\|_{\mathrm{L}^2(\Omega)}.$$

Case 2: We assume $n = 2$ *and* $q > 2$.

We define $q_* := 2q/(2+q)$ and note that $q_* \in (1, 2)$ as well as $(q_*)^* = q$. Moreover, we have $\mathbf{1}_\Omega |\nabla u| \in \mathrm{L}^{q_*}(\mathbb{R}^n)$ since Ω is bounded. Thus, another application of (10.6) and Theorem 10.7, followed by Hölder's inequality, gives

$$\begin{aligned}
\|u - (u)_E\|_{\mathrm{L}^q(\Omega)} &\leq \frac{\operatorname{diam}(\Omega)^n}{n|E|} \|\mathcal{I}(\mathbf{1}_\Omega |\nabla u|)\|_{\mathrm{L}^q(\Omega)} \\
&\lesssim \frac{\operatorname{diam}(\Omega)^n}{|E|} \|\nabla u\|_{\mathrm{L}^{q_*}(\Omega)} \\
&\leq \frac{\operatorname{diam}(\Omega)^n}{|E|} |\Omega|^{1/q_* - 1/2} \|\nabla u\|_{\mathrm{L}^2(\Omega)}.
\end{aligned}$$

The claim follows since $1/q_* - 1/2 = 1/q$.

Case 3: We assume $n = 1$ *and* $q = \infty$.

In this case we can use the fundamental theorem of calculus directly and none of the material from this section is actually needed. We leave the (hopefully instructive) comparison of methods as Exercise 10.2. □

Example 10.18. A typical example for Theorem 10.17 is when $\Omega = B(x,r)$ is a ball and $|E| \geq \gamma |B(x,r)|$ for some $\gamma \in (0,1)$. For instance, E could be a ball with radius comparable to that of $B(x,r)$. In that case, the Sobolev–Poincaré inequality becomes

$$\|u - (u)_E\|_{\mathrm{L}^q(B(x,r))} \leq C r^{1+n(\frac{1}{q}-\frac{1}{2})} \|\nabla u\|_{\mathrm{L}^2(B(x,r))}.$$

Eventually, the Sobolev–Poincaré inequality implies a Sobolev embedding on suitable open subsets of $\mathbb{R}^n$.

Corollary 10.19. *In the setting of Theorem 10.17 there exists a constant $C \geq 0$ such that*

$$\|u\|_{\mathrm{L}^q(\Omega)} \leq C\big(\|\nabla u\|_{\mathrm{L}^2(\Omega)} + \|u\|_{\mathrm{L}^1(E)}\big) \qquad (u \in \mathrm{H}^1(\Omega)).$$

Proof. Write $u = (u - (u)_E) + (u)_E$ and use that $\|(u)_E\|_{\mathrm{L}^q(\Omega)} \leq \frac{|\Omega|^{1/q}}{|E|}\|u\|_{\mathrm{L}^1(E)}$. □

10.3. Further applications of the Meyers–Serrin theorem

The Meyers–Serrin theorem opens the door to proving further, more involved properties of weak derivatives by approximation. To showcase the general strategy, we prove a version of the chain rule. It will be somewhat more convenient to only work with *real-valued* functions.

Theorem 10.20 (Chain rule for H^1)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be a non-empty open set and let $u \in \mathrm{H}^1(\Omega)$ be real-valued. Suppose that $\phi\colon \mathbb{R} \to \mathbb{R}$ is continuously differentiable with bounded derivative and $\phi(0) = 0$. Then $\phi \circ u \in \mathrm{H}^1(\Omega)$ and in the weak sense*

$$\nabla(\phi \circ u)(x) = \phi'(u(x))\nabla u(x) \qquad (a.e.\ x \in \Omega).$$

Proof. We set $C \coloneqq \|\phi'\|_{\mathrm{L}^\infty(\mathbb{R})}$ so that $|\phi(t)| \leq C|t|$ by the mean value theorem. Given a real-valued $u \in \mathrm{H}^1(\Omega)$, the bounds on ϕ imply

$$\|\phi \circ u\|_{\mathrm{L}^2(\Omega)} \leq C\|u\|_{\mathrm{L}^2(\Omega)} \quad \text{and} \quad \|(\phi' \circ u)\nabla u\|_{\mathrm{L}^2(\Omega)} \leq C\|\nabla u\|_{\mathrm{L}^2(\Omega)}. \tag{10.7}$$

Now, we take an approximating sequence $(u_j) \subseteq \mathrm{C}^\infty(\Omega) \cap \mathrm{H}^1(\Omega)$ as in the Meyers–Serrin theorem. Up to switching to a subsequence, we can assume that (u_j) tends to u a.e. and by construction we can take the u_j real-valued. Since the chain rule holds for the continuously differentiable functions $\phi \circ u_j$, we conclude from (10.7) that $\phi \circ u_j \in \mathrm{H}^1(\Omega)$. The mean value theorem once again yields

$$\|\phi \circ u_j - \phi \circ u\|_{\mathrm{L}^2(\Omega)} \leq C\|u_j - u\|_{\mathrm{L}^2(\Omega)},$$

which shows $\phi \circ u_j \to \phi \circ u$ in $\mathrm{L}^2(\Omega)$ in the limit as $j \to \infty$. Furthermore, we have

$$\begin{aligned}
&\big\|\nabla(\phi \circ u_j) - (\phi' \circ u)\nabla u\big\|_{\mathrm{L}^2(\Omega)} \\
&\qquad \leq \big\|(\phi' \circ u_j)(\nabla u_j - \nabla u)\big\|_{\mathrm{L}^2(\Omega)} + \big\|(\phi' \circ u_j - \phi' \circ u)\nabla u\big\|_{\mathrm{L}^2(\Omega)} \\
&\qquad \leq C\big\|\nabla u_j - \nabla u\big\|_{\mathrm{L}^2(\Omega)} + \big\|(\phi' \circ u_j - \phi' \circ u)\nabla u\big\|_{\mathrm{L}^2(\Omega)},
\end{aligned}$$

where the first term on the right tends to zero. By dominated convergence, so does the second term since ϕ' is bounded and continuous and (u_j) tends to u a.e. on Ω. In total, we have also shown $\nabla(\phi \circ u_j) \to (\phi' \circ u)\nabla u$ in $\mathrm{L}^2(\Omega)$. Hence, $\phi \circ u_j$ converges in $\mathrm{H}^1(\Omega)$ and the weak gradient of its limit $\phi \circ u$ can be computed by the chain rule. □

10.4. Exercises

Exercise 10.1 (An integral on the real line)**.** Let $u \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R})$ and $x \in \mathbb{R}$. Show that[2]

$$u(x) = \frac{1}{2}\int_{\mathbb{R}} u'(y)\,\mathrm{sgn}(x-y)\,\mathrm{d}y.$$

Then compare with Lemma 10.3.

Exercise 10.2 (Sobolev–Poincaré inequality in $\mathbb{R}$)**.** Use the fundamental theorem of calculus to prove Theorem 10.17 in dimension $n = 1$.

Exercise 10.3 (Riesz potentials at the endpoints)**.** Prove that in dimension $n \geq 2$ the boundedness of the Riesz potential $\mathcal{I} : \mathrm{L}^p(\mathbb{R}^n) \to \mathrm{L}^{p^*}(\mathbb{R}^n)$ fails at the endpoints $p = 1$ and $p = n$. What happens in dimension $n = 1$?

Hint: Remember ‘les intégrales de Bertrand’ $\int_0^{1/\mathrm{e}} \frac{1}{r|\log(r)|^\alpha}\,\mathrm{d}r \ldots$

Exercise 10.4 (Gagliardo–Nirenberg inequality)**.** In this exercise, you are going to establish the *Gagliardo–Nirenberg inequality* stated in Theorem 10.11:

$$\|u\|_{\mathrm{L}^q(\mathbb{R}^n)} \leq C\|u\|_{\mathrm{L}^2(\mathbb{R}^n)}^{1+n/q-n/2}\|\nabla u\|_{\mathrm{L}^2(\mathbb{R}^n)}^{\frac{n}{2}-n/q} \qquad (u \in \mathrm{H}^1(\mathbb{R}^n)), \tag{10.8}$$

whenever q satisfies

- $2 \leq q \leq 2^*$ if $n \geq 3$,
- $2 \leq q < \infty$ if $n = 2$,

[2] As usual, $\mathrm{sgn} := \mathbf{1}_{(0,\infty)} - \mathbf{1}_{(-\infty,0)}$ is the sign function.

- $2 \leq q \leq \infty$ if $n = 1$.

We start with the case $n \geq 3$.

(a) Establish (10.8) as follows: Use Hölder's inequality to control $\|u\|_{L^q(\mathbb{R}^n)}$ by $\|u\|_{L^2(\mathbb{R}^n)}$ and $\|u\|_{L^{2^*}(\mathbb{R}^n)}$ and apply the Sobolev embedding afterwards.

We proceed with the critical case $n = 2$. Let $2 \leq q < \infty$.

(b) Why does the same reasoning as in part (a) not apply when $n = 2$ with $2^* = \infty$, even though it would lead exactly to the required estimate?

(c) Let $u \in C_c^\infty(\mathbb{R}^2)$. Use Lemma 10.15 for $\Omega = E = B(x, r)$ and Lemma 10.6 to establish for any $r \in (0, \infty)$ and $\theta \in (0, 1)$ the bound

$$|u(x)| \leq C\Big(r^{-2\theta/q}\|u\|_{L^q(\mathbb{R}^n)}^\theta (\mathcal{M}u)(x)^{1-\theta} + r(\mathcal{M}\nabla u)(x)\Big).$$

(d) Minimize the inequality derived in (c) with respect to $r > 0$ to establish a multiplicative counterpart of this inequality.

(e) Take the L^q-norm of the inequality derived in (d) and apply Hölder's inequality with exponents p and p' on the right-hand side, where also $p \in (1, \infty)$ needs to be chosen yet.

(f) Get rid of the maximal operator and find unique choices for θ and p from the requirement that $\|\nabla u\|_{L^2(\mathbb{R}^n)}^{1-2/q}$ should appear on the right-hand side in (10.8).

(g) Complete the proof of (10.8).

Finally, we study the case $n = 1$.

(f) Use Corollary 3.14 to prove for any $r > 0$ that

$$|u(x)| \leq (2r)^{-1/2}\|u\|_{L^2(\mathbb{R})} + (2r)^{1/2}\|u'\|_{L^2(\mathbb{R})} \qquad (x \in \mathbb{R}).$$

(g) Establish (10.8) in the case $q = \infty$ by minimizing with respect to r.

(h) Conclude (10.8) for $2 \leq q < \infty$.

Exercise 10.5 (Truncation of Sobolev functions)**.** Let $\Omega \subseteq \mathbb{R}^n$ be a non-empty open set. Prove that $H^1(\Omega) \cap L^\infty(\Omega)$ is dense in $H^1(\Omega)$.

Exercise 10.6 (Dirichlet conditions and the punctured space)**.** In this exercise we come back to a claim from Lecture 3, namely that in dimension $n \geq 2$ one point is 'too small to be seen' by Sobolev functions. More precisely, we ask you to prove that on $\Omega := \mathbb{R}^n \setminus \{0\}$ we have $H^1(\Omega) = H_0^1(\Omega)$.

Hint: Take $u \in H^1(\Omega) \cap L^\infty(\Omega)$ and construct a sequence $(u_j) \subseteq C_c^\infty(\Omega)$ that is bounded in $H^1(\Omega)$ and converges to u in $L^2(\Omega)$. Then, use functional analysis.

11. Off-diagonal behavior

In this lecture, we come back to elliptic operators L in divergence form. To streamline the presentation, we shall put boundary conditions aside and only work in $\mathbb{R}^n$. So far, we have regarded L as an example for various, more general concepts in Hilbert space theory: sesquilinear forms, functional calculus and quadratic estimates. As a matter of fact, all estimates involving L have been of a global nature, taking the norm on $\mathrm{L}^2(\mathbb{R}^n)$, and rarely have we used that we work on a space of functions that may have special properties on certain subregions $E, F \subseteq \mathbb{R}^n$. In this lecture we are interested in finer mapping properties of the resolvent $(1+L)^{-1}$ relative to given E, F:

If u is supported in E, what can be said about the size of $(1+L)^{-1}u$ on F — in particular, if E and F are far away from each other?

Once again, the negative Laplacian $-\Delta$ on $\mathbb{R}^n$ will be our guiding example but, compared to Lecture 4, we work in the 'state variable' x rather than the 'frequency variable' ξ.

Notation 11.1. Throughout the lecture, $L = -\operatorname{div}(A\nabla\,\cdot)$ denotes an elliptic operator in divergence form on $\mathbb{R}^n$ and λ, Λ are the lower and upper bound for its coefficients, see (3.9) and (3.4). We write $\operatorname{dist}(E, F)$ for the (Euclidean) distance of sets $E, F \subseteq \mathbb{R}^n$.

11.1. The Bessel kernel and the negative Laplacian

For starters, we consider $L = -\Delta$ the negative Laplacian on $\mathbb{R}^n$. In this case, explicit computations can be made. It all relies on the following representation of the resolvent as a convolution operator.

Lemma 11.2. *Let $G : \mathbb{R}^n \to (0, \infty)$ be given by*

$$G(x) := \frac{1}{4\pi} \int_0^\infty \mathrm{e}^{-\pi|x|^2/s} \mathrm{e}^{-s/(4\pi)} s^{-n/2} \,\mathrm{d}s.$$

Then $\|G\|_{\mathrm{L}^1(\mathbb{R}^n)} = 1$ and for every $u \in \mathrm{L}^2(\mathbb{R}^n)$ we have that $(1-\Delta)^{-1}u = G * u$.

Proof. First, we compute the norm of G as

$$\int_{\mathbb{R}^n} G(x)\,\mathrm{d}x \overset{\text{Tonelli}}{=} \frac{1}{4\pi}\int_0^\infty \mathrm{e}^{-s/4\pi}\Big(s^{-n/2}\int_{\mathbb{R}^n}\mathrm{e}^{-\pi|x|^2/s}\,\mathrm{d}x\Big)\,\mathrm{d}s$$
$$\overset{x=\sqrt{s}y}{=} \frac{1}{4\pi}\int_0^\infty \mathrm{e}^{-s/4\pi}\cdot 1\,\mathrm{d}s \quad = \quad 1.$$

To show the identity for the resolvent, we introduce $m = 4\pi^2|\cdot|^2$ as in Lecture 4 and recall from Corollary 4.8 and Example 1.16 that $(1-\Delta)^{-1}u = \mathcal{F}^{-1}((1+m)^{-1}\widehat{u})$. Since the Fourier transform turns convolutions into multiplications, we have $G * u = \mathcal{F}^{-1}(\widehat{G}\widehat{u})$ in $\mathrm{L}^2(\mathbb{R}^n)$ and our task is to prove $\widehat{G} = (1+m)^{-1}$. To this end, we let $v \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$ and compute

$$\int_{\mathbb{R}^n}\widehat{G}(\xi)\cdot\overline{v(\xi)}\,\mathrm{d}\xi \overset{\text{Plancherel}}{=} \int_{\mathbb{R}^n} G(x)\cdot\overline{\check{v}(x)}\,\mathrm{d}x$$
$$\overset{\text{Tonelli}}{=} \frac{1}{4\pi}\int_{\mathbb{R}^n}\int_0^\infty\int_{\mathbb{R}^n}\mathrm{e}^{-2\pi\mathrm{i}x\cdot\xi}\mathrm{e}^{-\pi|x|^2/s}\,\mathrm{d}x\,\mathrm{e}^{-s/4\pi}s^{-n/2}\,\mathrm{d}s\,\overline{v(\xi)}\,\mathrm{d}\xi.$$

The integral in x is the Fourier transform of $\mathrm{e}^{-\pi|\cdot|^2/s}$ at ξ. Since $\mathrm{e}^{-\pi|\cdot|^2}$ is a fixed point of the Fourier transform[1], the integral in x is equal to $s^{n/2}\mathrm{e}^{-\pi s|\xi|^2}$, see Proposition 4.2 (c). Now, we can continue by

$$= \frac{1}{4\pi}\int_{\mathbb{R}^n}\int_0^\infty \mathrm{e}^{-s(1/4\pi+\pi|\xi|^2)}\,\mathrm{d}s\,\overline{v(\xi)}\,\mathrm{d}\xi$$
$$= \int_{\mathbb{R}^n}\frac{1}{1+4\pi^2|\xi|^2}\cdot\overline{v(\xi)}\,\mathrm{d}\xi.$$

The fundamental lemma in the calculus of variations implies $\widehat{G} = (1+m)^{-1}$ and the proof is complete. □

The radially symmetric kernel G in the previous lemma is usually called *Bessel kernel*. It enjoys the following exponential estimate.

Lemma 11.3. *There exists $C > 0$ such that*

$$G(x) \le C\mathrm{e}^{-|x|/2} \qquad (|x| \ge 1).$$

Proof. We bound the integrand in the definition of G by

$$\mathrm{e}^{-\pi|x|^2/s}\mathrm{e}^{-s/4\pi} = \mathrm{e}^{-(\pi|x|^2/2s+s/8\pi)}\cdot\big(\mathrm{e}^{-\pi|x|^2/2s}\mathrm{e}^{-s/8\pi}\big) \le \mathrm{e}^{-|x|/2}\big(\mathrm{e}^{-\pi/2s}\mathrm{e}^{-s/8\pi}\big),$$

[1]Most likely, you have seen this in your calculus classes. Here is a recap of the argument. When $n = 1$, the functions $y := \mathrm{e}^{-\pi|\cdot|^2}$ and $\widehat{y}$ both solve the same initial value problem $z'(x) = -2\pi x z(x)$ with $z(0) = 1$. In higher dimensions, write $\mathrm{e}^{-2\pi\mathrm{i}x\cdot\xi}\mathrm{e}^{-\pi|x|^2} = \mathrm{e}^{-2\pi\mathrm{i}x_1\xi_1}\mathrm{e}^{-\pi x_1^2}\cdot\ldots\cdot\mathrm{e}^{-2\pi\mathrm{i}x_n\xi_n}\mathrm{e}^{-\pi x_n^2}$ and use Fubini's theorem.

where we have used the inequality between arithmetic and geometric mean $a+b/2 \geq \sqrt{ab}$ in the exponent of the first factor and the assumption $|x| \geq 1$ for the second one. The claim follows by integrating with respect to $s \in (0, \infty)$. □

In case of the negative Laplacian, our initial question can now easily be answered.

Proposition 11.4 (L^2-off-diagonal estimates for the negative Laplacian)**.** *We can find constants $C, c > 0$ such that for all measurable sets $E, F \subseteq \mathbb{R}^n$ and all $u \in \mathrm{L}^2(\mathbb{R}^n)$ with* $\operatorname{supp}(u) \subseteq E$ *we have*

$$\|(1 - \Delta)^{-1} u\|_{\mathrm{L}^2(F)} \leq C\mathrm{e}^{-c\operatorname{dist}(E,F)} \|u\|_{\mathrm{L}^2(E)}.$$

Proof. Let $u \in \mathrm{L}^2(\mathbb{R}^n)$ with $\operatorname{supp}(u) \subseteq E$. We distinguish two cases.

Case 1: $\operatorname{dist}(E, F) \leq 1$.

In this case we are not claiming any decay and simply use the global resolvent bound, coming from m-accretivity of $-\Delta$:

$$\|(1 - \Delta)^{-1} u\|_{\mathrm{L}^2(F)} \leq \|(1 - \Delta)^{-1} u\|_{\mathrm{L}^2(\mathbb{R}^n)} \leq \|u\|_{\mathrm{L}^2(\mathbb{R}^n)} \leq \mathrm{e} \cdot \mathrm{e}^{-\operatorname{dist}(E,F)} \|u\|_{\mathrm{L}^2(E)}.$$

Case 2: $\operatorname{dist}(E, F) > 1$.

We use the formula in Lemma 11.2. Since $\|G\|_{\mathrm{L}^1(\mathbb{R}^n)} = 1$ and u is supported in E, the Cauchy–Schwarz inequality yields

$$\begin{aligned}
\big|\big((1 - \Delta)^{-1} u\big)(x)\big|^2 &= \left|\int_E G(x - y)u(y) \,\mathrm{d}y\right|^2 \\
&\leq \left(\int_E G(x - y)^{1/2} G(x - y)^{1/2} |u(y)| \,\mathrm{d}y\right)^2 \\
&\leq \int_E G(x - y)|u(y)|^2 \,\mathrm{d}y.
\end{aligned}$$

We integrate over $x \in F$, use Tonelli's theorem and switch to polar coordinates, in order to find

$$\begin{aligned}
\int_F |(1 - \Delta)^{-1} u|^2 \,\mathrm{d}x &\leq \int_E \int_F G(x - y) \,\mathrm{d}x \,|u(y)|^2 \,\mathrm{d}y \\
&\leq \sigma_{n-1} \int_E \int_{\operatorname{dist}(E,F)}^\infty r^{n-1} G(re_1) \,\mathrm{d}r \,|u(y)|^2 \,\mathrm{d}y,
\end{aligned} \tag{11.1}$$

where e_1 is the first unit vector in $\mathbb{R}^n$. For $r \geq 1$ we have $G(re_1) \lesssim \mathrm{e}^{-r/2}$ by Lemma 11.3. Hence, $r^{n-1} G(re_1) \lesssim r^{n-1} \mathrm{e}^{-r/4} \mathrm{e}^{-r/4} \lesssim \mathrm{e}^{-r/4}$ and by a computation of the $\mathrm{d}r$-integral in (11.1) we derive the required estimate

$$\int_F |(1 - \Delta)^{-1} u|^2 \,\mathrm{d}x \lesssim \mathrm{e}^{-\frac{\operatorname{dist}(E,F)}{4}} \int_E |u(y)|^2 \,\mathrm{d}y. \qquad \square$$

The resolvent decay with respect to $\operatorname{dist}(E, F)$ in Proposition 11.4 has been a consequence of the pointwise bound for the resolvent kernel $(x, y) \mapsto G(x - y)$ away from the 'diagonal' $\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : x = y\}$, hence the name *off-diagonal estimates*. There might be other reasons why such a bound is true and we turn the outcome of Proposition 11.4 into a definition. The concept, though for semigroups rather than resolvents, dates back to Gaffney's results for the heat equation on Riemannian manifolds [Gaf59] and was popularized by Davies in the context of general elliptic operators [Dav95]. Their fundamental contributions are reflected in the terminology: L^2-off-diagonal estimates are also known as *Davies–Gaffney estimates*.

Definition 11.5. *A family* $(T_t)_{t>0} \subseteq \mathcal{L}(\mathrm{L}^2(\mathbb{R}^n))$ *satisfies* L^2-off-diagonal estimates *if there exist* $C, c > 0$ *such that for all measurable sets* $E, F \subseteq \mathbb{R}^n$ *and all* $u \in \mathrm{L}^2(\mathbb{R}^n)$ *with* $\operatorname{supp}(u) \subseteq E$ *we have*

$$\|T_t u\|_{\mathrm{L}^2(F)} \leq C \mathrm{e}^{-c \frac{\operatorname{dist}(E,F)}{t}} \|u\|_{\mathrm{L}^2(E)}.$$

The same terminology is used for families of linear operators acting between tuples of $\mathrm{L}^2(\mathbb{R}^n)$*-functions.*

Remark 11.6. At this point the scaling parameter t does not play an essential role and readers should consider $t = 1$ and a single operator $T = T_1$ on a first reading. In fact, for the families we are going to be interested in, the dependence of the right-hand side with respect to t can also be obtained by an a posteriori scaling argument, see Exercise 11.5.

11.2. Off-diagonal estimates for elliptic operators

When we replace $-\Delta$ by L, there is no reason to believe that the resolvent is still given by an integrable kernel in any reasonable sense. (We will support that claim of ours by some further evidence later during this lecture.) However, L^2-off-diagonal estimates remains valid as we shall see next. Of course, the strategy of proof has to be dramatically different.

We start with uniform bounds with respect to the scaling parameter t, which in any case is necessary since we may take $E = F = \mathbb{R}^n$ in Definition 11.5. In the following, it will be convenient to use *second order scaling* and work with $(1 + t^2 L)^{-1}$ instead of the usual $(1 + tL)^{-1}$ for $t > 0$.

Lemma 11.7. *For all* $u \in \mathrm{L}^2(\mathbb{R}^n)$ *and all* $t > 0$ *we have the following uniform bounds:*

(a) $\|(1 + t^2 L)^{-1} u\|_{\mathrm{L}^2(\mathbb{R}^n)} \leq \|u\|_{\mathrm{L}^2(\mathbb{R}^n)}$,

(b) $\|t \nabla (1 + t^2 L)^{-1} u\|_{\mathrm{L}^2(\mathbb{R}^n)} \leq \sqrt{2/\lambda} \|u\|_{\mathrm{L}^2(\mathbb{R}^n)}$.

Proof. Part (a) is a direct consequence of m-accretivity (Theorem 2.21). As for (b), we set $w := (1+t^2L)^{-1}u$ and use ellipticity and the definition of L to bound

$$\lambda\|t\nabla w\|^2_{\mathrm{L}^2(\mathbb{R}^n)} \leq t^2 \operatorname{Re}\big(a(w)\big) = t^2 \operatorname{Re}\big(\langle Lw, w\rangle_{\mathrm{L}^2(\mathbb{R}^n)}\big) \leq \|t^2Lw\|_{\mathrm{L}^2(\mathbb{R}^n)}\|w\|_{\mathrm{L}^2(\mathbb{R}^n)}.$$

Now, $t^2Lw = u - w$, so that according to (a) the right-hand side is bounded by $2\|u\|^2_{\mathrm{L}^2(\mathbb{R}^n)}$. This proves (b). □

Theorem 11.8 (L^2-off-diagonal estimates for L)**.** *The following families of operators satisfy* L^2*-off-diagonal estimates:*

(a) $\big((1+t^2L)^{-1}\big)_{t>0}$,

(b) $\big(t\nabla(1+t^2L)^{-1}\big)_{t>0}$.

The implicit constants c and C as in Definition 11.5 depend only on n, λ and Λ.

Proof. We fix measurable sets $E, F \subseteq \mathbb{R}^n$, a function $u \in \mathrm{L}^2(\mathbb{R}^n)$ with support in E and a scaling parameter $t > 0$. We abbreviate $w := (1+t^2L)^{-1}u$ and $d := \operatorname{dist}(E, F)$.

Off-diagonal estimates only provide additional information (as compared to uniform boundedness) when d/t is large and similar to the proof for the negative Laplacian in Proposition 11.4 we distinguish two cases. However, it will be advantageous to keep the threshold $\alpha > 0$ variable, depending on n, λ and Λ.

Case 1: $d \leq \alpha t$.

Lemma 11.7 directly yields the required bound

$$\begin{aligned}\|w\|_{\mathrm{L}^2(F)} + \|t\nabla w\|_{\mathrm{L}^2(F)} \leq \|w\|_{\mathrm{L}^2(\mathbb{R}^n)} + \|t\nabla w\|_{\mathrm{L}^2(\mathbb{R}^n)} &\leq (1+\sqrt{2/\lambda})\|u\|_{\mathrm{L}^2(E)} \\ &\leq (1+\sqrt{2/\lambda})\mathrm{e}^{\alpha}\cdot \mathrm{e}^{-\frac{d}{t}}\|u\|_{\mathrm{L}^2(E)}.\end{aligned}$$

Case 2: $d > \alpha t$.

We fix a real-valued, bounded function $\rho \in \mathrm{C}^\infty(\mathbb{R}^n)$ with $\rho = 1$ on F, $\rho = 0$ on E and

$$d\|\nabla\rho\|_{\mathrm{L}^\infty(\mathbb{R}^n)} \leq C \tag{11.2}$$

for some constant C that only depends on n.[2] The principal idea of the proof is to test the elliptic equation $(1+t^2L)w = u$ with the following function:

$$v := w\eta^2, \quad \text{where} \quad \eta := \mathrm{e}^{\frac{d}{\alpha t}\rho} - 1.$$

[2]Here is one construction. Take $\phi \in \mathrm{C}_\mathrm{c}^\infty(B(0,1))$ with integral 1 and smooth out the characteristic function of $\Omega_F := \{x \in \mathbb{R}^n \mid \operatorname{dist}(x, F) \leq d/3\}$ — the function $\rho := \mathbf{1}_{\Omega_F} * \phi_{d/3}$ does the job.

We note that $v \in \mathrm{H}^1(\mathbb{R}^n)$ since η is smooth and bounded with bounded derivative (Lemma 3.9) and in addition we have

$$\begin{aligned} \eta &= 0 \quad (\text{on } E), \\ \eta &= \mathrm{e}^{\frac{d}{\alpha t}} - 1 \geq \tfrac{1}{2}\mathrm{e}^{\frac{d}{\alpha t}} \quad (\text{on } F), \\ |\nabla \eta| &= |\tfrac{d}{\alpha t}\nabla\rho \cdot \mathrm{e}^{\frac{d}{\alpha t}\rho}| \leq \tfrac{C}{\alpha t}|\eta + 1|, \end{aligned} \tag{11.3}$$

where we have used $d > \alpha t$ and the inequality $\mathrm{e}^s \geq 2$ for $s \geq 1$ to derive the exponential lower bound on F.

Since $u = (1 + t^2 L)w$, we find by definition of L that

$$\begin{aligned} \langle u, v\rangle_{\mathrm{L}^2(\mathbb{R}^n)} &= \langle w, v\rangle_{\mathrm{L}^2(\mathbb{R}^n)} + t^2 a(w, v) \\ &= \int_{\mathbb{R}^n} |w|^2\eta^2 \,\mathrm{d}x + t^2 \int_{\mathbb{R}^n} A\nabla w \cdot \overline{\nabla(w\eta^2)} \,\mathrm{d}x \\ &= \int_{\mathbb{R}^n} |w|^2\eta^2 \,\mathrm{d}x + t^2 \int_{\mathbb{R}^n} A\nabla w \cdot \eta^2\overline{\nabla w} + A\nabla w \cdot 2\eta\overline{w}\nabla\eta \,\mathrm{d}x. \end{aligned}$$

As $v = w\eta^2 = 0$ on $E \supseteq \operatorname{supp}(u)$, the left-hand side vanishes and we get

$$\int_{\mathbb{R}^n} |w|^2\eta^2 \,\mathrm{d}x + t^2 \int_{\mathbb{R}^n} (A\nabla w \cdot \overline{\nabla w})\eta^2 \,\mathrm{d}x = -t^2 \int_{\mathbb{R}^n} A\eta\nabla w \cdot 2\overline{w}\nabla\eta \,\mathrm{d}x.$$

We control the real part of the left-hand side from below by ellipticity of A and the absolute value of the right-hand side from above by boundedness of A. Consequently,

$$\int_{\mathbb{R}^n} |w|^2\eta^2 \,\mathrm{d}x + \lambda t^2 \int_{\mathbb{R}^n} |\nabla w|^2\eta^2 \,\mathrm{d}x \leq t^2 \int_{\mathbb{R}^n} |\eta\nabla w| \cdot 2\Lambda|w\nabla\eta| \,\mathrm{d}x$$

and the 'Peter–Paul inequality'[3] $ab \leq {\varepsilon a^2}/{2} + {b^2}/{2\varepsilon}$ for positive numbers yields

$$\leq \frac{\lambda t^2}{2} \int_{\mathbb{R}^n} |\nabla w|^2\eta^2 \,\mathrm{d}x + \frac{2\Lambda^2 t^2}{\lambda} \int_{\mathbb{R}^n} |w|^2|\nabla\eta|^2 \,\mathrm{d}x.$$

We have chosen the prefactor for the first term on the right in such a way that it can be absorbed back into the left-hand side, resulting in the bound

$$\int_{\mathbb{R}^n} |w|^2\eta^2 \,\mathrm{d}x + \frac{\lambda t^2}{2} \int_{\mathbb{R}^n} |\nabla w|^2\eta^2 \,\mathrm{d}x \leq \frac{2\Lambda^2 t^2}{\lambda} \int_{\mathbb{R}^n} |w|^2|\nabla\eta|^2 \,\mathrm{d}x.$$

Since terms with η are potentially very large, see (11.3), we need to move all η-dependency to the left. To this end, the algebraic properties of η come in handy. Indeed, we have $|\nabla\eta|^2 \leq \frac{2C^2}{\alpha^2 t^2}(\eta^2 + 1)$ by (11.3) and therefore

$$\int_{\mathbb{R}^n} |w|^2\eta^2 \,\mathrm{d}x + \frac{\lambda t^2}{2} \int_{\mathbb{R}^n} |\nabla w|^2\eta^2 \,\mathrm{d}x \leq \frac{4\Lambda^2 C^2}{\lambda\alpha^2} \int_{\mathbb{R}^n} |w|^2(\eta^2 + 1) \,\mathrm{d}x,$$

[3] One must 'rob Peter to pay Paul'.

so that fixing the threshold $\alpha := \sqrt{8/\lambda}\,\Lambda C$ indeed leads us to

$$\frac{1}{2}\int_{\mathbb{R}^n} |w|^2\eta^2 \,\mathrm{d}x + \frac{\lambda t^2}{2}\int_{\mathbb{R}^n} |\nabla w|^2\eta^2 \,\mathrm{d}x \leq \frac{1}{2}\int_{\mathbb{R}^n} |w|^2 \,\mathrm{d}x.$$

Lemma 11.7 allows us to estimate the L^2-norm of w on the right by the one of u and on the left we can use the exponential lower bound for η on F in (11.3) to conclude

$$\frac{1}{8}\mathrm{e}^{\frac{2d}{\alpha t}}\int_F |w|^2 \,\mathrm{d}x + \frac{\lambda}{8}\mathrm{e}^{\frac{2d}{\alpha t}}\int_F |t\nabla w|^2 \,\mathrm{d}x \leq \frac{1}{2}\int_E |u|^2 \,\mathrm{d}x.$$

By definition of w, this is the required off-diagonal estimate. □

There is a straightforward duality principle for off-diagonal estimates.

Lemma 11.9. *If $(T_t)_{t>0} \subseteq \mathcal{L}(\mathrm{L}^2(\mathbb{R}^n))$ satisfies L^2-off-diagonal estimates, then so does the adjoint family $(T_t^*)_{t>0}$ and the implicit constants $C, c > 0$ can be taken the same.*

Proof. Fix $t > 0$, measurable sets $E, F \subseteq \mathbb{R}^n$ and a function $u \in \mathrm{L}^2(\mathbb{R}^n)$ with support in E. For any $v \in \mathrm{L}^2(\mathbb{R}^n)$ with support in F we have

$$\begin{aligned} |\langle T_t^* u, v\rangle_{\mathrm{L}^2(\mathbb{R}^n)}| = |\langle u, T_t v\rangle_{\mathrm{L}^2(\mathbb{R}^n)}| &\leq \|u\|_{\mathrm{L}^2(E)}\|T_t v\|_{\mathrm{L}^2(E)} \\ &\leq \|u\|_{\mathrm{L}^2(E)} C\mathrm{e}^{-c\frac{\operatorname{dist}(F,E)}{t}}\|v\|_{\mathrm{L}^2(F)}, \end{aligned}$$

and taking $v = \mathbf{1}_F T_t^* u$ yields the claim. □

Applying the duality principle to the resolvent family leads us to the resolvents of the adjoint operator $L^* = -\operatorname{div}(A^*\nabla\,\cdot)$, which is an operator of the same type as L, compare with Lemma 6.7 and Theorem 3.29 (c). Things are different for the gradient of the resolvent, whose adjoint will play an important role in the upcoming lectures. To describe this family properly, we need the *very weak divergence operator*

$$\operatorname{div}\colon \mathrm{L}^2(\mathbb{R}^n)^n \to \mathrm{H}^{-1}(\mathbb{R}^n), \quad \langle \operatorname{div} U, v\rangle_{\mathrm{H}^{-1}(\mathbb{R}^n),\mathrm{H}^1(\mathbb{R}^n)} = -\int_{\mathbb{R}^n} U \cdot \overline{\nabla v} \,\mathrm{d}x. \tag{11.4}$$

It is clearly bounded and its name stems from the fact that if $U = (U_1, \ldots, U_n) \in \mathrm{H}^1(\mathbb{R}^n)^n$, then integration by parts via Corollary 3.19 reveals that $\operatorname{div} U = \sum_{j=1}^n \partial_j U_j$ is the divergence taken in the sense of weak derivatives. Implicitly, this definition has always been around. Namely, the definition of

$$\mathscr{L}\colon \mathrm{H}^1(\mathbb{R}^n) \to \mathrm{H}^{-1}(\mathbb{R}^n), \quad \langle \mathscr{L}u, v\rangle_{\mathrm{H}^{-1}(\mathbb{R}^n),\mathrm{H}^1(\mathbb{R}^n)} = \int_{\mathbb{R}^n} A\nabla u \cdot \overline{\nabla v} \,\mathrm{d}x$$

says precisely that $\mathscr{L} = -\operatorname{div}(A\nabla\,\cdot)$ can be understood as a three-fold composition of operators: weak gradient, multiplication by A, negative very weak divergence.

Lemma 11.10. *Let $t > 0$. The adjoint of* $\nabla(1 + t^2 L^*)^{-1} \subseteq \mathrm{L}^2(\mathbb{R}^n) \times \mathrm{L}^2(\mathbb{R}^n)^n$ *is given by* $-(1 + t^2 \mathscr{L})^{-1} \operatorname{div}$.

Proof. Since $\nabla(1 + t^2 L^*)^{-1}$ is bounded (Lemma 11.7), so is its adjoint (Proposition 1.20 (h)). The operator $-(1 + t^2 \mathscr{L})^{-1} \operatorname{div} \subseteq \mathrm{L}^2(\mathbb{R}^n)^n \times \mathrm{L}^2(\mathbb{R}^n)$ is also bounded as a composition of bounded operators (Corollary 2.20). Given $U \in \mathrm{H}^1(\mathbb{R}^n)^n$ and $v \in \mathrm{L}^2(\mathbb{R}^n)$, integration by parts and duality for resolvents as above yield

$$\begin{aligned}\big\langle U, \nabla(1 + t^2 L^*)^{-1} v\big\rangle_{\mathrm{L}^2(\mathbb{R}^n)^n} &= \big\langle -\operatorname{div} U, (1 + t^2 L^*)^{-1} v\big\rangle_{\mathrm{L}^2(\mathbb{R}^n)} \\ &= \big\langle -(1 + t^2 L)^{-1} \operatorname{div} U, v\big\rangle_{\mathrm{L}^2(\mathbb{R}^n)} \\ &= \big\langle -(1 + t^2 \mathscr{L})^{-1} \operatorname{div} U, v\big\rangle_{\mathrm{L}^2(\mathbb{R}^n)},\end{aligned}$$

where in the last step we have used that $(1 + t^2 L)^{-1}$ is the restriction of $(1 + t^2 \mathscr{L})^{-1}$ to $\mathrm{L}^2(\mathbb{R}^n)$, see (the proof of) Theorem 2.21. By density, this identity extends to all $U \in \mathrm{L}^2(\mathbb{R}^n)^n$ and the claim follows. □

In view of the preceding two lemmas, we can add the following result to Theorem 11.8 (applied to L^*).

Corollary 11.11. *The family* $(t(1 + t^2 \mathscr{L})^{-1} \operatorname{div})_{t>0}$ *satisfies* L^2*-off-diagonal estimates. Implicit constants c and C as in Definition 11.5 depend only on n, λ and Λ.*

11.3. Consequences of off-diagonal estimates

Consider again the negative Laplacian. By Lemma 11.2, its resolvent is given by

$$(1 - \Delta)^{-1} u = G * u,$$

where G is the Bessel kernel. Since G is integrable, the right-hand side makes sense for all bounded functions and this provides a way to *define* $(1 - \Delta)^{-1} b$ for $b \in \mathrm{L}^\infty(\mathbb{R}^n)$. This extension can also be recovered from the L^2-theory by approximation as follows. We fix any ball $B \subseteq \mathbb{R}^n$ and consider the L^2-functions $\mathbf{1}_{2^k B} b$. Then for almost every $x \in \mathbb{R}^n$ we have by dominated convergence that

$$\begin{aligned}\big((1 - \Delta)^{-1} b\big)(x) := \int_{\mathbb{R}^n} G(x - y) b(y) \,\mathrm{d}y &= \lim_{k \to \infty} \int_{\mathbb{R}^n} G(x - y) \mathbf{1}_{2^k B} b(y) \,\mathrm{d}y \\ &= \lim_{k \to \infty} \big((1 - \Delta)^{-1} \mathbf{1}_{2^k B} b\big)(x).\end{aligned}$$

Using off-diagonal estimates, the same construction can be made in much greater generality.

Proposition 11.12. *Let $(T_t)_{t>0} \subseteq \mathcal{L}(\mathrm{L}^2(\mathbb{R}^n))$ satisfy L^2-off-diagonal estimates and let $b \in \mathrm{L}^\infty(\mathbb{R}^n)$. For any ball $B \subseteq \mathbb{R}^n$ the limit*

$$\lim_{k\to\infty} T_t(\mathbf{1}_{2^k B} b)$$

exists in $\mathrm{L}^2(K)$ for any compact subset $K \subseteq \mathbb{R}^n$ and is independent of the choice of B. Moreover, we obtain the same limit if B is a cube with sides parallel to the coordinate axes (that is, a ball with respect to the ℓ^∞-norm on $\mathbb{R}^n$).

In the proof we will see for the first time how off-diagonal estimates interact nicely with annular decompositions of $\mathbb{R}^n$.

Definition 11.13. *Given a ball $B \subseteq \mathbb{R}^n$, we introduce the annuli*

$$C_1(B) \coloneqq 4B \quad \textit{and} \quad C_\ell(B) \coloneqq 2^{\ell+1}B \setminus 2^\ell B \qquad (\ell \geq 2).$$

We use the same notation if B is a cube with sides parallel to the coordinate axes.

Proof of Proposition 11.12. Let $K \subseteq \mathbb{R}^n$ be compact and $\ell_0 \in \mathbb{N}$ be such that $K \subseteq 2^{\ell_0}B$. Let r be the radius of B. If $\ell \in \mathbb{N}$ is such that $\ell > \ell_0$, then

$$\operatorname{dist}(K, C_\ell(B)) \geq 2^\ell r - 2^{\ell_0} r \geq 2^{\ell-1} r.$$

For $k > j > \ell_0$ we write

$$\mathbf{1}_{2^k B} b - \mathbf{1}_{2^j B} b = \sum_{\ell=j}^{k-1} \mathbf{1}_{C_\ell(B)} b$$

and note that $\|\mathbf{1}_{C_\ell(B)} b\|_{\mathrm{L}^2(\mathbb{R}^n)} \leq (2^{\ell+1} r)^{n/2} \omega_n^{1/2} \|b\|_{\mathrm{L}^\infty(\mathbb{R}^n)}$. Off-diagonal estimates yield

$$\begin{aligned}
\|T_t(\mathbf{1}_{2^k B} b) - T_t(\mathbf{1}_{2^j B} b)\|_{\mathrm{L}^2(K)} &\leq \sum_{\ell=j}^{k-1} \|T_t(\mathbf{1}_{C_\ell(B)} b)\|_{\mathrm{L}^2(K)} \\
&\leq \sum_{\ell=j}^{k-1} C \mathrm{e}^{-c \frac{2^{\ell-1} r}{t}} (2^{\ell+1} r)^{n/2} \omega_n^{1/2} \|b\|_{\mathrm{L}^\infty(\mathbb{R}^n)},
\end{aligned}$$

where the right-hand side is a piece of a convergent series. This proves that $(T_t(\mathbf{1}_{2^k B} b))_k$ is as Cauchy sequence in $\mathrm{L}^2(K)$ and hence convergent. B-independence of the limit as well as that B could be replaced by a cube follows by an analogous pattern and is left as Exercise 11.2. $\square$

Definition 11.14 (L^∞-extension of off-diagonal families)**.** *In the setting of Proposition 11.12 we define $T_t b \coloneqq \lim_{k\to\infty} T_t(\mathbf{1}_{2^k B} b)$.*

Since the Bessel potential G is positive and has integral 1, the (extended) resolvent of $-\Delta$ has the *conservation property*

$$(1-\Delta)^{-1}\mathbf{1}_{\mathbb{R}^n} = \mathbf{1}_{\mathbb{R}^n}.$$

Extending the conservation property to general elliptic operators in divergence form is one of the many applications for off-diagonal estimates. We are sure that you will enjoy figuring out the surprisingly simple proof on your own and have prepared a guided exercise (Exercise 11.3).

Theorem 11.15 (Conservation property)**.** *For any* $t > 0$ *we have* $(1+t^2L)^{-1}\mathbf{1}_{\mathbb{R}^n} = \mathbf{1}_{\mathbb{R}^n}$.

Besides the possibility of merely defining $(1-\Delta)^{-1}u$ for broader classes of functions u, the representation via the Bessel kernel also provides additional estimates in $\mathrm{L}^p(\mathbb{R}^n)$. In fact, by the convolution inequality in Proposition 3.3 (c) we have

$$\|(1-\Delta)^{-1}u\|_{\mathrm{L}^p(\mathbb{R}^n)} = \|G * u\|_{\mathrm{L}^p(\mathbb{R}^n)} \leq C\|u\|_{\mathrm{L}^p(\mathbb{R}^n)}$$

for any $u \in \mathrm{L}^p(\mathbb{R}^n) \cap \mathrm{L}^2(\mathbb{R}^n)$, where $1 \leq p \leq \infty$ and $C = \|G\|_{\mathrm{L}^1(\mathbb{R}^n)} = 1$. Once again, we may ask for the same type of estimate when $-\Delta$ is replaced by a general elliptic operator in divergence form . . .

The surprising answer is that the range of admissible exponents p depends on the operator. When the coefficients A are real, all exponents $p \in [1,\infty]$ are admissible but for complex coefficients in dimension $n \geq 3$ the range can be different. In particular, there exist elliptic operators in divergence form whose resolvents are not given by kernels that behave similarly to G. As of now, we refer to [AE23, Sec. 1.3] for a historical account and hope to learn more about it during the project phase. However, when $p \in (1,\infty)$ is such that

$$\left|\frac{1}{p} - \frac{1}{2}\right| < \frac{1}{n}, \tag{11.5}$$

then L^p-extrapolation of resolvents always works.

Theorem 11.16 (L^p-extrapolation of resolvents)**.** *If* $p \in (1,\infty)$ *satisfies* (11.5)*, then there exists* $C > 0$ *such that for all* $t > 0$ *and all* $u \in \mathrm{L}^p(\mathbb{R}^n) \cap \mathrm{L}^2(\mathbb{R}^n)$ *we have*

$$\|(1+t^2L)^{-1}u\|_{\mathrm{L}^p(\mathbb{R}^n)} \leq C\|u\|_{\mathrm{L}^p(\mathbb{R}^n)}.$$

The proof relies on Sobolev embeddings and a beautiful application of off-diagonal estimates. It follows a general pattern that can be applied to a variety of other operators on $\mathrm{L}^2(\mathbb{R}^n)$ that have a smoothing property in the Sobolev scale and a decay property in $\mathrm{L}^2(\mathbb{R}^n)$.

Proof of Theorem 11.16. We present the argument in dimension $n \geq 3$ and refer to Exercise 11.6 for the modifications in dimensions 1 and 2.

The proof is divided into four steps. Since $n \geq 3$, we can introduce $q := 2^* \in (2, \infty)$ as in Definition 10.5 and the range for p in (11.5) is precisely $q' < p < q$. In the first three steps we assume $2 < p < q$ and in the final step we treat the remaining exponents by duality. Throughout the proof, we write $R_t := (1 + t^2 L)^{-1}$. Implicit constants will only depend on n, λ, Λ and p.

Step 1: We prove an L^2-L^q*-estimate.*

Let $t > 0$ and $u \in \mathrm{L}^2(\mathbb{R}^n)$. We simply join the Sobolev embedding in Corollary 10.10 with Lemma 11.7:

$$\|R_t u\|_{\mathrm{L}^q(\mathbb{R}^n)} \lesssim \|\nabla R_t u\|_{\mathrm{L}^2(\mathbb{R}^n)} \lesssim t^{-1} \|u\|_{\mathrm{L}^2(\mathbb{R}^n)}.$$

Step 2: We interpolate to get an L^2-L^p*-estimate with decay.*

We fix $p \in (2, q)$ and define $\theta \in (0,1)$ via $1/p = (1-\theta)/2 + \theta/q$. By definition of q, we also have $1/p = 1/2 - \theta/n$.

Let $E, F \subseteq \mathbb{R}^n$ be measurable, $u \in \mathrm{L}^2(\mathbb{R}^n)$ be supported in E and $t > 0$. We use Hölder's inequality

$$\|R_t u\|_{\mathrm{L}^p(F)} \leq \left\| |R_t u|^{1-\theta} \right\|_{\mathrm{L}^{2/(1-\theta)}(F)} \left\| |R_t u|^{\theta} \right\|_{\mathrm{L}^{q/\theta}(F)} = \|R_t u\|_{\mathrm{L}^2(F)}^{1-\theta} \|R_t u\|_{\mathrm{L}^q(F)}^{\theta}$$

and then apply off-diagonal estimates (Theorem 11.8) to the first and the L^2–L^q-bound from Step 1 to the second term to find

$$\|R_t u\|_{\mathrm{L}^p(F)} \lesssim t^{-\theta} \mathrm{e}^{-c(1-\theta)\frac{\operatorname{dist}(E,F)}{t}} \|u\|_{\mathrm{L}^2(E)}. \tag{11.6}$$

The power of t is tied to the dimension via $\theta = n/2 - n/p > 0$. Such an estimate is also called an L^2-L^p*-off-diagonal estimate*. In the following, we will write $2c$ instead of $c(1-\theta)$ for the positive factor in the exponential function, the precise value of which does not matter in our proof.

Step 3: We use off-diagonal decay to deduce the L^p*-estimate.*

We fix $t > 0$ and $u \in \mathrm{L}^p(\mathbb{R}^n) \cap \mathrm{L}^2(\mathbb{R}^n)$. We split $\mathbb{R}^n = \bigcup_k Q_k$ into the union of the pairwise disjoint, half-open cubes of sidelength t given by $Q_k := tk + [0,t)^n$ with $k \in \mathbb{Z}^n$, and accordingly, we split

$$\|R_t u\|_{\mathrm{L}^p(\mathbb{R}^n)}^p = \sum_k \|R_t u\|_{\mathrm{L}^p(Q_k)}^p.$$

For fixed k, we write $R_t u = \sum_{\ell=1}^{\infty} R_t(\mathbf{1}_{C_\ell(Q_k)} u)$. This sum converges in $\mathrm{L}^p(\mathbb{R}^n)$, because $R_t : \mathrm{L}^2(\mathbb{R}^n) \to \mathrm{L}^p(\mathbb{R}^n)$ is bounded by (11.6) with $E = F = \mathbb{R}^n$ and $u = \sum_{\ell=1}^{\infty} \mathbf{1}_{C_\ell(Q_k)} u$ in $\mathrm{L}^2(\mathbb{R}^n)$. Hence, we can continue by

$$\leq \sum_k \left(\sum_{\ell=1}^{\infty} \|R_t(\mathbf{1}_{C_\ell(Q_k)} u)\|_{\mathrm{L}^p(Q_k)} \right)^p.$$

At this point we can use (11.6) with the sets $E = C_\ell(Q_k)$ and $F = Q_k$. We have $\operatorname{dist}(E,F) = (2^\ell - 1)\frac{t}{2}$ if $\ell \geq 2$ and $\operatorname{dist}(E,F) = 0$ for $\ell = 1$. Hence, we have $\operatorname{dist}(E,F) \geq (2^\ell - 2)\frac{t}{2}$ in any case, which leads us to

$$\lesssim \sum_k \bigg(\sum_{\ell=1}^\infty t^{n/p-n/2} \mathrm{e}^{-c(2^\ell-2)} \|u\|_{\mathrm{L}^2(C_\ell(Q_k))} \bigg)^p.$$

This is still an L^2–L^p-estimate, but on the sets $C_\ell(Q_k) \subseteq 2^{\ell+1}Q_k$ of measure bounded by $(2^{\ell+1}t)^n$. We can use Hölder's inequality in order to get L^p-norms and cancel all factors of t:

$$\begin{aligned} &\leq \sum_k \bigg(\sum_{\ell=1}^\infty t^{n/p-n/2} \mathrm{e}^{-c(2^\ell-2)} (2^{\ell+1}t)^{n/2-n/p} \|u\|_{\mathrm{L}^p(2^{\ell+1}Q_k)} \bigg)^p \\ &\simeq \sum_k \bigg(\sum_{\ell=1}^\infty 2^{\ell(n/2-n/p)} \mathrm{e}^{-c2^\ell} \|u\|_{\mathrm{L}^p(2^{\ell+1}Q_k)} \bigg)^p. \end{aligned}$$

In order to pull the p-th power into the inner sum, we split $\mathrm{e}^{-c2^\ell} = \mathrm{e}^{-\frac{c}{p'}2^\ell} \mathrm{e}^{-\frac{c}{p}2^\ell}$ and use Hölder's inequality for the summation in ℓ. One factor is just a convergent (numerical) series in ℓ and the remaining one is

$$\begin{aligned} &\lesssim \sum_k \sum_{\ell=1}^\infty \mathrm{e}^{-c2^\ell} \|u\|^p_{\mathrm{L}^p(2^{\ell+1}Q_k)} \\ &= \sum_{\ell=1}^\infty \mathrm{e}^{-c2^\ell} \int_{\mathbb{R}^n} \bigg(\sum_k \mathbf{1}_{2^{\ell+1}Q_k} \bigg) |u|^p \, \mathrm{d}x, \end{aligned}$$

where we have used Tonelli's theorem and monotone convergence in the second step. By a simple counting argument (Exercise 11.4), every $x \in \mathbb{R}^n$ is contained in exactly $(2^{\ell+1})^n$ of the enlarged cubes $2^{\ell+1}Q_k$. Hence, we can conclude

$$= \sum_{\ell=1}^\infty (2^{\ell+1})^n \mathrm{e}^{-c2^\ell} \int_{\mathbb{R}^n} |u|^p \, \mathrm{d}x =: C\|u\|^p_{\mathrm{L}^p(\mathbb{R}^n)}$$

and the value C of the numerical series is finite. This is the required L^p-estimate.

Step 4: We treat $p < 2$ by duality.

Finally, we let $p \in (q', 2)$ and $u \in \mathrm{L}^p(\mathbb{R}^n) \cap \mathrm{L}^2(\mathbb{R}^n)$. Step 3 applied to L^* in place of L on $\mathrm{L}^{p'}(\mathbb{R}^n)$, duality for the resolvents and Hölder's inequality yield for all $v \in \mathrm{C}_c^\infty(\mathbb{R}^n)$ with $\|v\|_{\mathrm{L}^{p'}(\mathbb{R}^n)} \leq 1$ the bound

$$\begin{aligned} \big|\langle R_t u, v\rangle_{\mathrm{L}^2(\mathbb{R}^n)}\big| &= \big|\langle u, (1+t^2L^*)^{-1}v\rangle_{\mathrm{L}^2(\mathbb{R}^n)}\big| \\ &\leq \|u\|_{\mathrm{L}^p(\mathbb{R}^n)} \|(1+t^2L^*)^{-1}v\|_{\mathrm{L}^{p'}(\mathbb{R}^n)} \lesssim \|u\|_{\mathrm{L}^p(\mathbb{R}^n)}. \end{aligned}$$

The L^p-estimate on $R_t u$ follows by taking the supremum over all such v. □

11.4. Exercises

Exercise 11.1. Let $u \in \mathrm{L}^2(\mathbb{R}^n)$ satisfy $u \geq 0$ a.e. and $u > 0$ on a set of strictly positive measure. Prove that $(1 - \Delta)^{-1} u > 0$ a.e. on $\mathbb{R}^n$.

Exercise 11.2. Prove that the limit in Proposition 11.12 is independent of the chosen ball. Then argue that we obtain the same limit when B is a cube with sides parallel to the coordinate axes.

Hint: Given two balls $B, B' \subseteq \mathbb{R}^n$, start by picking $k \in \mathbb{N}$ such that $B' \subseteq 2^k B$.

Exercise 11.3 (Proof of the conservation property)**.** In this exercise you are going to prove Theorem 11.15 by proceeding as follows.

(a) We set $u := (1 + t^2 L)^{-1} \mathbf{1}_{\mathbb{R}^n}$. Make a formal computation that supports the claim that $u = \mathbf{1}_{\mathbb{R}^n}$.

For a rigorous proof, we fix a function $\phi \in \mathrm{C}_c^\infty(B(0,2))$ with $\phi = 1$ on $B(0,1)$ and set $\phi^k := \phi(2^{-k} \,\cdot)$ for $k \in \mathbb{N}$.

(b) Show that for every compact set $K \subseteq \mathbb{R}^n$ we have

$$u = \lim_{k \to \infty} (1 + t^2 L)^{-1} \phi^k \qquad (\text{in } \mathrm{L}^2(K)).$$

This is a smoothed version of Proposition 11.12.

(c) Prove that

$$\phi^k - (1 + t^2 L)^{-1} \phi^k = -t^2 \big((1 + t^2 \mathscr{L})^{-1} \operatorname{div}\big)(A \nabla \phi^k).$$

(d) Use off-diagonal estimates to prove that in the limit as $k \to \infty$ the left-hand side in (c) tends to 0 in a suitable sense.

(e) Conclude.

Exercise 11.4 (Counting cubes . . .)**.** For $t > 0$, consider the partition of $\mathbb{R}^n$ into the pairwise disjoint, half-open cubes $Q_k := tk + [0, t)^n$ with $k \in \mathbb{Z}^n$. Given $\ell \in \mathbb{N}$, prove that each point $x \in \mathbb{R}^n$ is contained in exactly $(2^\ell)^n$ of the cubes $2^\ell Q_k$.

Exercise 11.5 (Rescaling of the coefficients)**.** For $t > 0$ we consider the dilation operator $\delta_t u(x) = u(tx)$ acting on measurable functions. We set $A_t := \delta_t A$ and introduce the elliptic operators in divergence form $L_t := -\operatorname{div}(A_t \nabla \,\cdot)$.

(a) Quickly recall that the coefficients A_t are elliptic and bounded with the same parameters λ, Λ.

(b) Prove that $L_t = t^2 \delta_t L \delta_{t^{-1}}$.

(c) Conclude that $(1 + L_t)^{-1} = \delta_t (1 + t^2 L)^{-1} \delta_{t^{-1}}$.

The similarity relations in (b) and (c) allow us to restrict ourselves to resolvent parameter $t = 1$ in certain estimates, as long as implied constants depend on L only through n, λ and Λ. Here is an example, which brings us back to Remark 11.6.

(d) Suppose we had only shown the off-diagonal bound

$$\|(1+L)^{-1}u\|_{\mathrm{L}^2(F)} \leq C\mathrm{e}^{-c\operatorname{dist}(E,F)}\|u\|_{\mathrm{L}^2(E)}$$

for all elliptic operators L in divergence form, all measurable sets $E, F \subseteq \mathbb{R}^n$ and all $u \in \mathrm{L}^2(E)$, but with constants $C, c > 0$ that only depend on n, λ and Λ. Prove that $((1+t^2L)^{-1})_{t>0}$ satisfies L^2-off-diagonal estimates without resorting to Theorem 11.8.

Exercise 11.6 (L^p-extrapolation in low dimensions)**.** In this exercise you are going to complete the proof of Theorem 11.16 by treating the case of dimension $n \leq 2$.

(a) Carefully go through the proof of Theorem 11.16 once again and convince yourself that the only open task in dimension $n \leq 2$ is to prove the L^2–L^q-estimate in Step 1 for all $q \in (2, \infty)$.

(b) Find a substitute for the Sobolev embedding to complete Step 1 for an arbitrary $q \in (2, \infty)$.

12. Square roots of elliptic operators

The Kato property for square roots has appeared in three of the previous episodes: First, in Lecture 4 for the negative Laplacian on $\mathbb{R}^n$, next for self-adjoint operators as part of Kato's second representation theorem in Lecture 6, and then in Lecture 8, where we have also learned that relying exclusively on functional calculus is inadequate for determining whether or not, for a given operator L associated with a bounded, elliptic and sectorial sesquilinear form a, the domains of $\sqrt{L}$ and a are the same. In the remaining three lectures, we shall provide an affirmative answer to this question for elliptic operators in divergence form on $\mathbb{R}^n$ through the following theorem.

Theorem 12.1 (The solution of the Kato conjecture)**.** *Let* $L = -\operatorname{div}(A\nabla\,\cdot)$ *be an elliptic operator in divergence form on* $\mathbb{R}^n$*. Then* $\operatorname{dom}(\sqrt{L}) = \mathrm{H}^1(\mathbb{R}^n)$ *and*

$$\|\sqrt{L}u\|_{\mathrm{L}^2(\mathbb{R}^n)} \simeq \|\nabla u\|_{\mathrm{L}^2(\mathbb{R}^n)} \qquad (u \in \mathrm{H}^1(\mathbb{R}^n)). \tag{12.1}$$

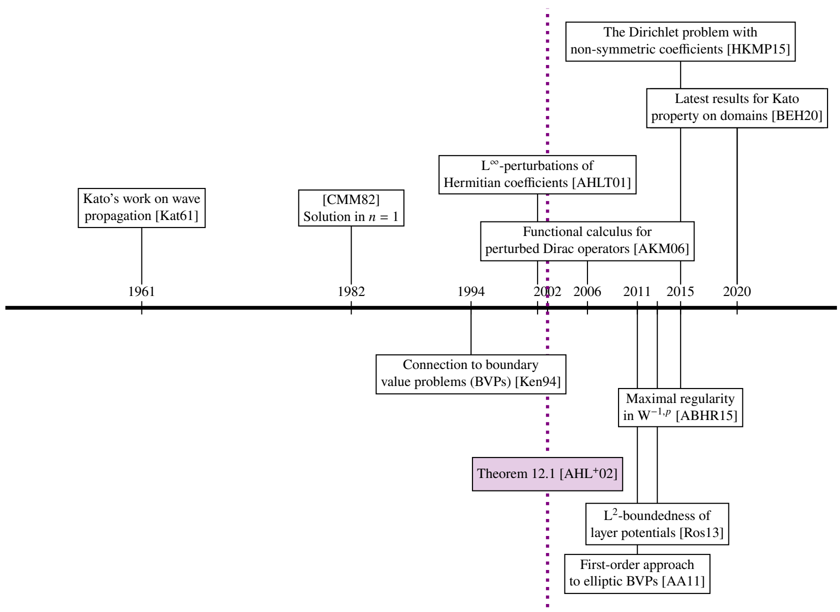

Figure 12.1.: The timeline towards the solution of the Kato conjecture by Auscher–Hofmann–Lacey–McIntosh–Tchamitchian in 2002 and a personal selection of implications and follow-ups related to the material of these lecture notes.

This fundamental theorem, proved by Auscher–Hofmann–Lacey–McIntosh–Tchamitchian in the early 2000s [AHL+02], represents the pinnacle of nearly four decades of research. Acknowledging all individuals who have contributed to the solution is as unfeasible as enumerating all the breakthrough results that have ensued. Nonetheless, we have tried to present a personal timeline in Figure 12.1 with an emphasis on results that are closely connected to the lecture notes and might be studied in more detail during the project phase.

In this lecture, we shall discuss in which sense Theorem 12.1 is the critical regularity result for elliptic operators with L^∞-coefficients and how its proof can be reduced to a very classical concept in harmonic analysis — a square function estimate.

12.1. Why square roots are critical

Let $L = -\operatorname{div}(A\nabla\,\cdot)$ be an elliptic operator in divergence form. Since L is a differential operator of second order, we would naively expect that its domain consists of functions with second-order derivatives in $\mathrm{L}^2(\mathbb{R}^n)$. While this is true for the negative Laplacian (Theorem 4.6), we have seen in Exercise 4.6 that it is plain false in general.[1] Just as naive it seems to 'interpolate' regularity and ask for $\operatorname{dom}(L^\alpha) = \mathrm{H}^{2\alpha}(\mathbb{R}^n)$ when $\alpha \in (0,1)$, but we have seen at several occasions that this actually works. Figure 12.2 summarizes what we know so far on Sobolev regularity for fractional powers of L, including the following new statement.

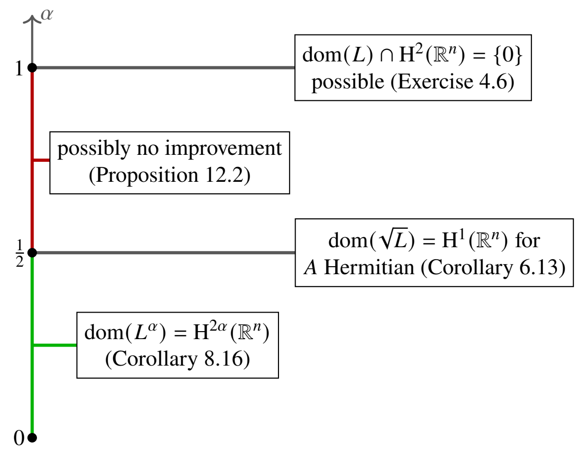

Figure 12.2.: Schematic representation of Sobolev regularity for the domains of the fractional powers L^α. The exponent $\alpha = 1/2$ is critical: There is an optimal result for $\alpha < 1/2$ and Proposition 12.2 provides an operator for which even $\operatorname{dom}(L) \subseteq \mathrm{H}^{2\alpha}(\mathbb{R}^n)$ fails simultaneously for *every* $\alpha \in (1/2, 1)$.

[1] This example was for an operator with Dirichlet boundary conditions on $\Omega = (0,1)$, but it could also be adapted to $\Omega = \mathbb{R}$ by extending the coefficients 1-periodically.

Proposition 12.2. *There exists an elliptic operator L in divergence form on $\mathbb{R}$ and a function $u \in \mathrm{dom}(L)$ such that $u \notin \mathrm{H}^{1+\varepsilon}(\mathbb{R})$ for every $\varepsilon > 0$. In particular, we have* $\mathrm{dom}(L^\alpha) \not\subseteq \mathrm{H}^{2\alpha}(\mathbb{R})$ *for* every $\alpha \in (1/2, 1)$.

Remark 12.3. We mostly work in dimension $n = 1$ for simplicity but once the strategy of proof is in place, the example can be adapted to higher dimensions, see Exercise 12.3. Different, somewhat more complicated examples in dimension $n \geq 2$ can be found in [AT98, Sec. 2.3].

Proposition 12.2 justifies, in retrospective, why we called L^α with $\alpha \in (0, 1/2)$ a 'subcritical fractional power' in Section 8.2 and why square roots are 'critical': It is the exact moment on the fractional power scale where regularity gain in the Sobolev scale breaks down for general L. For the proof, we need two lemmas on fractional Sobolev spaces. Readers may want to recall Definition 4.11 and Proposition 4.13 beforehand. Fractional spaces and norms for the gradient should be interpreted componentwise.

Lemma 12.4 (Lifting property)**.** *For all $\varepsilon \in (0, 1)$ we have*

$$\mathrm{H}^{1+\varepsilon}(\mathbb{R}^n) = \{u \in \mathrm{H}^1(\mathbb{R}^n) \mid \nabla u \in \mathrm{H}^\varepsilon(\mathbb{R}^n)^n\}$$

and $\big(\|\cdot\|^2_{\mathrm{H}^1(\mathbb{R}^n)} + [\nabla\cdot]^2_{\varepsilon,2}\big)^{1/2}$ *defines an equivalent norm on* $\mathrm{H}^{1+\varepsilon}(\mathbb{R}^n)$.

Proof. Let $u \in \mathrm{L}^2(\mathbb{R}^n)$. Then $\widehat{u} \in \mathrm{L}^2(\mathbb{R}^n)$ and all given additional properties are (or can be) characterized via the Fourier transform, see Definition 4.11 and Proposition 4.5:

$$\begin{aligned} u \in \mathrm{H}^{1+\varepsilon}(\mathbb{R}^n) &\iff |\cdot|^{1+\varepsilon}\widehat{u} \in \mathrm{L}^2(\mathbb{R}^n), \\ u \in \mathrm{H}^1(\mathbb{R}^n) &\iff |\cdot|\widehat{u} \in \mathrm{L}^2(\mathbb{R}^n), \\ \nabla u \in \mathrm{H}^\varepsilon(\mathbb{R}^n)^n &\iff |\cdot|^\varepsilon|\cdot\widehat{u}| \in \mathrm{L}^2(\mathbb{R}^n). \end{aligned}$$

Hence, the inclusion '$\supseteq$' is immediate and for '$\subseteq$' we use in addition that $|\xi| \leq 1+|\xi|^{1+\varepsilon}$ for all $\xi \in \mathbb{R}^n$. In passing, we have also proved that an equivalent norm on $\mathrm{H}^{1+\varepsilon}(\mathbb{R}^n)$ is given by $\big(\|\cdot\|^2_{\mathrm{H}^1(\mathbb{R}^n)} + \|\nabla\cdot\|^2_{\mathrm{H}^\varepsilon(\mathbb{R}^n)}\big)^{1/2}$ and, taking into account Proposition 4.13, we can replace $\|\nabla\cdot\|^2_{\mathrm{H}^\varepsilon(\mathbb{R}^n)}$ by $\|\cdot\|^2_{\mathrm{L}^2(\mathbb{R}^n)} + [\nabla\cdot]^2_{\varepsilon,2}$. ☐

The key step is to produce wildly oscillating coefficients. Similar to Exercise 4.6, we will achieve this by adding 1 to the characteristic function of a particularly nasty subset of $[0, 1]$.

Lemma 12.5. *There exists an open subset $E \subseteq [0, 1]$ such that for every $\varepsilon \in (0, 1)$ we have*

$$\int_0^1\int_0^1 \frac{|\mathbf{1}_E(x) - \mathbf{1}_E(y)|^2}{|x-y|^{2\varepsilon}}\,\frac{\mathrm{d}x\,\mathrm{d}y}{|x-y|} = \infty. \tag{12.2}$$

In particular, $\mathbf{1}_E$ is not contained in any fractional Sobolev space $\mathrm{H}^\varepsilon(\mathbb{R})$ with $\varepsilon \in (0, 1)$.

Proof. For a bounded open set E we have $\mathbf{1}_E \in \mathrm{L}^2(\mathbb{R})$ and therefore the first statement implies the second one by Proposition 4.13.

For the construction of E, we let (a_j) be a decreasing sequence of positive numbers with $\sum_{j=1}^{\infty} a_j = 1$. We divide $[0,1]$ into intervals according to the partial sums $s_k := \sum_{j=1}^{k} a_j$, $k \in \mathbb{N}_0$, and define E by selecting every second interval:

$$E := \bigcup_{k=0}^{\infty} (s_{2k}, s_{2k+1}),$$

see Figure 12.3 for a visualization. In order to achieve (12.2), the case of large ε is

Figure 12.3.: An up-to-scale visualization of the set E (purple color) in Lemma 12.5.

easy and would actually work by simply taking $E = (0, s_1)$.

Case 1: $\varepsilon \geq 1/2$.

We simply estimate

$$\begin{aligned}
\int_0^1 \int_0^1 \frac{|\mathbf{1}_E(x) - \mathbf{1}_E(y)|^2}{|x-y|^{2\varepsilon}} \frac{\mathrm{d}x \, \mathrm{d}y}{|x-y|} &\geq \int_0^{s_1} \int_{s_1}^{s_2} \frac{1}{(x-y)^{1+2\varepsilon}} \, \mathrm{d}x \, \mathrm{d}y \\
&= \frac{1}{2\varepsilon} \bigg(\int_0^{s_1} \frac{1}{(s_1 - y)^{2\varepsilon}} \, \mathrm{d}y - \int_0^{s_1} \frac{1}{(s_2 - y)^{2\varepsilon}} \, \mathrm{d}y \bigg).
\end{aligned}$$

Since $s_2 > s_1$, the second integral is finite but the first one is infinite by the assumption $2\varepsilon \geq 1$.

When $\varepsilon < 1/2$, we need to take into account all intervals forming E and make a rather particular choice of (a_j).

Case 2: $\varepsilon < 1/2$.

We split both integrals in (12.2) into the union of the intervals (s_k, s_{k+1}) and discard all interactions but the ones coming from $y \in (s_{2k}, s_{2k+1}) \subseteq E$ and $x \in (s_{2k+1}, s_{2k+2}) \subseteq [0,1] \setminus E$:

$$\begin{aligned}
\int_0^1 \int_0^1 \frac{|\mathbf{1}_E(x) - \mathbf{1}_E(y)|^2}{|x-y|^{2\varepsilon}} \frac{\mathrm{d}x \, \mathrm{d}y}{|x-y|} &\geq \sum_{k=0}^{\infty} \int_{s_{2k}}^{s_{2k+1}} \int_{s_{2k+1}}^{s_{2k+2}} \frac{1}{|x-y|^{1+2\varepsilon}} \, \mathrm{d}x \, \mathrm{d}y \\
&\geq \sum_{k=0}^{\infty} \int_{s_{2k+1}-a_{2k+2}}^{s_{2k+1}} \int_{s_{2k+1}}^{s_{2k+1}+a_{2k+2}} \frac{1}{|x-y|^{1+2\varepsilon}} \, \mathrm{d}x \, \mathrm{d}y,
\end{aligned}$$

where in the second step we have used that the sequence (a_j) is decreasing. Now, $|x-y| \leq 2a_{2k+2}$ on the domain of integration for fixed k, so that we obtain the lower bound

$$\begin{aligned} &\geq \sum_{k=0}^{\infty} \frac{(a_{2k+2})^2}{(2a_{2k+2})^{1+2\varepsilon}} \\ &= 2^{-1-2\varepsilon} \sum_{k=0}^{\infty} (a_{2k+2})^{1-2\varepsilon}. \end{aligned}$$

As $\varepsilon \in (0, 1/2)$, we have $1-2\varepsilon \in (0,1)$, and all we have to do for (12.2) is to name one decreasing sequence (a_j) of positive numbers for which $\sum_{j=1}^{\infty} a_j = 1$, while $\sum_{k=0}^{\infty}(a_{2k+2})^{\alpha} = \infty$ for every $\alpha \in (0,1)$. To this end, we let

$$C := \sum_{j=1}^{\infty} \frac{1}{(j+\mathrm{e}) \log^2(j+\mathrm{e})}$$

be the value of a convergent Bertrand series. Then $a_j := \frac{1}{C(j+\mathrm{e})\log^2(j+\mathrm{e})}$ does the job. □

Remark 12.6. In the language of geometric measure theory, Lemma 12.5 provides an example of a set E that has infinite ε-fractional perimeter in $\mathbb{R}$ for every $\varepsilon > 0$, compare with [CRS10], where this notion has been introduced.

We are ready for the

Proof of Proposition 12.2. The fractional Sobolev spaces are decreasingly ordered in ε. This follows, for instance, from the general properties of fractional powers in Proposition 6.3 (a). Hence, we only have to consider an arbitrary $\varepsilon \in (0,1)$. Likewise, we have $\mathrm{dom}(L) \subseteq \mathrm{dom}(L^{\alpha})$ for all $\alpha \in (1/2, 1)$ and the first part of the claim implies the second one by taking $\varepsilon = 2\alpha - 1$.

For the construction of L, we let $E \subseteq [0,1]$ be as in Lemma 12.5 and consider the elliptic operator

$$L = -\frac{\mathrm{d}}{\mathrm{d}x}\Big(A \frac{\mathrm{d}}{\mathrm{d}x}\,\cdot\Big), \quad \text{where} \quad A := 1 + \mathbf{1}_E.$$

Since $1 \leq A \leq 2$, this operator is elliptic. We define a continuous function by

$$v\colon \mathbb{R} \to \mathbb{R}, \quad v(x) = \int_0^x A^{-1}(y)\,\mathrm{d}y$$

and note that by Theorem 3.12 we have $v' = A^{-1}$ on $\mathbb{R}$ in the weak sense. Next, we pick a function $\eta \in \mathrm{C}_\mathrm{c}^{\infty}(\mathbb{R})$ with $\eta = 1$ on $[0,1]$ and define $u := \eta v$. By construction, u is continuous and satisfies

$$u' = \eta' v + \eta A^{-1} \tag{12.3}$$

in the weak sense, see Lemma 3.9 (c). Since v is continuous and A^{-1} is bounded, both u and u' are bounded with compact support, so that in particular we have $u \in \mathrm{H}^1(\mathbb{R})$.

Next, we claim $u \in \operatorname{dom}(L)$, which by definition is the same as saying $Au' \in \mathrm{H}^1(\mathbb{R})$, see also Exercise 4.6. We have taken $\eta = 1$ on $[0, 1] \supseteq E$ in order to guarantee $A\eta' = \eta'$. Consequently, (12.3) implies $Au' = \eta' v + \eta$, which is again weakly differentiable with derivative

$$(Au')' = \eta'' v + \eta' v' + \eta' \in \mathrm{L}^2(\mathbb{R}).$$

On the other hand, we have $u' = A^{-1} = 1 - \frac{1}{2}\mathbf{1}_E$ on $[0, 1]$ and therefore

$$\int_0^1 \int_0^1 \frac{|u'(x) - u'(y)|^2}{|x-y|^{2\varepsilon}} \, \frac{\mathrm{d}x \, \mathrm{d}y}{|x-y|} = \frac{1}{4} \int_0^1 \int_0^1 \frac{|\mathbf{1}_E(x) - \mathbf{1}_E(y)|^2}{|x-y|^{2\varepsilon}} \, \frac{\mathrm{d}x \, \mathrm{d}y}{|x-y|} = \infty$$

by the defining property of E. According to Proposition 4.13, this means $u' \notin \mathrm{H}^\varepsilon(\mathbb{R})$ and by Lemma 12.4 we conclude $u \notin \mathrm{H}^{1+\varepsilon}(\mathbb{R})$ as required. □

12.2. The Kato property and square functions

The first step in the proof of Theorem 12.1 will be a connection between the Kato property and a certain square function estimate. This works not only for operators in divergence form but in full generality for injective sectorial operators as in Section 2.2.

We let H be a Hilbert space with inner product $\langle \cdot\,, \cdot \rangle$ and norm $\| \cdot \|$, V be a second Hilbert space that is continuously and densely embedded into H, and $a \colon V \times V \to \mathbb{C}$ be a bounded, elliptic and accretive sesquilinear form. As usual, L is the m-accretive operator associated with a in H. In addition, we assume for simplicity that L is injective. Let us recall from Corollary 8.10 that L satisfies quadratic estimates, which in second order scaling, changing variables $s = t^2$ so that $\frac{\mathrm{d}s}{s} = 2\frac{\mathrm{d}t}{t}$, means that for every $f \in \mathrm{H}_0^\infty(\mathrm{S}_\varphi) \setminus \{0\}$ on a sector of angle $\varphi > \pi/2$ we have

$$\int_0^\infty \|f(t^2 L)u\|^2 \, \frac{\mathrm{d}t}{t} \simeq \|u\|^2 \qquad (u \in H). \tag{12.4}$$

Moreover, we recall from Proposition 2.24 and Lemma 6.7 that L^* is an operator of the same type as L.

Proposition 12.7. *In the setting above, the following are equivalent:*

(a) $V \subseteq \operatorname{dom}(\sqrt{L})$ *and* $\|\sqrt{L}u\|^2 \lesssim \operatorname{Re}\big(a(u)\big)$ *for all* $u \in V$.

(b) *The square function estimate*

$$\int_0^\infty \|tL(1+t^2L)^{-1}u\|^2 \, \frac{\mathrm{d}t}{t} \lesssim \operatorname{Re}\big(a(u)\big)$$

holds for all $u \in V$.

Moreover, if L *and* L^* *both satisfy (b), then* $\operatorname{dom}(\sqrt{L}) = V = \operatorname{dom}(\sqrt{L^*})$ *with*

$$\|\sqrt{L}u\|^2 \simeq \operatorname{Re}\big(a(u)\big) \simeq \|\sqrt{L^*}u\|^2 \qquad (u \in V).$$

Proof. We divide the proof into the three (obvious) steps.

Step 1: '(a) $\Longrightarrow$ *(b)':*

Let $u \in V$. Since we have $u \in \operatorname{dom}(\sqrt{L})$ by assumption, we can write

$$tL(1+t^2L)^{-1}u = t\sqrt{L}(1+t^2L)^{-1}\sqrt{L}u = f(t^2L)(\sqrt{L}u),$$

where $f := \sqrt{\mathbf{z}}(1+\mathbf{z})^{-1}$ is of class $\mathrm{H}_0^\infty(\mathrm{S}_\varphi)$ on any sector. Now, (12.4) yields

$$\int_0^\infty \|tL(1+t^2L)^{-1}u\|^2 \, \frac{\mathrm{d}t}{t} \simeq \|\sqrt{L}u\|^2 \tag{12.5}$$

and by assumption, the right-hand side is controlled by $\operatorname{Re}\big(a(u)\big)$. In passing, we remark that we have obtained (12.5) for all $u \in \operatorname{dom}(\sqrt{L})$, independently of the assumption (a).

Step 2: '(b) $\Longrightarrow$ *(a)':*

Let us first prove

$$\|\sqrt{L}u\|^2 \lesssim \operatorname{Re}\big(a(u)\big) \qquad (u \in \operatorname{dom}(L)) \tag{12.6}$$

with an implicit constant that does not depend on the additional assumption that $u \in \operatorname{dom}(L)$. In this case we have $u \in \operatorname{dom}(\sqrt{L})$ by Proposition 6.3. Consequently, (12.5) is applicable and yields

$$\|\sqrt{L}u\|^2 \simeq \int_0^\infty \|tL(1+t^2L)^{-1}u\|^2 \, \frac{\mathrm{d}t}{t}.$$

But as we also have $u \in V$, the assumption controls the right-hand side by $\operatorname{Re}\big(a(u)\big)$ and (12.6) follows.

It remains to remove the a priori assumption $u \in \operatorname{dom}(L)$. To this end, we let $u \in V$ and, according to Proposition 2.23, we pick a sequence $(u_j) \subseteq \operatorname{dom}(L)$ with $u_j \to u$ in V in the limit as $j \to \infty$. In particular, $u_j \to u$ in H and (12.6) implies that $(\sqrt{L}u_j)$ is a Cauchy sequence in H. Since $\sqrt{L}$ is closed, $u \in \operatorname{dom}(\sqrt{L})$ with the same estimate as in (12.6) follows. This proves (a).

Step 3: We suppose (b) (and hence also (a)) for L and L^*.

By symmetry of the assumption, we only have to prove $\operatorname{dom}(\sqrt{L}) \subseteq V$ with the bound $\operatorname{Re}\big(a(u)\big) \lesssim \|\sqrt{L}u\|^2$ for all $u \in \operatorname{dom}(\sqrt{L})$.

If in addition we have $u \in \operatorname{dom}(L)$, then we can use the definition of L, duality for fractional powers (Corollary 6.11) and — since $u \in V$ — also (a) for L^* to find

$$\operatorname{Re}\big(a(u)\big) = \operatorname{Re}\big\langle Lu, u\big\rangle = \operatorname{Re}\big\langle \sqrt{L}u, \sqrt{L^*}u\big\rangle \leq \|\sqrt{L}u\|\|\sqrt{L^*}u\| \lesssim \|\sqrt{L}u\| \operatorname{Re}\big(a(u)\big)^{1/2},$$

which implies indeed that

$$\operatorname{Re}\big(a(u)\big) \lesssim \|\sqrt{L}u\|^2. \tag{12.7}$$

Now, we argue as in the proof of Theorem 6.12: Since a is elliptic, we also have

$$\|u\|_V^2 \lesssim \operatorname{Re}\big(a(u)\big) + \|u\|^2 \lesssim \|\sqrt{L}u\|^2 + \|u\|^2$$

for all $u \in \operatorname{dom}(L)$, and as the latter is a core for $\sqrt{L}$, we conclude $\operatorname{dom}(\sqrt{L}) \subseteq V$ and that (12.7) extends to all $u \in \operatorname{dom}(\sqrt{L})$ by density. □

Remark 12.8. In Proposition 12.7 (b) we have chosen the specific square function that we are actually going to bound in order to prove Theorem 12.1, but apart from that, we could have been much more general. We will leave the task of formulating a broader statement as Exercise 12.2.

Elliptic operators in divergence form on $\mathbb{R}^n$ are injective (Theorem 3.29 (b)) and this class is invariant under taking adjoints. Consequently, we have the following corollary, which will be our starting point in proving Theorem 12.1 in the next lecture.

Corollary 12.9 (Reduction to a square function estimate)**.** *In order to establish Theorem 12.1, it is necessary and sufficient to prove the square function estimate*

$$\int_0^\infty \|tL(1+t^2L)^{-1}u\|_{\mathrm{L}^2(\mathbb{R}^n)}^2 \frac{\mathrm{d}t}{t} \lesssim \|\nabla u\|_{\mathrm{L}^2(\mathbb{R}^n)}^2 \qquad (u \in \mathrm{H}^1(\mathbb{R}^n))$$

for every *elliptic operator in divergence form on* $\mathrm{L}^2(\mathbb{R}^n)$.

We close the lecture by pointing out that the Kato property

$$\|\sqrt{L}u\|^2 \simeq \operatorname{Re}\big(a(u)\big) \qquad (u \in \operatorname{dom}(L)) \tag{12.8}$$

for operators L as above with sectoriality angle $\varphi_L < \pi/2$, not necessarily in divergence form, directly implies quadratic estimates (and hence boundedness of the H^∞-calculus) for L. Of course, this is somewhat pointless from a logical standpoint because we already know the latter property and still struggle to prove the former one, even when

L is a divergence form operator on $\mathbb{R}^n$. We believe that the argument can give some more intuition on why the Kato conjecture is a hard problem, nonetheless.

We use the holomorphic semigroup generated by $-L$ and its properties from Proposition 6.6. Given $u \in H$, we compute

$$\begin{aligned} -\frac{\mathrm{d}}{\mathrm{d}t}\|\mathrm{e}^{-tL}u\|^2 &\overset{6.6\,(\mathrm{d})}{=} 2\,\mathrm{Re}\big\langle L\mathrm{e}^{-tL}u, \mathrm{e}^{-tL}u\big\rangle \\ &= 2\,\mathrm{Re}\big(a(\mathrm{e}^{-tL}u)\big) \\ &\overset{(12.8)}{\simeq} \|\sqrt{L}\mathrm{e}^{-tL}u\|^2. \end{aligned}$$

We integrate both sides with respect to $t \in (0, \infty)$. By Proposition 6.6 (b) and (c), the integral of the left-hand side is equal to $\|u\|^2$ and consequently, we find

$$\|u\|^2 \simeq \int_0^\infty \|\sqrt{tL}\mathrm{e}^{-tL}u\|^2 \,\frac{\mathrm{d}t}{t},$$

which is indeed a quadratic estimate for L with auxiliary function $\sqrt{\mathbf{z}}\mathrm{e}^{-\mathbf{z}}$.

12.3. Exercises

Exercise 12.1 (Weak quadratic estimates)**.** Let $V \subseteq H$ be two Hilbert spaces with continuous and dense embedding, let $a\colon V \times V \to \mathbb{C}$ be a bounded, elliptic and sectorial sesquilinear form and let L be the operator associated with a in H. Prove that

$$\int_0^\infty |\langle tL\mathrm{e}^{-tL}u, v\rangle| \,\frac{\mathrm{d}t}{t} \lesssim \|u\|\|v\| \qquad (u, v \in \overline{\mathrm{ran}(L)}). \tag{12.9}$$

Remark: This is sometimes called a *weak quadratic estimate*. Using specific properties of the function $\mathbf{z}\mathrm{e}^{-\mathbf{z}}$, Cowling–Doust–McIntosh–Yagi [CDMY96, Thm. 4.6] proved that if an injective sectorial operator L in H of angle $\varphi_L < \pi/2$ satisfies (12.9), then L has a bounded H^∞-calculus of any angle $\varphi > \pi/2$. This approach avoids von Neumann's inequality.

Exercise 12.2. In the setting of Proposition 12.7, suggest a general form of square function estimates that follow from (a) by the same proof. Then show that each of them is actually equivalent to (a).

Exercise 12.3 (Failure of $\mathrm{H}^{1+\varepsilon}$-regularity in higher dimensions)**.** Prove that counterexamples as in Proposition 12.2 exist in any dimension $n \geq 1$.

Hint: You can built your construction on an easy(!) extension of the one-dimensional example.

Exercise 12.4 (Perturbed Dirac Operators in a concrete setting)**.** In this exercise, we come back to the perturbed Dirac operators from Exercises 2.4 and 2.5 but eventually we choose D and B concretely and related to elliptic operators in divergence form. This construction is the starting point for the first order approach to elliptic boundary value problems in [AA11].

In the following, we consider the weak gradient ∇ as an unbounded operator from $\mathrm{L}^2(\mathbb{R}^n)$ to $\mathrm{L}^2(\mathbb{R}^n)^n$ with domain $\mathrm{dom}(\nabla) := \mathrm{H}^1(\mathbb{R}^n)$.

(a) Convince yourself that ∇ is closed.

(b) Recall the definition of the very weak divergence operator div from (11.4). Then prove that ∇^* coincides with the maximal restriction of $-\operatorname{div}$ to an operator from $\mathrm{L}^2(\mathbb{R}^n)^n$ to $\mathrm{L}^2(\mathbb{R}^n)$.

(c) Show that

$$D := \begin{pmatrix} 0 & \nabla^* \\ \nabla & 0 \end{pmatrix}$$

with domain $\mathrm{dom}(D) := \mathrm{dom}(\nabla) \times \mathrm{dom}(\nabla^*)$ is a self-adjoint operator in $\mathrm{L}^2(\mathbb{R}^n) \times \mathrm{L}^2(\mathbb{R}^n)^n$.

(d) Let $A \colon \mathbb{R}^n \to \mathcal{L}(\mathbb{C}^n)$ be measurable, essentially bounded and elliptic. Define a bounded linear operator B in $\mathrm{L}^2(\mathbb{R}^n) \times \mathrm{L}^2(\mathbb{R}^n)^n$ by multiplication with the matrix-valued function

$$\begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix}.$$

Use Exercise 2.5 to prove that $(DB)^2$ is sectorial.

(e) Show that

$$(DB)^2 = \begin{pmatrix} L & 0 \\ 0 & K \end{pmatrix}$$

for some operator K and $L = -\operatorname{div}(A\nabla\,\cdot)$ and conclude again that L is sectorial.

Remark: This gives an alternative proof for the sectoriality of L that does not use the Lax–Milgram lemma and sesquilinear forms.

Exercise 12.5 (Multiplicative perturbations of elliptic operators)**.** Let $a \in \mathrm{L}^\infty(\mathbb{R}^n)$ be accretive, that is, there exists $\kappa > 0$ such that $\operatorname{Re} a \geq \kappa$ almost everywhere. As usual, M_a is the associated multiplication operator.

(a) Use Exercise 12.4 to prove that if L is an elliptic operator in divergence form on $\mathbb{R}^n$, then $M_a L$ is sectorial.

(b) Provide an example in which the angle of sectoriality is larger than $\pi/2$.

13. The solution of the Kato conjecture: Part I

We have previously seen in Corollary 12.9 that the Kato conjecture is equivalent to proving the square function estimate

$$\int_0^\infty \|tL(1+t^2L)^{-1}u\|_{\mathrm{L}^2(\mathbb{R}^n)}^2 \frac{\mathrm{d}t}{t} \lesssim \|\nabla u\|_{\mathrm{L}^2(\mathbb{R}^n)^n}^2 \qquad (u \in \mathrm{H}^1(\mathbb{R}^n)) \tag{13.1}$$

for every elliptic operator L in divergence form on $\mathbb{R}^n$. In this lecture, we are gearing up to explore further this particular strategy.

Notation 13.1. Throughout the lecture, $L = -\operatorname{div}(A\nabla\,\cdot)$ is an elliptic operator in divergence form on $\mathbb{R}^n$. The symbols C, c are reserved for positive constants that depend only on n, λ and Λ. We abbreviate $\|\cdot\|_{\mathrm{L}^2(\mathbb{R}^n)}$ by $\|\cdot\|_2$ and use the same notation for the norm of N-tupels of functions $F = (F_j)_{j=1}^N \in \mathrm{L}^2(\mathbb{R}^n)^N$.

In (13.1) we seek control of a square function by the norm of $F := \nabla u$, so it seems wise to write the left-hand side in terms of the same variable as

$$tL(1+t^2L)^{-1}u = t(1+t^2\mathscr{L})^{-1}\mathscr{L}u =: \Theta_t(\nabla u), \tag{13.2}$$

using the following family of operators.

Definition 13.2. *For $t > 0$ define bounded operators $\Theta_t \colon \mathrm{L}^2(\mathbb{R}^n)^n \to \mathrm{L}^2(\mathbb{R}^n)$ by*

$$\Theta_t F := -t(1+t^2\mathscr{L})^{-1}\operatorname{div}(AF),$$

where div *is the very weak divergence operator.*

Remark 13.3. Since multiplication by the bounded, matrix-valued function A does not enlarge the support of functions, the family $(\Theta_t)_{t>0}$ satisfies L^2-off-diagonal estimates by Corollary 11.11.

Naturally, this leads us to working with tuples of functions in the Hilbert spaces $\mathrm{L}^2(\mathbb{R}^n)^N$ and we agree on applying bounded operators on $\mathrm{L}^2(\mathbb{R}^n)$ componentwise in this 'vector-valued' setting.

13.1. About square functions and Carleson measures

It all begins with an idea that was implicit in the '$T(1)$ and $T(b)$ theorems for square functions' due to Christ–Journé [CJ87] and Semmes [Sem90]. The terminology will become clearer during the final Lecture 14. As of now, it suffices to know that these are criteria to prove estimates of the form

$$\int_0^\infty \|\Theta_t F\|_2^2 \,\frac{\mathrm{d}t}{t} \lesssim \|F\|_2^2, \tag{13.3}$$

where, rather than in our setting above, F is scalar-valued and Θ_t are integral operators with kernels that have reasonably good pointwise bounds and some regularity. The function $\gamma_t := \Theta_t \mathbf{1}_{\mathbb{R}^n}$, called *principal part* of Θ_t, plays a decisive role in their analysis. Namely, it turns out that in this setting the estimate

$$\int_0^\infty \|\Theta_t F - \gamma_t \cdot \mathcal{A}_t F\|_2^2 \,\frac{\mathrm{d}t}{t} \lesssim \|F\|_2^2 \tag{13.4}$$

always holds. Here, $\mathcal{A}_t$ is the dyadic averaging operator at scale t from Lecture 9. Put differently, on each scale $t > 0$, they find a sound approximation for $\Theta_t F$ by a multiplication operator acting on the dyadic averages of F. By virtue of the *principal part approximation* (13.4), the square function estimate (13.3) is equivalent to

$$\int_0^\infty \int_{\mathbb{R}^n} |(\mathcal{A}_t F)(x)|^2 |\gamma_t(x)|^2 \,\frac{\mathrm{d}x \,\mathrm{d}t}{t} \lesssim \|F\|_2^2, \tag{13.5}$$

which brings us back to familiar ground. We have seen in Theorem 9.19 and the discussion preceding it that (13.5) holds if and only if the principal part gives rise to a Carleson measure $\mathrm{d}\nu(x,t) = |\gamma_t(x)|^2 \,\frac{\mathrm{d}x\,\mathrm{d}t}{t}$. This is the criterion used by Christ–Journé and Semmes.

Most of the outline above does not apply to the operators Θ_t from Definition 13.2 'off-the-shelf', existence of good integral kernels being the most severe obstruction. Nonetheless, following ideas that first appeared in the work of Auscher–Tchamitchian [AT98], the same reduction to a Carleson measure estimate of a principal part works in case of the square function bound (13.1), which is our goal for this lecture. Demonstrating the Carleson measure estimate will remain for the grande finale.

13.2. The principal part

Since the operators Θ_t in Definition 13.2 act on n-tuples of functions, we need one principal part in each coordinate direction. Off-diagonal estimates allow us to define $\Theta_t b$ for each $b \in \mathrm{L}^\infty(\mathbb{R}^n)^n$ as in Definition 11.14 and in particular, the following definition is meaningful.

Definition 13.4 (Principal part of Θ_t)**.** *Identify the standard unit vectors* $e_1, \dots, e_n \in \mathbb{C}^n$ *with the respective constant functions on* $\mathbb{R}^n$*. For* $t > 0$ *define*

$$\gamma_t := \big(\Theta_t(e_j)\big)_{j=1}^n \in \mathrm{L}^2_{\mathrm{loc}}(\mathbb{R}^n)^n.$$

In order to become more acquainted with the construction, let us prove that the potential approximants $\gamma_t \cdot \mathcal{A}_t F$ for a principal part approximation are uniformly bounded in $\mathrm{L}^2(\mathbb{R}^n)$ with respect to $t > 0$. Here, $F \in \mathrm{L}^2(\mathbb{R}^n)^n$ and, according to our general agreement, the dyadic averaging operators act componentwise on F. A convenient way to compute $\gamma_t \cdot \mathcal{A}_t F$ is as follows. On each fixed dyadic cube $Q \in \square_t$ we have

$$\gamma_t \cdot \mathcal{A}_t F = \sum_{j=1}^n \Theta_t(e_j)\,(F_j)_Q = \Theta_t\bigg(\sum_{j=1}^n e_j\,(F_j)_Q\bigg) = \Theta_t\big(\mathbf{1}_{\mathbb{R}^n}(F)_Q\big)$$

by construction. Taking Q itself as the reference region B in Definition 11.14, we obtain with convergence in $\mathrm{L}^2(Q)$ that

$$\begin{aligned}\gamma_t \cdot \mathcal{A}_t F &= \lim_{k\to\infty} \Theta_t\big(\mathbf{1}_{2^k Q}(F)_Q\big)\\ &= \lim_{k\to\infty} \sum_{\ell=1}^{k-1} \Theta_t\big(\mathbf{1}_{C_\ell(Q)}(F)_Q\big)\\ &= \sum_{\ell=1}^{\infty} \Theta_t\big(\mathbf{1}_{C_\ell(Q)}(F)_Q\big).\end{aligned} \tag{13.6}$$

We will primarily approach the analysis of the right-hand side using off-diagonal estimates. In this context, readers should keep in mind the inequality

$$\frac{\mathrm{dist}(C_\ell(Q), Q)}{t} \geq \frac{(2^\ell - 2)\frac{t}{2}}{t} = \frac{2^\ell - 2}{2},$$

given that the sidelength of Q is at least t.

Lemma 13.5. *For all* $t > 0$ *we have*

$$\|\gamma_t \cdot \mathcal{A}_t F\|_2 \leq C\|F\|_2 \qquad (F \in \mathrm{L}^2(\mathbb{R}^n)^n).$$

Proof. It suffices to prove

$$\|\gamma_t \cdot \mathcal{A}_t F\|_{\mathrm{L}^2(Q)} \leq C\|F\|_{\mathrm{L}^2(Q)} \tag{13.7}$$

for all cubes $Q \in \square_t$. Then the claim follows by summing the squares of both sides with respect to Q. As for (13.7), we start from (13.6) and apply off-diagonal estimates for Θ_t in order to obtain

$$\begin{aligned}\|\gamma_t \cdot \mathcal{A}_t F\|_{\mathrm{L}^2(Q)} &\leq \sum_{\ell=1}^{\infty} \big\|\Theta_t\big(\mathbf{1}_{C_\ell(Q)}(F)_Q\big)\big\|_{\mathrm{L}^2(Q)}\\ &\leq \sum_{\ell=1}^{\infty} C\mathrm{e}^{-c2^\ell}\,\|(F)_Q\|_{\mathrm{L}^2(C_\ell(Q))}.\end{aligned} \tag{13.8}$$

By Hölder's inequality, we have

$$\begin{aligned}\|(F)_Q\|_{\mathrm{L}^2(C_\ell(Q))} &= \frac{|C_\ell(Q)|^{1/2}}{|Q|}\left|\int_Q F \,\mathrm{d}y\right| \\ &\leq \frac{|C_\ell(Q)|^{1/2}}{|Q|^{1/2}}\|F\|_{\mathrm{L}^2(Q)} \\ &\leq 2^{(\ell+1)n/2}\|F\|_{\mathrm{L}^2(Q)},\end{aligned}$$

which, substituted back into (13.8), yields (13.7). □

We continue with a first comparison between Θ_t and its principal part.

Lemma 13.6. *For all $t > 0$ we have*

$$\|(\Theta_t - \gamma_t \cdot \mathcal{A}_t)F\|_2 \leq Ct\|\nabla F\|_2 \qquad (F \in \mathrm{H}^1(\mathbb{R}^n)^n).$$

Proof. We argue in two steps: First we work on a fixed dyadic cube, then we handle the summation over a partition of $\mathbb{R}^n$ into cubes.

Step 1: Estimate on a fixed cube $Q \in \square_t$.

Since Θ_t is a bounded operator on $\mathrm{L}^2(\mathbb{R}^n)^n$, we have

$$\Theta_t F = \sum_{\ell=1}^{\infty} \Theta_t\big(\mathbf{1}_{C_\ell(Q)}F\big).$$

Taking the difference with (13.6) on the same cube and applying off-diagonal estimates as usual, we find

$$\begin{aligned}\|(\Theta_t - \gamma_t \cdot \mathcal{A}_t)F\|_{\mathrm{L}^2(Q)} &\leq \sum_{\ell=1}^{\infty}\big\|\Theta_t\big(\mathbf{1}_{C_\ell(Q)}(F-(F)_Q)\big)\big\|_{\mathrm{L}^2(Q)} \\ &\lesssim \sum_{\ell=1}^{\infty} \mathrm{e}^{-c2^\ell}\big\|F-(F)_Q\big\|_{\mathrm{L}^2(2^{\ell+1}Q)}.\end{aligned}$$

We have generously enlarged the domain of integration from the annulus $C_\ell(Q)$ to the cube $2^{\ell+1}Q$, enabling us to apply the Sobolev–Poincaré inequality from Theorem 10.17 componentwise with $q = 2$ on the interior of $2^{\ell+1}Q$: There is a constant C, depending only on n, such that

$$\begin{aligned}\big\|F-(F)_Q\big\|_{\mathrm{L}^2(2^{\ell+1}Q)} &\leq C\bigg(\frac{\operatorname{diam}(2^{\ell+1}Q)^n}{|Q|}\cdot|2^{\ell+1}Q|^{1/n}\bigg)\|\nabla F\|_{\mathrm{L}^2(2^{\ell+1}Q)} \\ &\leq C\Big((2^{\ell+1}\sqrt{n})^n\cdot 2^{\ell+2}t\Big)\|\nabla F\|_{\mathrm{L}^2(2^{\ell+1}Q)}.\end{aligned} \tag{13.9}$$

This was the key step, leading us to

$$\|(\Theta_t - \gamma_t \cdot \mathcal{A}_t)F\|_{\mathrm{L}^2(Q)} \lesssim t \sum_{\ell=1}^{\infty} 2^{\ell(n+1)} \mathrm{e}^{-c2^\ell} \|\nabla F\|_{\mathrm{L}^2(2^{\ell+1}Q)}. \tag{13.10}$$

Step 2: Summing up in Q.

Since the bound in Step 1 comes with rapid decay with respect to ℓ, we can complete the proof by the same argument that already appeared in Step 3 of the proof of Theorem 11.16. Let us recapitulate the strategy. First, we sum the square of (13.10) over all Q, then we use the Cauchy–Schwarz inequality for the summation in ℓ and finally, we bound the overlap of the enlarged cubes via Exercise 11.4 in order to conclude

$$\begin{aligned}
\|(\Theta_t - \gamma_t \cdot \mathcal{A}_t)F\|_2^2 &\lesssim t^2 \sum_{Q \in \square_t} \Big(\sum_{\ell=1}^{\infty} 2^{\ell(n+1)} \mathrm{e}^{-c2^\ell} \|\nabla F\|_{\mathrm{L}^2(2^{\ell+1}Q)} \Big)^2 \\
&\lesssim t^2 \sum_{Q \in \square_t} \sum_{\ell=1}^{\infty} \mathrm{e}^{-c2^\ell} \|\nabla F\|_{\mathrm{L}^2(2^{\ell+1}Q)}^2 \\
&= t^2 \sum_{\ell=1}^{\infty} \mathrm{e}^{-c2^\ell} \int_{\mathbb{R}^n} \Big(\sum_{Q \in \square_t} \mathbf{1}_{2^{\ell+1}Q} \Big) |\nabla F|^2 \,\mathrm{d}x \\
&= t^2 \sum_{\ell=1}^{\infty} (2^{\ell+1})^n \mathrm{e}^{-c2^\ell} \int_{\mathbb{R}^n} |\nabla F|^2 \,\mathrm{d}x \\
&=: C t^2 \|\nabla F\|_2^2
\end{aligned}$$

with a finite value C for the numerical series. □

Lemma 13.6 is still insufficient for a square function estimate like (13.4). First, we only get decay in t when t is small, which is natural given that the method of proof was to compare F with its averages at scale t. Second, we already have the gradient on F on the right-hand side even though eventually we want to use $F = \nabla u$. The way out here is to apply Lemma 13.5 not directly to $F = \nabla u$ but to a smoothed version $P_t \nabla u$ that is conducive for square function estimates.

Definition 13.7. *For $t > 0$ define bounded operators for the respective L^2-norms by*

$$P_t := (1 - t^2\Delta)^{-1} \quad \textit{and} \quad Q_t := t\nabla(1 - t^2\Delta)^{-1}.$$

Remark 13.8. Letting P_t act componentwise on tuples of functions, the commutation property $tP_t\nabla u = Q_t u$ for all $u \in \mathrm{H}^1(\mathbb{R}^n)$ follows immediately by taking the Fourier transform — both sides are equivalent to multiplication by $\xi \mapsto 2\pi\mathrm{i}t\xi(1 + t^2 m(\xi))^{-1}$, where $m(\xi) = 4\pi^2|\xi|^2$ is as in Lecture 4.

We will need the following three properties.

Lemma 13.9. *For all $F \in \mathrm{L}^2(\mathbb{R}^n)^n$ and $u \in \mathrm{H}^1(\mathbb{R}^n)$, the operators Q_t give rise to:*

(a) The square function estimate

$$\int_0^\infty \|Q_t F\|_2^2 \, \frac{\mathrm{d}t}{t} \lesssim \|F\|_2^2.$$

(b) The adjoint square function estimate

$$\int_0^\infty \|Q_t^* \nabla u\|_2^2 \, \frac{\mathrm{d}t}{t} \lesssim \|\nabla u\|_2^2.$$

(c) The reproducing formula

$$\int_0^\infty 2 Q_t Q_t^* \nabla u \, \frac{\mathrm{d}t}{t} = \nabla u.$$

Proof. (a) We use the Kato property for $-\Delta$ (Proposition 4.12) componentwise to write

$$\big\|Q_t F\big\|_2^2 = \big\|t\nabla(1 - t^2\Delta)^{-1} F\big\|_2^2 = \big\|\sqrt{-t^2\Delta}(1 - t^2\Delta)^{-1} F\big\|_2^2,$$

whereupon the claim follows from quadratic estimates for self-adjoint operators (Lemma 8.3) with auxiliary function $\sqrt{\mathbf{z}}(1 + \mathbf{z})^{-1}$.

(b) According to Lemma 11.10, we have $Q_t^* = -(1 - t^2\Delta)^{-1} t \operatorname{div}$, where we slightly abuse notation and do not distinguish between the unbounded operator in $\mathrm{L}^2(\mathbb{R}^n)$ and the bounded operator on $\mathrm{H}^1(\mathbb{R}^n)$ in case of the negative Laplacian. Hence,

$$\begin{aligned} Q_t^* \nabla u &= (1 - t^2\Delta)^{-1}(-t\Delta)u \\ &= (-t\Delta)(1 - t^2\Delta)^{-1} u. \end{aligned} \tag{13.11}$$

The negative Laplacian has the Kato property (Proposition 4.12) and the claim follows on using ‘(a) $\Longrightarrow$ (b)’ in Proposition 12.7.

(c) Applying Q_t to (13.11) yields

$$Q_t Q_t^* \nabla u = (-t^2\Delta)(1 - t^2\Delta)^{-2} \nabla u;$$

this is probably best seen by commuting the corresponding bounded multiplication operators in Fourier space as in Remark 13.8. Hence, the claim is nothing but the Calderón reproducing formula for the injective sectorial operator $-\Delta$ with auxiliary function $f = \mathbf{z}(1 + \mathbf{z})^{-2}$, but in second order scaling, using $f(-t^2\Delta)$. The different scaling results in a factor 2 on the left-hand side and the proof is complete. □

We assemble the bounds that we have obtained so far in the following smoothed principal part approximation. This only uses the square function estimate in part (a) of Lemma 13.9 and hence works for general n-tuples F of functions.

Proposition 13.10 (Smoothed principal part approximation)**.** *We have the square function estimate*

$$\int_0^\infty \|(\Theta_t - \gamma_t \cdot \mathcal{A}_t)P_t F\|_2^2 \,\frac{\mathrm{d}t}{t} \leq C\|F\|_2^2 \qquad (F \in \mathrm{L}^2(\mathbb{R}^n)^n).$$

Proof. According to Lemma 13.6, we can bound

$$\|(\Theta_t - \gamma_t \cdot \mathcal{A}_t)P_t F\|_2 \lesssim t\|\nabla P_t F\|_2 = \|Q_t F\|_2$$

and we conclude by Lemma 13.9 (a). □

13.3. Principal part approximation

To establish a principal part approximation similar to (13.4), we need yet to remove the smoothing operators P_t in Proposition 13.10. To this end, we split

$$\Theta_t - \gamma_t \cdot \mathcal{A}_t = (\Theta_t - \gamma_t \cdot \mathcal{A}_t)P_t + \Theta_t(1 - P_t) - \gamma_t \cdot \mathcal{A}_t(1 - P_t). \tag{13.12}$$

Each term is a bounded operator acting on functions $F \in \mathrm{L}^2(\mathbb{R}^n)^n$, but eventually we only need to consider $F = \nabla u$ with $u \in \mathrm{H}^1(\mathbb{R}^n)$.

The first term is the one, whose square function is already under control by Proposition 13.10.

The second term in (13.12) is notably easy to handle. In fact, the subsequent proposition relies solely on two ingredients: m-accretivity of L and square function estimates for Q_t^*.

Proposition 13.11. *We have the square function estimate*

$$\int_0^\infty \|\Theta_t(1 - P_t)\nabla u\|_2^2 \,\frac{\mathrm{d}t}{t} \leq C\|\nabla u\|_2^2 \qquad (u \in \mathrm{H}^1(\mathbb{R}^n)).$$

Proof. We compute

$$\begin{aligned}
\Theta_t(1 - P_t)\nabla u &\overset{\text{Rem. 13.8}}{=} (\Theta_t \nabla)(1 - P_t)u \\
&\overset{(13.2)}{=} \big(tL(1 + t^2L)^{-1}\big)\big(-t^2\Delta(1 - t^2\Delta)^{-1}u\big) \\
&= \big(1 - (1 + t^2L)^{-1}\big)\big(-t\Delta(1 - t^2\Delta)^{-1}u\big) \\
&\overset{(13.11)}{=} \big(1 - (1 + t^2L)^{-1}\big)Q_t^*\nabla u.
\end{aligned}$$

Since L is m-accretive, we conclude

$$\|\Theta_t(1-P_t)\nabla u\|_2 \le 2\|Q_t^*\nabla u\|_2$$

and Lemma 13.9 (b) yields the claim. □

The third term in (13.12) is more complicated. However, similar to the second term, we shall rely on rather general harmonic analysis that has almost nothing to do with the principal part γ_t itself. To get to the gist of the matter, we start with the simple observation that $\mathcal{A}_t^2 F = \mathcal{A}_t F$ for every $F \in \mathrm{L}^2(\mathbb{R}^n)^n$ — once a function is constant on each dyadic cube of generation t, further averaging on the same cubes does not change anything anymore. Consequently, we can write

$$\gamma_t \cdot \mathcal{A}_t(1-P_t) = \gamma_t \cdot \mathcal{A}_t(\mathcal{A}_t(1-P_t)). \tag{13.13}$$

As far as the principal part is concerned, we are going to simply use the uniform bound for $\gamma_t \cdot \mathcal{A}_t$ in Lemma 13.5 in order to discard it and concentrate on $\mathcal{A}_t(1-P_t)$.

We will need the following auxiliary result to control dyadic averages of gradient fields.

Lemma 13.12. *There is a constant $C > 0$ such that for all dyadic cubes $Q \in \square$ and all $u \in \mathrm{H}^1(\mathbb{R}^n)$ we have*

$$\left|\fint_Q \nabla u \,\mathrm{d}x\right|^2 \le \frac{C}{\ell(Q)}\left(\fint_Q |u|^2 \,\mathrm{d}x\right)^{1/2}\left(\fint_Q |\nabla u|^2 \,\mathrm{d}x\right)^{1/2}.$$

Proof. We work with a threshold $\alpha \in (0,1)$ that will be chosen as the proof unfolds further and pick a function $\eta \in \mathrm{C}_\mathrm{c}^\infty(Q)$ such that $\eta = 1$ on $(1-\alpha)Q$ and

$$\|1-\eta\|_{\mathrm{L}^\infty(\mathbb{R}^n)} + \alpha\ell(Q)\|\nabla\eta\|_{\mathrm{L}^\infty(\mathbb{R}^n)} \le C \tag{13.14}$$

for some constant C that only depends on n.[1] Let us also write the inequality in question as

$$|X|^2 \le \frac{C}{\ell(Q)} Y^{1/2} Z^{1/2}, \tag{13.15}$$

where X, Y and Z correspond to the averages of ∇u, $|u|^2$ and $|\nabla u|^2$ on Q, respectively. We can assume $Z > 0$ since otherwise we have $\nabla u = 0$ a.e. on Q and there is nothing more to prove.

We begin by splitting $\nabla u = \eta\nabla u + (1-\eta)\nabla u$ and integrate the first term by parts to obtain

$$\fint_Q \nabla u \,\mathrm{d}x = -\fint_Q (\nabla\eta)u \,\mathrm{d}x + \fint_Q (1-\eta)\nabla u \,\mathrm{d}x,$$

[1] This can be achieved by the usual construction: Take $\phi \in \mathrm{C}_\mathrm{c}^\infty(B(0,1))$ with integral 1 and set $\eta := \mathbf{1}_{(1-\alpha)Q} * \phi_\varepsilon$ for $\varepsilon := \alpha\ell(Q)/4$.

where both integrands on the right take non-zero values only on $E := Q \setminus (1-\alpha)Q$. Thus, the uniform bounds (13.14) followed by Hölder's inequality yield

$$\begin{aligned} |X| &\leq \frac{C}{\alpha\ell(Q)} \fint_Q \mathbf{1}_E |u| \,\mathrm{d}x + C \fint_Q \mathbf{1}_E |\nabla u| \\ &\leq \frac{C}{\alpha\ell(Q)} \Big(\frac{|E|}{|Q|}\Big)^{1/2} Y^{1/2} + C \Big(\frac{|E|}{|Q|}\Big)^{1/2} Z^{1/2}. \end{aligned}$$

We square both sides and estimate $\frac{|E|}{|Q|} = (1-(1-\alpha)^n) \leq \alpha n$ by Bernoulli's inequality, to arrive at

$$|X|^2 \leq 2C^2 n \Big(\frac{Y}{\alpha\ell(Q)^2} + \alpha Z\Big).$$

At this stage we want to pick the threshold α such that on the right-hand side both terms in brackets coincide, that is, we fix $\alpha := Y^{1/2}Z^{-1/2}\ell(Q)^{-1}$, leading us directly to (13.15).

In case you did not notice, there is one caveat with this argument: We cannot be sure that our preferred choice of α is smaller than 1! However, in case it is not, we have $Z^{1/2} \leq Y^{1/2}\ell(Q)^{-1}$ and since $|X|^2 \leq Z = Z^{1/2}Z^{1/2}$ by a direct application of the Cauchy–Schwarz inequality, the claim (13.15) follows once again. □

We are in a position to control the final term in (13.12) via the strategy outlined around (13.13).

Proposition 13.13. *We have the square function estimate*

$$\int_0^\infty \|\mathcal{A}_t(1-P_t)\nabla u\|_2^2 \,\frac{\mathrm{d}t}{t} \leq C\|\nabla u\|_2^2 \qquad (u \in \mathrm{H}^1(\mathbb{R}^n)).$$

In particular, we have the same bound for the square function of $\gamma_t \cdot \mathcal{A}_t(1-P_t)\nabla u$.

Proof. The argument comes in two steps.

Step 1: Reduction via a Schur-type bound.

In Step 2 we shall show the bound

$$\|\mathcal{A}_s(1-P_s)Q_t F\|_2 \leq C \min\Big(\frac{s}{t}, \frac{t}{s} + \frac{t^{1/2}}{s^{1/2}}\Big)\|F\|_2 \tag{13.16}$$

for all $F \in \mathrm{L}^2(\mathbb{R}^n)^n$ and all $s, t > 0$. Let us first see, how this implies the claim by a Schur-type bound.

It starts with the reproducing formula

$$\nabla u = 2\int_0^\infty Q_t Q_t^* \nabla u \,\frac{\mathrm{d}t}{t}$$

from Lemma 13.9 (c). The operators $\mathcal{A}_s(1 - P_s)$ are (uniformly) bounded on $\mathrm{L}^2(\mathbb{R}^n)$: this is clear for $(1 - P_s)$ and for $\mathcal{A}_s$ we can use, for instance, the pointwise bound by the Hardy–Littlewood maximal operator in Proposition 9.16, see also Exercise 13.1. Hence,

$$\mathcal{A}_s(1 - P_s)\nabla u = 2\int_0^\infty \mathcal{A}_s(1 - P_s)Q_t Q_t^* \nabla u \,\frac{\mathrm{d}t}{t}$$

and, defining $\zeta \in \mathrm{L}^1((0,\infty); \frac{\mathrm{d}\tau}{\tau})$ by $\zeta(\tau) := \min(\tau, \tau^{-1} + \tau^{-1/2})$, we conclude

$$\begin{aligned}
\int_0^\infty \|\mathcal{A}_s(1 - P_s)\nabla u\|_2^2 \,\frac{\mathrm{d}s}{s} &\leq 4\int_0^\infty \left(\int_0^\infty \|\mathcal{A}_s(1 - P_s)Q_t Q_t^* \nabla u\|_2 \,\frac{\mathrm{d}t}{t}\right)^2 \frac{\mathrm{d}s}{s}\\
&\overset{(13.16)}{\leq} 4C^2 \int_0^\infty \left(\int_0^\infty \zeta(st^{-1})\|Q_t^* \nabla u\|_2 \,\frac{\mathrm{d}t}{t}\right)^2 \frac{\mathrm{d}s}{s}\\
&\overset{(\heartsuit)}{\leq} 4C^2 \left(\int_0^\infty \zeta(\tau)\,\frac{\mathrm{d}\tau}{\tau}\right)^2 \left(\int_0^\infty \|Q_t^* \nabla u\|_2^2 \,\frac{\mathrm{d}t}{t}\right)\\
&\overset{\text{Lem. 13.9 (b)}}{\lesssim} \|\nabla u\|_2^2.
\end{aligned}$$

Here, $(\heartsuit)$ refers to the combination of the Cauchy–Schwarz inequality and Tonelli's theorem from the proof of Lemma 8.7.[2]

Step 2: Proof of (13.16).

Let $F \in \mathrm{L}^2(\mathbb{R}^n)^n$ and $s, t > 0$. Just as $\mathcal{A}_t$, also the operators P_t and Q_t are uniformly bounded on $\mathrm{L}^2(\mathbb{R}^n)^n$ with respect to $t > 0$, see Lemma 11.7. Via the Fourier transform, we have the following unitary equivalences ($\cong$) of bounded operators to multiplication operators by functions of the Fourier variable $\xi \in \mathbb{R}^n$, where we write $m(\xi) = 4\pi^2|\xi|^2$ as usual:

$$\begin{aligned}
P_s Q_t &\cong 2\pi\mathrm{i}t\xi \,\frac{1}{1 + t^2 m(\xi)}\,\frac{1}{1 + s^2 m(\xi)},\\
(1 - P_s)Q_t &\cong 2\pi\mathrm{i}ts^2 m(\xi)\xi \,\frac{1}{1 + t^2 m(\xi)}\,\frac{1}{1 + s^2 m(\xi)},
\end{aligned}$$

from which we can read off the transformation rules

$$P_s Q_t = \frac{t}{s} P_t Q_s \quad \text{and} \quad (1 - P_s)Q_t = \frac{s}{t}(1 - P_t)Q_s.$$

When $s \leq t$, we can obtain (13.16) by the following simple use of uniform L^2-bounds:

$$\begin{aligned}
\|\mathcal{A}_s(1 - P_s)Q_t F\|_2 &\lesssim \|(1 - P_s)Q_t F\|_2\\
&= \frac{s}{t}\|(1 - P_t)Q_s F\|_2\\
&\lesssim \frac{s}{t}\|F\|_2.
\end{aligned}$$

[2]Or, in case you are familiar with it, Young's convolution inequality on the multiplicative group $(0,\infty)$. #DiscussionForum

When $t \leq s$, we can at least bound

$$\begin{aligned} \|\mathcal{A}_s(1-P_s)Q_tF\|_2 &\lesssim \|\mathcal{A}_sQ_tF\|_2 + \|P_sQ_tF\|_2 \\ &= \|\mathcal{A}_sQ_tF\|_2 + \frac{t}{s}\|P_tQ_sF\|_2 \\ &\lesssim \|\mathcal{A}_sQ_tF\|_2 + \frac{t}{s}\|F\|_2. \end{aligned} \tag{13.17}$$

To estimate $\|\mathcal{A}_sQ_tF\|_2$, we recall that by definition of the dyadic averaging operator, the function $\mathcal{A}_sQ_tF$ is constant on each cube $Q \in \square_s$. Hence, we can write (the square of) its $\mathrm{L}^2(\mathbb{R}^n)$-norm as

$$\begin{aligned} \|\mathcal{A}_sQ_tF\|_2^2 &= \sum_{Q\in\square_s} |Q| \left| \fint_Q Q_tF \,\mathrm{d}x \right|^2 \\ &= \sum_{Q\in\square_s} |Q| \left| \fint_Q t\nabla P_tF \,\mathrm{d}x \right|^2 \end{aligned}$$

and Lemma 13.12 controls the averages of gradients by

$$\begin{aligned} &\lesssim \sum_{Q\in\square_s} |Q| \frac{1}{s} \left(\fint_Q |tP_tF|^2 \,\mathrm{d}x \right)^{1/2} \left(\fint_Q |t\nabla P_tF|^2 \,\mathrm{d}x \right)^{1/2} \\ &= \sum_{Q\in\square_s} \frac{t}{s} \|P_tF\|_{\mathrm{L}^2(Q)} \|Q_tF\|_{\mathrm{L}^2(Q)}. \end{aligned}$$

Now, it suffices to apply the Cauchy–Schwarz inequality to the summation in Q in order to get control by

$$\begin{aligned} &\leq \frac{t}{s} \left(\sum_{Q\in\square_s} \|P_tF\|_{\mathrm{L}^2(Q)}^2 \right)^{1/2} \left(\sum_{Q\in\square_s} \|Q_tF\|_{\mathrm{L}^2(Q)}^2 \right)^{1/2} \\ &= \frac{t}{s}\|P_tF\|_2\|Q_tF\|_2 \\ &\lesssim \frac{t}{s}\|F\|_2^2. \end{aligned}$$

Together with (13.17), this completes the proof of (13.16) also in the case $t \leq s$. □

Remark 13.14. Since $\mathcal{A}_t(1-P_t) = \mathcal{A}_t(\mathcal{A}_t - P_t)$, we could go one step further and ask for square function bounds involving only $\mathcal{A}_t - P_t$. While this is true, more classical, and even renders much more clearly the idea that the third term in (13.12) is about comparing different types of averages, its proof nevertheless seems to require harder arguments and explicit kernel computations, compare with [AT98, App. C]. We have learned the trick to stick with $\mathcal{A}_t(1-P_t)$ from Axelsson(Rosén)–Keith–McIntosh [AKM06].

Combining Proposition 13.10, 13.11 and 13.13, we control square functions of all three terms in (13.12) applied to ∇u. The result is the following principal part approximation.

Proposition 13.15 (Principal part approximation)**.** *We have the square function estimate*

$$\int_0^\infty \|(\Theta_t - \gamma_t \cdot \mathcal{A}_t)\nabla u\|_2^2 \, \frac{\mathrm{d}t}{t} \leq C\|\nabla u\|_2^2 \qquad (u \in \mathrm{H}^1(\mathbb{R}^n)).$$

This should be considered as half the way toward the proof of the Kato conjecture and here is what remains to be done.

Corollary 13.16 (Reduction to a Carleson measure estimate)**.** *In order to prove Theorem 12.1, it is necessary and sufficient to prove that* $|\gamma_t(x)|^2 \, \frac{\mathrm{d}x \, \mathrm{d}t}{t}$ *is a Carleson measure for every elliptic operator in divergence form on* $\mathrm{L}^2(\mathbb{R}^n)$.

Proof. Thanks to the principal approximation, the square function bound (13.1) at the beginning of the lecture is now equivalent to proving for all $u \in \mathrm{H}^1(\mathbb{R}^n)$ that

$$\int_0^\infty \int_{\mathbb{R}^n} |\gamma_t(x) \cdot (\mathcal{A}_t \nabla u)(x)|^2 \, \frac{\mathrm{d}x \, \mathrm{d}t}{t} \lesssim \|\nabla u\|_2^2.$$

The Cauchy–Schwarz inequality yields $|\gamma_t(x) \cdot (\mathcal{A}_t \nabla u)(x)|^2 \leq |(\mathcal{A}_t \nabla u)(x)|^2 |\gamma_t(x)|^2$, so sufficiency follows from Carleson's lemma (applied componentwise).

Proving necessity is left to you as Exercise 13.3, as for the Kato conjecture, it functions more as a (nice) gimmick. □

To be continued . . .

13.4. Exercises

Exercise 13.1 (L^2-bounds for the dyadic averaging operators)**.** Let $F \in \mathrm{L}^2(\mathbb{R}^n)^n$ and $t > 0$. Prove the inequality

$$\|\mathcal{A}_t F\|_2 \leq \|F\|_2.$$

Does that mean that the dyadic maximal operator $\mathcal{A}$ does not increase L^2-norms?

Exercise 13.2 (Smooth vs. rough averages)**.** We denote the rough averages at scale $t > 0$ of a given function $F \in \mathrm{L}^2(\mathbb{R}^n)^n$ by

$$(\mathcal{M}_t F)(x) := \fint_{B(x,t)} F \, \mathrm{d}y,$$

compare with Lecture 9. Square functions involving $\mathcal{M}_t$, as opposed to $\mathcal{A}_t$, are easier to investigate thanks to the convolution structure of $\mathcal{M}_t$. In this exercise, we ask you to prove the bound

$$\int_0^\infty \|(\mathcal{M}_t - P_t)F\|_2^2 \, \frac{\mathrm{d}t}{t} \lesssim \|F\|_2^2 \qquad (F \in \mathrm{L}^2(\mathbb{R}^n)^n),$$

compare with Remark 13.14.

Hint: Carefully investigate the Fourier transform of the characteristic function of the unit ball.

Exercise 13.3 (Necessity of the Carleson estimate)**.** Prove necessity of the Carleson condition as stated in Corollary 13.16.

Hint: Read again the paragraph preceding Theorem 9.19 before you start.

Exercise 13.4 (Dyadically bootstrapping Poincaré)**.** Let $Q \subseteq \mathbb{R}^n$ be a cube of sidelength t, let $\ell \in \mathbb{N}$ and $u \in \mathrm{H}^1(\mathbb{R}^n)$. In (13.9) we have seen, using a slightly different nomenclature, the Poincaré inequality

$$\|u - (u)_Q\|_{\mathrm{L}^2(2^\ell Q)} \leq C\Phi(\ell)t\|\nabla u\|_{\mathrm{L}^2(2^\ell Q)},$$

where C depends only on n and $\Phi(\ell) = 2^{\ell(n+1)}$.

While this was sufficient for our purpose, the dependence on ℓ can be improved:

$$\Phi(\ell) = \begin{cases} 2^{\ell n/2} & \text{if } n \geq 3, \\ \ell 2^\ell & \text{if } n = 2, \\ 2^\ell & \text{if } n = 1. \end{cases}$$

Prove it!

Exercise 13.5 (Endpoint extrapolation in dimension $n = 1$)**.** In this exercise, we work in dimension $n = 1$. In Theorem 11.16 and Exercise 11.6 you have seen an L^p-extrapolation for the resolvents: For every $p \in (1, \infty)$ there is a constant C such that

$$\|(1 + t^2 L)^{-1} u\|_{\mathrm{L}^p(\mathbb{R})} \leq C\|u\|_{\mathrm{L}^p(\mathbb{R})} \tag{13.18}$$

for all $t > 0$ and all $u \in \mathrm{L}^p(\mathbb{R}) \cap \mathrm{L}^2(\mathbb{R})$. The goal of this exercise is to refine the strategy of proof in order to extend the result to the endpoints $p = 1$ and $p = \infty$.

We proceed as follows. The constants C, c are only allowed to depend on λ and Λ.

(a) For $t > 0$ and $u \in \mathrm{L}^2(\mathbb{R})$ prove the L^2–L^∞-bound

$$\|(1 + t^2 L)^{-1} u\|_{\mathrm{L}^\infty(\mathbb{R})} \leq Ct^{-1/2}\|u\|_{\mathrm{L}^2(\mathbb{R})}.$$

(b) Let $t > 0$ and $E, F \subseteq \mathbb{R}$ be measurable and set $d \coloneqq \operatorname{dist}(E, F)$. Upgrade part (a) to the L^2–L^∞-off-diagonal estimate

$$\|(1 + t^2 L)^{-1} u\|_{\mathrm{L}^\infty(F)} \leq C t^{-1/2} \mathrm{e}^{-c \frac{d}{t}} \|u\|_{\mathrm{L}^2(E)}$$

for $u \in \mathrm{L}^2(\mathbb{R})$ with $\operatorname{supp}(u) \subseteq E$.

Hint: In the case $d/t \geq 1$, try to use a bounded function $\rho \in \mathrm{C}^\infty(\mathbb{R}^n)$ with $\operatorname{dist}(E, \operatorname{supp}(\rho)) \geq d/2$. Of course, you also have to construct it ;-)

(c) Mimic a familiar argument to prove (13.18) for $p = \infty$.

(d) Complete the proof by treating the case $p = 1$.

14. The solution of the Kato conjecture: Part II

As the grande finale of our course, we are going to complete the proof of the Kato conjecture. Notation is as before and we shall make the following conventions.

Notation 14.1. The symbols C, c are reserved for positive constants that depend only on n, λ and Λ. The Carleson box over a dyadic cube $Q \in \square$ is denoted by $R(Q) := Q \times (0, \ell(Q)]$.

In Corollary 13.16 of the foregoing lecture, we have reduced the Kato conjecture to proving that $|\gamma_t(x)|^2 \frac{\mathrm{d}x\,\mathrm{d}t}{t}$ is a Carleson measure. Let us draw a parallel with the theory of singular integral operators, while simultaneously explaining the terminology '$T(1)$ and $T(b)$ theorem' in Section 13.1. The reduction step in Corollary 13.16 was on the square function estimate

$$\int_0^\infty \|\Theta_t \nabla u\|_2^2 \frac{\mathrm{d}t}{t} \lesssim \|\nabla u\|_2^2,$$

which, from Lecture 12, was already known to be equivalent to the Kato conjecture. Instead of bounding the left-hand side for all $F = \nabla u$, we have therefore reduced the matter to investigating $\Theta_t F$ for very specific functions, namely $F = e_1, \ldots, e_n$ corresponding to $\gamma_t = (\Theta_t e_j)_{j=1}^n$ and, in particular, $F = 1$ in dimension $n = 1$. Reducing the L^2-boundedness of an operator T to a property of $T(1)$, the suitably defined application of T to the constant-1-function, is called a '$T(1)$ theorem' ever since the famous result of David–Journé [DJ84] for singular integral operators. We refer to [AGP18] for a historical account. However, calculating $T(1)$ might just be impossible. This is what happens with our Θ_t.

The idea behind a '$T(b)$ theorem' is to replace the constant-1-function by other, tailor-suited test functions b for T. For singular integral operators, such criteria first appeared in the work of David–Journé–Semmes [DJS85]. In our case, and in analogy with the '$T(b)$ theorem for square functions' in [Sem90], we shall design test functions b allowing us to verify the '$T(1)$ condition' for Θ_t, namely that $|\gamma_t(x)|^2 \frac{\mathrm{d}x\,\mathrm{d}t}{t}$ is a Carleson measure.

14.1. A $T(b)$ strategy for the Kato conjecture

Let us explain a common strategy employed in a $T(b)$ argument through the proof of the Carleson measure estimate in dimension $n = 1$.

Proposition 14.2. *In dimension $n = 1$, we have the Carleson measure estimate*

$$\iint_{R(Q)} |\gamma_t(x)|^2 \frac{\mathrm{d}x\,\mathrm{d}t}{t} \le C|Q| \qquad (Q \in \square).$$

We need a simple, yet important observation that is already specific to the one-dimensional setting. We leave the proof for you as Exercise 14.1.

Lemma 14.3. *In dimension $n = 1$, the principal part approximation in Proposition 13.15 remains valid for all $F \in \mathrm{L}^2(\mathbb{R})$ rather than $F = \nabla u$ with $u \in \mathrm{H}^1(\mathbb{R})$.*

Proof of Proposition 14.2. We fix $Q \in \square$ and a cut-off function $\eta_Q \in \mathrm{C}_c^\infty(2Q)$ with the usual requirements that $\eta = 1$ on Q and $\|\eta_Q\|_\infty + \ell(Q)\|\nabla\eta_Q\|_\infty \le C$.[1] Keeping in mind that in dimension $n = 1$ the coefficients A are a function $A\colon \mathbb{R} \to \mathbb{C}$ with $|A| \le \Lambda$ and $\mathrm{Re}(A) \ge \lambda$ almost everywhere, we define the '$T(b)$-type test function'

$$b_Q \coloneqq A^{-1}\eta_Q. \tag{14.1}$$

The rest of the proof comes in five steps.

Step 1: Accretivity estimate for the test function.

We claim the fundamental estimate

$$\mathrm{Re}\big((\mathcal{A}_t b_Q)(x)\big) \ge \frac{\lambda}{\Lambda^2} \qquad ((x,t) \in R_Q). \tag{14.2}$$

To this end, let $Q' \in \square_t$ be the unique dyadic cube of generation t that contains x. Now, $(x,t) \in R_Q$ precisely means that $t \le \ell(Q)$ and $x \in Q$. Consequently, we have $Q' \subseteq Q$ and

$$\mathrm{Re}(b_Q) = \mathrm{Re}(A^{-1}) = \frac{\mathrm{Re}(\overline{A})}{|A|^2} \ge \frac{\lambda}{\Lambda^2},$$

almost everywhere on Q'. Averaging both sides over Q' yields (14.2).

Step 2: Back to the dyadic averaging operator.

[1] If you are not yet familiar with the construction, flip back to the proof of Theorem 11.8.

We make use of (14.2) by bounding

$$\frac{\lambda^2}{\Lambda^4}\iint_{R(Q)} |\gamma_t(x)|^2 \,\frac{\mathrm{d}x\,\mathrm{d}t}{t} \leq \iint_{R(Q)} |\gamma_t(x)\cdot(\mathcal{A}_t b_Q)(x)|^2 \,\frac{\mathrm{d}x\,\mathrm{d}t}{t}. \tag{14.3}$$

Now that we have re-introduced the dyadic averaging operator, we use the principal part approximation, backwards in a way, and write

$$\gamma_t \cdot \mathcal{A}_t = (\gamma_t \cdot \mathcal{A}_t - \Theta_t) + \Theta_t.$$

Generously enlarging the domain of integration on the right-hand side of (14.3), we find ourselves left with

$$\begin{aligned}\frac{\lambda^2}{2\Lambda^4}\iint_{R(Q)} |\gamma_t(x)|^2 \,\frac{\mathrm{d}x\,\mathrm{d}t}{t} &\leq \int_0^{\ell(Q)} \|(\gamma_t\cdot\mathcal{A}_t - \Theta_t)b_Q\|_2^2 + \|\Theta_t b_Q\|_2^2 \,\frac{\mathrm{d}t}{t} \\ &\leq C\|b_Q\|_2^2 + \int_0^{\ell(Q)} \|\Theta_t b_Q\|_2^2 \,\frac{\mathrm{d}t}{t},\end{aligned} \tag{14.4}$$

where in the second step we have used the principal part approximation with $F = b_Q$ in its extended version stated in Lemma 14.3.

Step 3: L^2*-bound for the test function.*

The first term on the right in (14.4) is easy to control: By definition of b_Q we have

$$\|b_Q\|_2^2 \leq \|A^{-1}\|_\infty^2 \|\eta_Q\|_\infty^2 |2Q| \leq \frac{2C^2}{\lambda^2}|Q|. \tag{14.5}$$

Step 4: Truncated square function bound for the test function.

The whole argument hinges on the possibility to manually bound the (truncated) square function for the special function b_Q on the right of (14.4). To this end, we write out the definition of Θ_t (see Definition 13.2) and b_Q to find

$$\begin{aligned}\Theta_t b_Q &= -t(1+t^2\mathscr{L})^{-1}\operatorname{div}(Ab_Q) \\ &= -t(1+t^2L)^{-1}(\operatorname{div}\eta_Q).\end{aligned}$$

We have incorporated A^{-1} in the definition of b_Q so that the coefficients cancel in this computation. Here, $\operatorname{div}\eta_Q = (\eta_Q)'$ is a smooth function that is bounded by $C/\ell(Q)$ and has support in $2Q$. Controlling the resolvents of L simply by m-accretivity, we obtain

$$\|\Theta_t b_Q\|_2^2 \leq \|t \operatorname{div}\eta_Q\|_2^2 \leq t^2\frac{2C^2}{\ell(Q)^2}|Q|,$$

which is still enough to conclude the desired estimate

$$\int_0^{\ell(Q)} \|\Theta_t b_Q\|_2^2 \,\frac{\mathrm{d}t}{t} \leq \int_0^{\ell(Q)} t\frac{2C^2}{\ell(Q)^2}|Q| \,\mathrm{d}t = C^2|Q|. \tag{14.6}$$

Step 5: Conclusion.

We assemble (14.4), (14.5) and (14.6) to give the Carleson measure estimate

$$\iint_{R(Q)} |\gamma_t(x)|^2 \, \frac{\mathrm{d}x \, \mathrm{d}t}{t} \leq C|Q|,$$

with a possibly different constant C. $\square$

Together, Proposition 14.2 and Corollary 13.16 solve the Kato conjecture in dimension $n = 1$! This is due to Coifman–McIntosh–Meyer [CMM82] by a different, yet closely related argument. The proof provided above also foreshadows a potential strategy in all dimensions and, at the same time, it underscores the major difficulties that make the Kato conjecture so much harder in higher dimension. To this end, let us wrap up the essential features of the '$T(b)$-type test functions' b_Q associated with $Q \in \square$:

(a) Uniform L^2-bounds in (14.5).

(b) Pointwise control of $|\gamma_t|$ by $|\gamma_t \cdot \mathcal{A}_t b_Q|$ on $R(Q)$ in (14.3).

(c) Manual control over the truncated square function with $\Theta_t b_Q$ in (14.6) by adapting b_Q to L, and using the principal part approximation for $F = b_Q$ in (14.4).

In higher dimensions, (c) would necessitate constructing b_Q in the form of a gradient, but the pointwise control in (b) represents the most significant obstacle. Indeed, given $(x, t) \in R(Q)$, the $\mathbb{C}^n$-vectors $w := \gamma_t(x)$ and $\overline{\xi} := (\mathcal{A}_t b_Q)(x)$ should satisfy

$$c|w| \leq |w \cdot \overline{\xi}|. \tag{14.7}$$

In dimension $n = 1$, there is no orthogonality and (14.7) is the same as saying $|\overline{\xi}| \geq c$, which in turn followed from the pointwise bound $\mathrm{Re}(b_Q) \geq c$ on Q. However, in higher dimensions, (14.7) indicates that ξ must roughly align with the direction of w, about which we have limited information. In fact, we are trying to use a $T(b)$ argument precisely because we seem unable to compute $\gamma_t(x)$ accurately.

14.2. The sectorial decomposition of $\mathbb{C}^n$

A key idea in the resolution of the Kato conjecture in arbitrary dimensions was to use a sectorial decomposition of $\mathbb{C}^n$ to enforce the principal part on $R(Q)$ to lie within small cones with central axis $w = \xi$ and then construct test functions on Q that are close to $\overline{\xi}$ in average.

Definition 14.4. *Let $\varepsilon > 0$. For $\xi \in \mathbb{C}^n$ with $|\xi| = 1$, define the open cone with central axis ξ by*

$$\Gamma_\xi^\varepsilon := \left\{ w \in \mathbb{C}^n \setminus \{0\} \mid \left| \tfrac{w}{|w|} - \xi \right| \leq \varepsilon \right\}.$$

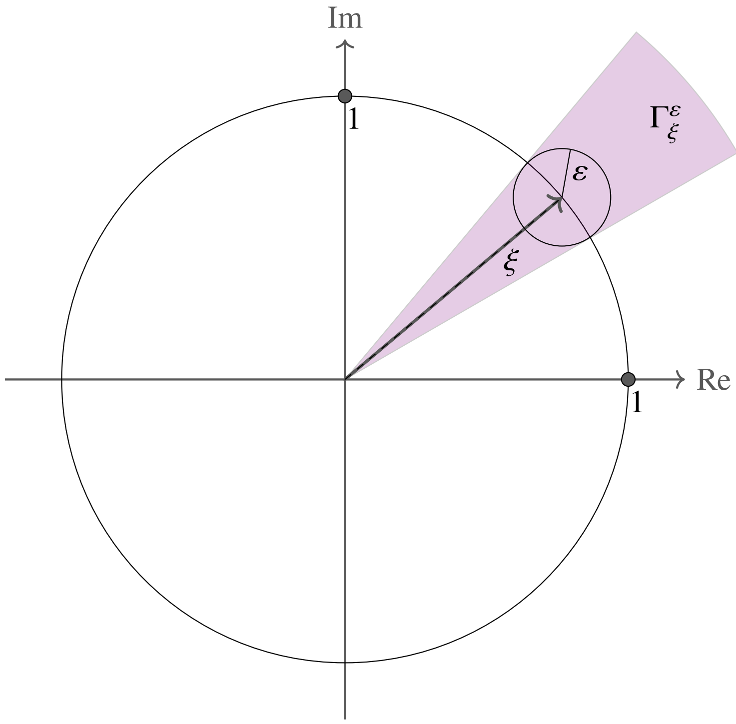

Figure 14.1.: Schematic representation of a cone $\Gamma_\xi^\varepsilon \subseteq \mathbb{C}^n \setminus \{0\}$ in dimension $n = 1$.

Our goal is then to prove the following Carleson measure estimate, where we restrict the values of the principal part to a cone Γ_ξ^ε.

Proposition 14.5 (Key estimate)**.** *There is a choice of $\varepsilon > 0$, depending only on n, λ and Λ, such that for every unit vector $\xi \in \mathbb{C}^n$ we have the Carleson measure estimate*

$$\iint_{R(Q)} |\gamma_t(x)\, \mathbf{1}_{\Gamma_\xi^\varepsilon}(\gamma_t(x))|^2 \,\frac{\mathrm{d}x\,\mathrm{d}t}{t} \leq C|Q| \qquad (Q \in \square).$$

The key estimate implies that $|\gamma_t(x)|^2 \frac{\mathrm{d}x\,\mathrm{d}t}{t}$ is a Carleson measure and hence concludes the proof of the Kato conjecture by virtue of Corollary 13.16!

Indeed, by compactness of the unit sphere, we can cover $\mathbb{C}^n \setminus \{0\}$ by a finite number of cones $\Gamma_{\xi_1}^\varepsilon, \ldots, \Gamma_{\xi_N}^\varepsilon$, where N depends only on n and ε, and we obtain

$$\iint_{R(Q)} |\gamma_t(x)|^2 \,\frac{\mathrm{d}x\,\mathrm{d}t}{t} \leq \iint_{R(Q)} \sum_{j=1}^N |\gamma_t(x)\, \mathbf{1}_{\Gamma_{\xi_j}^\varepsilon}(\gamma_t(x))|^2 \,\frac{\mathrm{d}x\,\mathrm{d}t}{t} \leq CN|Q|$$

for every cube $Q \in \square$ as required.

Hence, all that remains is to prove the key estimate. From now on, the unit vector ξ is fixed and we write $\Gamma^\varepsilon := \Gamma_\xi^\varepsilon$. We shall establish the claim of Proposition 14.5 by refining the strategy outlined in Section 14.1.

14.3. Construction of $T(b)$-type test functions

We will use the freedom of being able to pick ε small twice in order to steer the values of our test functions b_Q into the direction $\overline{\xi}$. The following proposition contains the construction of b_Q and adjusts the direction in average on the 'parent cube' Q itself.

Proposition 14.6. *There is a constant $\varepsilon_0 \in (0, 1]$ such that if $\varepsilon \leq \varepsilon_0$, then for each cube $Q \in \square$ we can construct a '$T(b)$-type test function' b_Q^ε with the following properties:*

(a) $\|b_Q^\varepsilon\|_2 \leq C|Q|^{1/2}$,

(b) $\operatorname{Re}\left(\xi \cdot \fint_Q b_Q^\varepsilon \, \mathrm{d}x\right) \geq 1$,

(c) $\displaystyle\iint_{R(Q)} |\gamma_t(x) \cdot (\mathcal{A}_t b_Q^\varepsilon)(x)|^2 \, \frac{\mathrm{d}x \, \mathrm{d}t}{t} \leq \frac{C}{\varepsilon^2}|Q|$.

Note that (a), (b) and (c) in Proposition 14.6 roughly align with their counterparts in Section 14.1, which share analogous labels. Notably, in (c) above, we refrain from explicitly stating the square function bound; however, we will leverage the manual control over $\Theta_t b_Q^\varepsilon$ in the proof.

We need the following elementary lemma on a family associated with L that appears for the first time.

Lemma 14.7. *For all $F \in \mathrm{L}^2(\mathbb{R}^n)^n$ and all $t > 0$ we have the uniform bound*

$$\|t^2\nabla(1 + t^2\mathscr{L})^{-1} \operatorname{div} F\|_2 \leq \frac{1}{\lambda}\|F\|_2.$$

Proof. The proof is similar to the one of Lemma 11.7. We set $w := (1 + t^2\mathscr{L})^{-1} \operatorname{div} F$ and note that by ellipticity of A we have

$$\lambda t^2\|\nabla w\|_2^2 \leq t^2 \operatorname{Re}\big(\langle \mathscr{L}w, w\rangle\big) \leq \operatorname{Re}\big(\langle (1 + t^2\mathscr{L})w, w\rangle\big) = \operatorname{Re}\big(\langle \operatorname{div} F, w\rangle\big),$$

where the angular brackets denote the $\mathrm{H}^{-1}(\mathbb{R}^n)$–$\mathrm{H}^1(\mathbb{R}^n)$-duality. By definition of the very weak divergence operator and the Cauchy–Schwarz inequality, we conclude

$$\lambda t^2\|\nabla w\|_2^2 \leq \|F\|_2\|\nabla w\|_2$$

and the claim follows. □

Proof of Proposition 14.6. We fix $Q \in \square$ and abbreviate $\ell := \ell(Q)$. Similarly, we simplify notation by omitting ε and Q when constructing the test function $b = b_Q^\varepsilon$.

We start by fixing $\eta \in \mathrm{C}_c^\infty(2Q)$ such that $\eta = 1$ on Q and $\|\eta\|_\infty + \ell\|\nabla\eta\|_\infty \leq C$ as in the proof of Proposition 14.2. With x_Q the center of Q, we introduce a smooth function with compact support, whose gradient is equal to $\overline{\xi}$ on Q by

$$\Phi(x) := \eta(x)(x - x_Q) \cdot \overline{\xi}, \tag{14.8}$$

and finally define

$$b := 2\nabla(1 + \varepsilon^2\ell^2 L)^{-1}\Phi. \tag{14.9}$$

Then, we have

$$\begin{aligned} \frac{1}{2}b - \nabla\Phi &= \nabla\big((1 + \varepsilon^2\ell^2 L)^{-1} - 1\big)\Phi \\ &= \nabla\big(-(1 + \varepsilon^2\ell^2 \mathscr{L})^{-1}\varepsilon^2\ell^2\mathscr{L}\big)\Phi \\ &= \varepsilon^2\ell^2\nabla(1 + \varepsilon^2\ell^2\mathscr{L})^{-1}\operatorname{div}(A\nabla\Phi), \end{aligned} \tag{14.10}$$

which conveys the idea that b is a gradient function adapted to L that has a chance of pointing in direction $\overline{\xi}$ on Q, compare with (b) and (c) in Section 14.1. Let us prove that we can pick ε small enough such that b has the stated properties.

(a) We begin by calculating

$$\begin{aligned} |\nabla\Phi(x)| &\leq |\nabla\eta(x)(x - x_Q) \cdot \overline{\xi}| + |\eta(x)\overline{\xi}| \\ &\leq \left(\frac{C}{\ell}\frac{\sqrt{n}\ell}{2} + C\right)\mathbf{1}_{2Q}(x). \end{aligned}$$

Hence, we obtain with a different constant C that

$$\|\nabla\Phi\|_2^2 \leq C|Q|. \tag{14.11}$$

Combining (14.10) and Lemma 14.7, we find

$$\|b - 2\nabla\Phi\|_2^2 \leq \frac{1}{\lambda^2}\|2A\nabla\Phi\|_2^2 \leq \frac{4C\Lambda^2}{\lambda^2}|Q|$$

and together, these two norm estimates imply (a).

(b) As $\nabla\Phi = \overline{\xi}$ on Q, we can write

$$⨍_Q b - 2\overline{\xi}\,\mathrm{d}x = ⨍_Q b - 2\nabla\Phi\,\mathrm{d}x =: 2\varepsilon^2\ell^2 ⨍_Q \nabla u\,\mathrm{d}x, \tag{14.12}$$

where in the second step we have used (14.10) and $u := (1+\varepsilon^2\ell^2\mathscr{L})^{-1}\operatorname{div}(A\nabla\Phi)$. The uniform L^2-bounds in Lemmas 11.7[2] and 14.7 in combination with (14.11) yield

$$\|u\|_2 \leq \frac{C}{\varepsilon\ell}|Q|^{1/2} \quad \text{and} \quad \|\nabla u\|_2 \leq \frac{C}{\varepsilon^2\ell^2}|Q|^{1/2},$$

[2]To be precise, the dual estimate to Lemma 11.7 (b), compare with Lemma 11.10.

so that estimating the average on the right-hand side of (14.12) by means of Lemma 13.12 leads us to

$$\left| \fint_Q b - 2\overline{\xi} \, \mathrm{d}x \right|^2 \le \frac{C\varepsilon^4 \ell^4}{\ell |Q|} \|u\|_2 \|\nabla u\|_2 \le C\varepsilon,$$

where C varies from step to step. Since $\xi \in \mathbb{C}^n$ is a unit vector, we conclude that

$$\mathrm{Re}\Big(\xi \cdot \fint_Q b \, \mathrm{d}x\Big) = 2 + \mathrm{Re}\Big(\xi \cdot \fint_Q b - 2\overline{\xi} \, \mathrm{d}x\Big) \ge 2 - \sqrt{C\varepsilon}.$$

Now, (b) follows upon taking $\varepsilon \le C^{-1}$.

(c) As in Step 2 in the proof of Proposition 14.2, we use the principal part approximation from Proposition 13.15 with $F = b$ backwards and bound

$$\begin{aligned} \frac{1}{2} \iint_{R(Q)} & |\gamma_t(x) \cdot (\mathcal{A}_t b)(x)|^2 \, \frac{\mathrm{d}x \, \mathrm{d}t}{t} \\ &\le \int_0^\ell \|(\gamma_t \cdot \mathcal{A}_t - \Theta_t) b\|_2^2 + \|\Theta_t b\|_2^2 \, \frac{\mathrm{d}t}{t} \\ &\le C\|b\|_2^2 + \int_0^\ell \|\Theta_t b\|_2^2 \, \frac{\mathrm{d}t}{t} \\ &\overset{\text{(a)}}{\le} C|Q| + \int_0^\ell \|\Theta_t b\|_2^2 \, \frac{\mathrm{d}t}{t}. \end{aligned} \tag{14.13}$$

Recall from (14.9) that b is indeed the gradient of a function in $\mathrm{H}^1(\mathbb{R}^n)$ and hence applying Proposition 13.15 was legitimate. Once again, all hinges on being able to compute $\Theta_t b$. By definition, we have

$$\begin{aligned} \Theta_t b &= -2t(1 + t^2 \mathscr{L})^{-1} \operatorname{div}\big(A\nabla(1 + \varepsilon^2 \ell^2 L)^{-1} \Phi\big) \\ &= 2t(1 + t^2 \mathscr{L})^{-1} \mathscr{L} (1 + \varepsilon^2 \ell^2 \mathscr{L})^{-1} \Phi \\ &= -t\big((1 + t^2 L)^{-1}\big)\big((1 + \varepsilon^2 \ell^2 \mathscr{L})^{-1} \operatorname{div}\big) 2A\nabla\Phi. \end{aligned}$$

Compared to Step 4 in the proof of Proposition 14.2, the coefficients A did not cancel but could be used to recombine the full operator $\mathscr{L}$ in the second line. It suffices to use uniform L^2-bounds (as in the proof of (b)) and (14.11) in order to control

$$\|\Theta_t b\|_2^2 \le \frac{2t^2}{\lambda \varepsilon^2 \ell^2} \|2A\nabla\Phi\|_2^2 \le \frac{8C\Lambda^2 t^2}{\lambda \varepsilon^2 \ell^2} |Q|.$$

Integration in t yields

$$\int_0^\ell \|\Theta_t b\|_2^2 \, \frac{\mathrm{d}t}{t} \le \frac{4C\Lambda^2}{\lambda \varepsilon^2} |Q|,$$

which we use back in (14.13) to conclude provided we take $\varepsilon \le 1$. □

14.4. Conclusion by a stopping time construction

The directional estimate in Proposition 14.6 (b) is not yet enough to implement the strategy outlined in Section 14.1, because we would need $(\mathcal{A}_t b_Q^\varepsilon)(x)$ to point roughly in direction $\overline{\xi}$ for $x \in Q$ and all $t \le \ell(Q)$. However, if $t \le {}^{\ell(Q)}/{}_2$, then we have

$$(\mathcal{A}_t b_Q^\varepsilon)(x) = \fint_{Q'} b_Q^\varepsilon \, \mathrm{d}y$$

for a *dyadic child* Q' of Q, that is, a proper dyadic subcube of Q. In dimension $n = 1$, all came automatically, since we had a pointwise lower bound for b_Q on the entire parent cube Q. Now, we only control the direction of the average on the parent cube. The next lemma asserts that a substantial number of dyadic children inherit this property, provided we choose ε to be even smaller.

It is at this point that we make essential use of the dyadic cube structure — the good children will be selected by maximality, or equivalently, a stopping time. The constant ε_0 below is the one from Proposition 14.6.

Lemma 14.8. *There is a choice of $\varepsilon \le \varepsilon_0$, depending only on n, λ and Λ, such that each dyadic cube $Q \in \square$ has pairwise disjoint dyadic children Q_j for which the sets*

$$E(Q) \coloneqq Q \setminus \bigcup_j Q_j \quad \textit{and} \quad E^*(Q) \coloneqq R(Q) \setminus \bigcup_j R(Q_j) \tag{14.14}$$

have the following properties:

(a) $\bigl|E(Q)\bigr| \ge \eta |Q|$, *for some $\eta > 0$ depending only on n, λ and Λ.*

(b) $\bigl|w \cdot (\mathcal{A}_t b_Q^\varepsilon)(x)\bigr| \ge \frac{1}{2}|w|$, *whenever $(x,t) \in E^*(Q)$ and $w \in \Gamma^\varepsilon$.*

Proof. Let $0 < \varepsilon \le \varepsilon_0$. Given Q, we let $(Q_j)_j$ be the family of maximal dyadic cubes $Q' \subseteq Q$ for which at least one of the following properties fails:

$$\fint_{Q'} |b_Q^\varepsilon| \, \mathrm{d}y \le \frac{1}{4\varepsilon}, \tag{14.15}$$

$$\operatorname{Re}\Bigl(\xi \cdot \fint_{Q'} b_Q^\varepsilon \, \mathrm{d}y\Bigr) \ge \frac{3}{4}. \tag{14.16}$$

The Q_j are pairwise disjoint by maximality. Let us check that they have the desired properties.

(b) Let $(x,t) \in E^*(Q)$ and $w \in \Gamma^\varepsilon$. For brevity, we set $v \coloneqq (\mathcal{A}_t b_Q^\varepsilon)(x) = \fint_{Q'} b_Q^\varepsilon \, \mathrm{d}y$, where by definition Q' is the unique dyadic cube of generation t that contains x.

We claim that Q' satisfies both (14.15) and (14.16). Suppose to the contrary that it does not. Then, by maximality, we would have $Q' \subseteq Q_j$ for some j.

Consequently, $x \in Q_j$ and $t \leq \ell(Q') \leq \ell(Q_j)$, but this would imply $(x,t) \in R(Q_j)$ in contradiction with the definition of $E^*(Q)$.

From (14.15) and (14.16) we obtain $|v| \leq 1/(4\varepsilon)$ and $\mathrm{Re}(\xi \cdot v) \geq 3/4$. Since ξ is the central axis of Γ^ε, we conclude

$$\left|\frac{w}{|w|} \cdot v\right| \geq |\xi \cdot v| - \left|\frac{w}{|w|} - \xi\right||v| \geq \frac{3}{4} - \varepsilon\frac{1}{4\varepsilon} = \frac{1}{2},$$

that is, $|w \cdot v| \geq |w|/2$.

(a) We will use the bounds for b_Q^ε on the parent cube in Proposition 14.6 to keep the measure of the cubes Q_j under control. Since eventually (a) will hold with $\eta > 0$, we automatically see that the parent cube is not one of the Q_j.

To this end, let $(Q_j)_{j\in J_1}$ and $(Q_j)_{j\in J_2}$ be the collections of those cubes Q_j, for which (14.15) and (14.16) fail, respectively. Let $F_1(Q)$ and $F_2(Q)$ denote their unions. Then $E(Q) = Q \setminus (F_1(Q) \cup F_2(Q))$ and, consequently, it suffices to prove the bound

$$|F_1(Q)| + |F_2(Q)| \leq (1-\eta)|Q| \tag{14.17}$$

for a suitable $\eta > 0$, using the choice of ε of course.

Since (14.15) fails for every cube in $(Q_j)_{j\in J_1}$, we obtain

$$|F_1(Q)| = \sum_{j\in J_1} |Q_j| \leq \sum_{j\in J_1} 4\varepsilon \int_{Q_j} |b_Q^\varepsilon| \, \mathrm{d}x \leq 4\varepsilon \int_Q |b_Q^\varepsilon| \, \mathrm{d}x.$$

By Hölder's inequality and Proposition 14.6 (a), we can now conclude that

$$|F_1(Q)| \leq 4C\varepsilon|Q|. \tag{14.18}$$

The other collection is slightly more complicated. We first use Proposition 14.6 (b) to bound

$$\begin{aligned}
1 &\leq \mathrm{Re}\left(\xi \cdot \fint_Q b_Q^\varepsilon \, \mathrm{d}x\right) \\
&= \frac{1}{|Q|}\int_{Q\setminus F_2(Q)} \mathrm{Re}\big(\xi \cdot b_Q^\varepsilon\big) \, \mathrm{d}x + \frac{1}{|Q|}\sum_{j\in J_2}\int_{Q_j} \mathrm{Re}\big(\xi \cdot b_Q^\varepsilon\big) \, \mathrm{d}x.
\end{aligned}$$

For the integrals on Q_j, we can bring the failure of (14.16) into play and to the integral on $Q \setminus F_2(Q)$ we simply apply Hölder's inequality. The result is

$$\begin{aligned}
&\leq \frac{|Q \setminus F_2(Q)|^{1/2}}{|Q|}\|b_Q^\varepsilon\|_2 + \frac{1}{|Q|}\sum_{j\in J_2}\frac{3}{4}|Q_j| \\
&\leq \frac{C|Q \setminus F_2(Q)|^{1/2}}{|Q|^{1/2}} + \frac{3}{4}
\end{aligned}$$

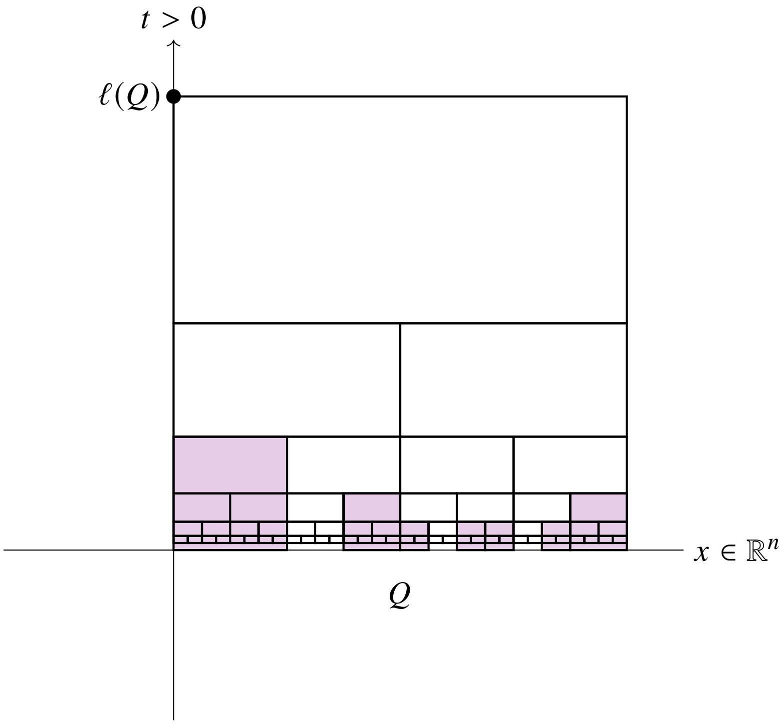

Figure 14.2.: A dyadic 'sawtooth' region created by removing from $R(Q)$ the purple Carleson boxes. They correspond to the dyadic children Q_j of Q in the scenario of Lemma 14.8. The remaining white region is called $E^*(Q)$.

and we rearrange terms to find

$$c|Q| := \frac{1}{(4C)^2}|Q| \leq |Q \setminus F_2(Q)| = |Q| - |F_2(Q)| \tag{14.19}$$

with some constant $c \in (0, 1)$.

Eventually, we combine (14.18) and (14.19) to give

$$|F_1(Q)| + |F_2(Q)| \leq \Big(4C\varepsilon + 1 - c\Big)|Q|,$$

whereupon the claim (14.17) follows with $\eta := c - 4C\varepsilon$, provided we take ε small enough to make this a positive quantity. □

Remark 14.9. Rather than relying on maximality, the same cubes Q_j can be selected through the use of a stopping time. This involves dyadically subdividing Q and halting the process each time either condition (14.15) or (14.16) fails. Hence, these Q_j are also referred to as *stopping time cubes*.

We have reached the stage of the proof where we fix ε as in Lemma 14.8 and henceforth drop ε from our notation by writing Γ and b_Q instead of Γ^ε and b_Q^ε, resprectively.

Combining Lemma 14.8 (b) and Proposition 14.6 (c), we see that for every dyadic cube Q we get

$$\begin{aligned}\iint_{E^*(Q)} |\gamma_t(x)\mathbf{1}_\Gamma(\gamma_t(x))|^2 \,\frac{\mathrm{d}x\,\mathrm{d}t}{t} &\le 4 \iint_{E^*(Q)} |\gamma_t(x)\cdot(\mathcal{A}_t b_Q)(x)|^2 \,\frac{\mathrm{d}x\,\mathrm{d}t}{t} \\ &\le \frac{4C}{\varepsilon^2}|Q|.\end{aligned} \tag{14.20}$$

The trade-off made thus far is that we get the Carleson measure estimate required for Proposition 14.5 not across the entire Carleson box, but only on its ample portion $E^*(Q)$ defined in Lemma 14.8. However, since all of this works on every cube, with uniform estimates, we can rely on the following, essential self-improvement principle.

Lemma 14.10 (John–Nirenberg lemma for Carleson measures)**.** *Let ν be a Borel measure on $\mathbb{R}^{n+1}_+$ and suppose that there exist constants $\kappa, \eta > 0$ with the following properties. Every dyadic cube $Q \in \square$ has pairwise disjoint dyadic children Q_j such that the sets $E(Q)$ and $E^*(Q)$ defined in* (14.14) *satisfy*

(a) $\big|E(Q)\big| \ge \eta|Q|$,

(b) $\nu\big(E^*(Q)\big) \le \kappa|Q|$.

Then ν is a Carleson measure with $\|\nu\|_C \le \kappa\eta^{-1}$.

Proof. Let us fix a dyadic cube $Q \in \square$, set $Q =: Q_{\alpha_0}$ and call Q the cube at 'stage 0'. We enumerate the dyadic children from the assumption by indices α_1. The pairwise disjoint children Q_{α_1} have sidelength $\ell(Q_{\alpha_1}) \le 2^{-1}\ell(Q)$ and leave out an ample portion of Q in the sense that by (a) we have

$$\frac{\sum_{\alpha_1}|Q_{\alpha_1}|}{|Q|} = \frac{|Q\setminus E(Q)|}{|Q|} = 1 - \frac{|E(Q)|}{|Q|} \le (1-\eta).$$

They are called cubes of 'stage 1'. The assumption re-applies simultaneously to each cube Q_{α_1}, leading to the cubes Q_{α_2} of 'stage 2' labeled by indices α_2. This process can be iterated. On the k-th stage we have cubes Q_{α_k} of sidelength at most $2^{-k}\ell(Q)$. Since with every selection we keep at most the $(1-\eta)$-th fraction of each cube for the next stage, they also satisfy

$$\frac{\sum_{\alpha_k}|Q_{\alpha_k}|}{|Q|} \le \frac{\sum_{\alpha_{k-1}}(1-\eta)|Q_{\alpha_{k-1}}|}{|Q|} \le \ldots \le (1-\eta)^k. \tag{14.21}$$

We make the important observation that for every point $(x,t) \in R(Q)$ there is a stage k in the iteration scheme at which (x,t) does no longer belong to a Carleson box of a cube of stage k. Indeed, $(x,t) \in R(Q_{\alpha_k})$ means that $0 < t \le \ell(Q_{\alpha_k}) \le 2^{-k}\ell(Q)$.

With this observation at hand, we can proceed to iteratively decompose

$$\begin{aligned}
R(Q) &= \Big(E^*(Q_{\alpha_0})\Big) \cup \Big(\bigcup_{\alpha_1} R(Q_{\alpha_1})\Big) \\
&= \dots \\
&= \Big(\bigcup_{k=0}^{m} \bigcup_{\alpha_k} E^*(Q_{\alpha_k})\Big) \cup \Big(\bigcup_{\alpha_{m+1}} R(Q_{\alpha_{m+1}})\Big) \\
&= \dots \\
&= \bigcup_{k=0}^{\infty} \bigcup_{\alpha_k} E^*(Q_{\alpha_k})
\end{aligned}$$

and conclude the desirable Carleson measure estimate

$$\begin{aligned}
\nu\big(R(Q)\big) &\le \sum_{k=0}^{\infty} \sum_{\alpha_k} \nu\big(E^*(Q_{\alpha_k})\big) \\
&\overset{\text{(b)}}{\le} \sum_{k=0}^{\infty} \sum_{\alpha_k} \kappa |Q_{\alpha_k}| \\
&\overset{(14.21)}{\le} \sum_{k=0}^{\infty} \kappa (1-\eta)^k |Q| \\
&= \frac{\kappa}{\eta} |Q|. \qquad \square
\end{aligned}$$

As far as the Kato conjecture is concerned, there is not much left to do.

Proof of Proposition 14.5. Apply the John–Nirenberg lemma to the measure given by $\mathrm{d}\nu(x,t) = |\gamma_t(x)\mathbf{1}_\Gamma(\gamma_t(x))|^2 \, \frac{\mathrm{d}x \, \mathrm{d}t}{t}$ — the assumptions have been verified beforehand in Lemma 14.8 (a) and (14.20). $\square$

In Section 14.2 we have already seen that Proposition 14.5 implies the solution of the Kato conjecture.

The end. For now.

14.5. Exercises

Exercise 14.1 (The principal part approximation for $n = 1$)**.**

(a) Prove that $\{\nabla u \mid u \in \mathrm{H}^1(\mathbb{R}^n)\}$ is dense in $\mathrm{L}^2(\mathbb{R}^n)^n$ if and only if $n = 1$.

(b) Conclude that in dimension $n = 1$ the principal part approximation can be extended as stated in Lemma 14.3

Exercise 14.2 (Operator-adapted Hodge decomposition)**.** We consider the bounded operators $\nabla : \mathrm{H}^1(\mathbb{R}^n) \to \mathrm{L}^2(\mathbb{R}^n)^n$ and $\operatorname{div}(A\,\cdot) : \mathrm{L}^2(\mathbb{R}^n)^n \to \mathrm{H}^{-1}(\mathbb{R}^n)$. The goal in this exercise is to prove that they induce the L-adapted topological *Hodge decomposition*

$$\mathrm{L}^2(\mathbb{R}^n)^n = \ker(\operatorname{div}(A\,\cdot)) \oplus \overline{\operatorname{ran}(\nabla)}. \tag{14.22}$$

When $L = -\Delta$ is the negative Laplacian, this is the Leray–Helmholtz decomposition of $\mathrm{L}^2(\mathbb{R}^n)^n$ into divergence free vector fields and gradient fields.

You may proceed as follows.

(a) Check that $\ker(\operatorname{div}(A\,\cdot)) \cap \overline{\operatorname{ran}(\nabla)} = \{0\}$.

(b) Prove that

$$P := -\nabla \mathscr{L}^{-1} \operatorname{div}(A\,\cdot)$$

is a bounded projection on $\mathrm{L}^2(\mathbb{R}^n)^n$.

(c) Show that $\overline{\operatorname{ran}(\nabla)} \subseteq \operatorname{ran}(P)$ and that $\ker(\operatorname{div}(A\,\cdot)) \subseteq \operatorname{ran}(1 - P)$.

(d) Deduce that (14.22) is indeed a topological decomposition.

Exercise 14.3 (An application of the Hodge decomposition)**.** The L-adapted Hodge decomposition offers a method for extending the quadratic estimate from the initial stage of the proof in Lecture 13 as follows.

(a) Prove that for all $F \in \mathrm{L}^2(\mathbb{R}^n)^n$ we have

$$\int_0^\infty \|\Theta_t F\|_2^2 \,\frac{\mathrm{d}t}{t} \lesssim \|F\|_2^2.$$

(b) Does the principal part approximation in Proposition 13.15 also extend in a similar manner a posteriori?

Exercise 14.4 (The bounded H^∞-calculus on $\mathrm{H}^1(\mathbb{R}^n)$)**.** Let $\varphi \in (\varphi_L, \pi)$. Prove the estimate

$$\|\nabla f(L)u\|_2 \lesssim \|f\|_{\infty,\varphi} \|\nabla u\|_2 \qquad (f \in \mathrm{H}^\infty(\mathrm{S}_\varphi),\ u \in \mathrm{H}^1(\mathbb{R}^n)).$$

A. Vector-valued integration

From Lecture 4 onwards, we will occasionally need integrals of vector-valued functions defined on $\mathbb{R}^n$. This self-contained appendix provides you with the essentials of the construction and we refer to [HvNVW16, ABHN01] for further background.

Notation A.1. In the following, X and Y will be Banach spaces (over the complex numbers).

A.1. Measurable functions

We begin by introducing simple functions and a first concept of measurability for functions valued in X.

Definition A.2. *A function $f\colon \mathbb{R}^n \to X$ is called* simple *if there exist finitely many disjoint Lebesgue measurable sets $E_k \subseteq \mathbb{R}^n$ with finite Lebesgue measure and vectors $u_k \in X$ such that*

$$f = \sum_k u_k \cdot \mathbf{1}_{E_k}.$$

Definition A.3. *A function $f\colon \mathbb{R}^n \to X$ is called* strongly measurable *if there exists a sequence of simple functions that converges to f almost everywhere.*

Example A.4. (a) Continuous functions can be approximated by simple functions uniformly on compact subsets of $\mathbb{R}^n$, hence are strongly measurable.

(b) If $g\colon \mathbb{R}^n \to \mathbb{C}$ is measurable and $u \in X$, then $f := u \cdot g$ is strongly measurable since g can be approximated by scalar-valued simple functions almost everywhere.

(c) If a sequence of strongly measurable functions converges almost everywhere, then (by diagonalization) the limit function is strongly measurable.

(d) If $f\colon \mathbb{R}^n \to X$ is strongly measurable and $\phi : X \to Y$ is continuous, then $\phi \circ f$ is strongly measurable. Indeed, if a sequence of simple functions (f_j) tends to f almost everywhere, then $(\phi \circ f_j)$ is a sequence of simple functions that tends to $\phi \circ f$ almost everywhere. In particular, $\|f\|_X\colon \mathbb{R}^n \to \mathbb{R}$ is (Lebesgue) measurable.

In the scalar case $X = \mathbb{C}$, strong measurability coincides with Lebesgue measurability. You may still wonder: Why not using the equivalent definition via pre-images of Borel sets? The reason is that contrarily to the separable space $\mathbb{C}$, a general Banach space and its associated Borel σ-algebra can be 'too large'. The following observation hopefully gives a feeling for the general issue and we refer to [HvNVW16, Sect. 1.1] for an extensive discussion.

Suppose that f is strongly measurable and fix an approximating sequence (f_j) of simple functions that tends to f outside of a null set N. Then

$$f(\mathbb{R}^n \setminus N) \subseteq \overline{\operatorname{span}\{f_j(\mathbb{R}^n) \mid j \in \mathbb{N}\}} =: Y,$$

where Y is separable. In other words: f almost everywhere takes values in a potentially small subspace of X if the latter is not separable itself. An example of a non-separable space that we will frequently encounter is $X = \mathcal{L}(H)$, where H is any infinite-dimensional Hilbert space. So, this property is a necessary condition for a function to be strongly measurable and we give it a name.

Definition A.5. *A function $f : \mathbb{R}^n \to X$ is called* almost separably-valued *if there exists a null set $N \subseteq \mathbb{R}^n$ such that $f(\mathbb{R}^n \setminus N)$ is contained in a separable closed subspace of X.*

By Example A.4 (d), strongly measurable functions also have the following property.

Definition A.6. *A function $f : \mathbb{R}^n \to X$ is called* weakly measurable *if for any fixed $\phi \in X'$ the function $\langle \phi, f \rangle_{X',X} : \mathbb{R}^n \to \mathbb{C}$ is Lebesgue measurable.*

Through the use of functionals on X, we can resort to the theory of scalar measurable functions. The following theorem nicely ties together these lines of thought. It will not be needed for the lectures and hence we only cite it without proof.

Theorem A.7 (Pettis, [ABHN01, Thm. 1.1.1]). *A function $f : \mathbb{R} \to X$ is strongly measurable if and only if it is weakly measurable and almost separably-valued. In particular, strong and weak measurability are equivalent in separable Banach spaces.*

A.2. The Bochner integral

With the 'correct' notion of measurability being set up, we can now follow the construction of the scalar Lebesgue integral. For a simple function $f = \sum_k u_k \cdot \mathbf{1}_{E_k}$ on $\mathbb{R}^n$ the *Bochner integral* is defined as

$$\int_{\mathbb{R}^n} f \, \mathrm{d}t := \sum_k u_k \cdot |E_k|.$$

It is routine to check that for simple functions $f, g\colon \mathbb{R} \to X$ and scalars $\alpha \in \mathbb{C}$ we have

$$\int_{\mathbb{R}^n} \big(\alpha f + g\big)\, \mathrm{d}t = \alpha \int_{\mathbb{R}^n} f\, \mathrm{d}t + \int_{\mathbb{R}^n} g\, \mathrm{d}t.$$

Moreover, the triangle inequality implies

$$\left\| \int_{\mathbb{R}^n} f\, \mathrm{d}t \right\|_X \le \int_{\mathbb{R}^n} \|f\|_X\, \mathrm{d}t. \tag{A.1}$$

Definition A.8. *A strongly measurable function $f\colon \mathbb{R}^n \to X$ is* Bochner integrable *if there exists a sequence of simple functions (f_j) such that*

$$\lim_{j\to\infty} \int_{\mathbb{R}^n} \|f_j - f\|_X\, \mathrm{d}t = 0.$$

First of all, this definition makes sense because the functions $\|f - f_j\|_X$ are strongly measurable by Example A.4 (d). Moreover, we see from

$$\begin{aligned} \left\| \int_{\mathbb{R}^n} f_j\, \mathrm{d}t - \int_{\mathbb{R}^n} f_k\, \mathrm{d}t \right\|_X &\le \int_{\mathbb{R}^n} \|f_j - f_k\|_X\, \mathrm{d}t \\ &\le \int_{\mathbb{R}^n} \|f_j - f\|_X\, \mathrm{d}t + \int_{\mathbb{R}^n} \|f - f_k\|_X\, \mathrm{d}t \end{aligned} \tag{A.2}$$

that the integrals $\int_{\mathbb{R}^n} f_j\, \mathrm{d}t$, $j \in \mathbb{N}$, form a Cauchy sequence in X. By completeness this sequence has a limit in X. With a calculation similar to (A.2) we also see that the limit is independent of the approximating sequence.

Definition A.9. *In the situation of Definition A.8, the* Bochner integral *is defined as*

$$\int_{\mathbb{R}^n} f\, \mathrm{d}t := \lim_{j\to\infty} \int_{\mathbb{R}^n} f_j\, \mathrm{d}t.$$

As in the case of strong measurability, there is a particularly useful criterion for Bochner integrability that fully reduces the matter to scalar-valued Lebesgue integration.

Theorem A.10 (Bochner)**.** *A strongly measurable function $f\colon \mathbb{R}^n \to X$ is Bochner integrable if and only if $\|f\|_X\colon \mathbb{R}^n \to \mathbb{R}$ is (Lebesgue) integrable and in this case we have*

$$\left\| \int_{\mathbb{R}^n} f\, \mathrm{d}t \right\|_X \le \int_{\mathbb{R}^n} \|f\|_X\, \mathrm{d}t.$$

Proof. If f is Bochner integrable and (f_j) is a sequence as in Definition A.8, then for any fixed j we have

$$\int_{\mathbb{R}^n} \|f\|_X \,\mathrm{d}t \le \int_{\mathbb{R}^n} \|f - f_j\|_X \,\mathrm{d}t + \int_{\mathbb{R}^n} \|f_j\|_X \,\mathrm{d}t < \infty.$$

Conversely, suppose that $\|f\|_X$ is integrable. Since f is strongly measurable, there exists a sequence of simple functions (f_j) that converges to f almost everywhere. The simple functions

$$g_j(t) := \begin{cases} f_j(t), & \text{if } \|f_j(t)\|_X \le 2\|f(t)\|_X, \\ 0, & \text{else,} \end{cases}$$

have the same properties but additionally they satisfy $\|g_j - f\|_X \le 3\|f\|_X$. The (scalar!) dominated convergence theorem yields

$$\lim_{j\to\infty} \int_{\mathbb{R}^n} \|g_j - f\| \,\mathrm{d}t = 0,$$

which means that f is Bochner integrable. Finally, the triangle inequality follows from (A.1) for the simple functions f_j and passing to the limit as $j \to \infty$. □

In many situations one does not integrate over all of $\mathbb{R}^n$ but only over a measurable subset. In this case measurability and integrability are defined through the respective properties of the extension by zero.

Definition A.11. *Let $E \subseteq \mathbb{R}^n$ be measurable and $f: E \to X$. Let $\widetilde{f}: \mathbb{R}^n \to X$ be the extension of f by zero. We say that*

(a) f is strongly measurable *if $\widetilde{f}$ is strongly measurable,*

(b) f is Bochner integrable on E *if $\widetilde{f}$ is Bochner integrable and in this case we define*

$$\int_E f \,\mathrm{d}t := \int_{\mathbb{R}^n} \widetilde{f} \,\mathrm{d}t.$$

Example A.12. A continuous function $f: I \to X$ on a compact interval I is Bochner integrable. This is a consequence of Bochner's theorem, since f is strongly measurable by Example A.4 (a) and $\|f\|_X$ is bounded, hence integrable on I.

One of the most fundamental properties of the Bochner integral concerns its interaction with bounded linear operators.

Proposition A.13. *Let $E \subseteq \mathbb{R}^n$ be measurable and $f: E \to X$ be Bochner integrable. If T is a bounded linear (or anti-linear) operator from X into Y, then $Tf: E \to Y$ is Bochner integrable and*

$$T \int_E f(t) \,\mathrm{d}t = \int_E Tf(t) \,\mathrm{d}t.$$

Proof. Without loss of generality we can let $E = \mathbb{R}^n$. If (f_j) is as in Definition A.3 on strong measurability of f, then (Tf_j) has the same properties with regard to Tf. Hence, Tf is strongly measurable. Likewise, if (f_j) is as in Definition A.8 on Bochner integrability of f, then (Tf_j) has the same properties with regard to Tf and the claim follows by Definition A.9 of the Bochner integral. □

A situation that we shall encounter every now and then during the lectures is that X is a space of functions depending on a variable x. Hence, also the Bochner integral is a function of x and we would like to pull the 'exterior variable' x into the integral. Here is an example, where this works without any difficulty.

Corollary A.14. *Let $E \subseteq \mathbb{R}^n$ be measurable and $f\colon \mathbb{R}^n \to \mathrm{C}(K)$ be Bochner integrable, where $K \subseteq \mathbb{R}^m$ is compact. Then $t \mapsto f(t)(x)$ is integrable for every $x \in K$ and we have*

$$\left(\int_E f(t)\,\mathrm{d}t\right)(x) = \int_E f(t)(x)\,\mathrm{d}t. \tag{A.3}$$

Proof. Simply apply Proposition A.13 to the point evaluations $\mathrm{ev}_x\colon \mathrm{C}(K) \to \mathbb{C}$, $\mathrm{ev}_x\, u \coloneqq u(x)$. □

For Bochner integrable functions $f\colon E \to \mathrm{L}^2(K)$ it is still reasonable to expect a rule like (A.3) for a.e. x, but the argument cannot be the same since point evaluation is not even well-defined on X. In this case already measurability of the scalar-valued functions $f(t)(x)$ can become an issue and we refer to [HvNVW16, Prop. 1.2.25] for a careful discussion of these points.

In the lectures we will be able to circumvent the problem completely by working only with strongly measurable functions $f\colon E \to \mathrm{L}^2(K)$ that take their values in the smaller space $\mathrm{C}(K)$ and are Bochner integrable with respect to the finer topology. In this case we can decide in which space we consider the Bochner integral and in particular we can chose $X = \mathrm{C}(K)$ so that (A.3) holds.

Corollary A.15. *Let $E \subseteq \mathbb{R}^n$ be measurable and let $f\colon E \to X$ be Bochner integrable. Suppose that $X \subseteq Y$ with continuous inclusion. Then f is also Bochner integrable as a function valued in Y and the Bochner integrals in X and Y are the same.*

Proof. This is Proposition A.13 applied to the inclusion map $X \hookrightarrow Y$. □

A very similar property for integrals of operator-valued functions reads as follows.

Corollary A.16. *Let $E \subseteq \mathbb{R}^n$ be measurable and $L\colon E \to \mathcal{L}(X,Y)$ be Bochner integrable. Then $t \mapsto L(t)x$ is Bochner integrable for every $x \in X$ and*

$$\left(\int_E L(t)\,\mathrm{d}t\right)x = \int_E L(t)x\,\mathrm{d}t.$$

Proof. This is Proposition A.13 applied to the evaluation map $\mathrm{ev}_x\colon \mathcal{L}(X,Y) \to Y$, $\mathrm{ev}_x L = Lx$. □

Many properties of the scalar Lebesgue integral that *do not* rely on order properties such as non-negativity or monotonicity have their counterparts for Bochner integrals. Below, we present the two most important examples for our course.

Theorem A.17 (Dominated convergence)**.** *Let $E \subseteq \mathbb{R}^n$ be measurable and $f_j\colon E \to X$, $j \in \mathbb{N}$, be Bochner integrable functions. If there exists a function $f\colon E \to X$ and a non-negative integrable function $g\colon E \to \mathbb{R}$ such that almost everywhere we have $\lim_{j\to\infty} f_j = f$ and $\|f_j\|_X \le g$ for every $j \in \mathbb{N}$, then f is Bochner integrable and*

$$\lim_{j\to\infty} \int_E \|f - f_j\|_X \,\mathrm{d}t = 0.$$

In particular, we have

$$\lim_{j\to\infty} \int_E f_j \,\mathrm{d}t = \int_E f \,\mathrm{d}t.$$

Proof. Measurability of f follows by Example A.4 (c) and as we have $\|f\|_X \le g$ almost everywhere, f is Bochner integrable thanks to Theorem A.10. As we have $\|f_j - f\|_X \le 2g$ almost everywhere, the claim follows from the scalar-valued dominated convergence theorem. □

Theorem A.18 (Fubini)**.** *Let $f\colon \mathbb{R}^m \times \mathbb{R}^n \to X$ be Bochner integrable. Then:*

(a) For almost every $s \in \mathbb{R}^m$ the function $f(s,\cdot)$ is Bochner integrable.

(b) For almost every $t \in \mathbb{R}^n$ the function $f(\cdot,t)$ is Bochner integrable.

(c) The functions $\int_{\mathbb{R}^n} f(\cdot,t)\,\mathrm{d}t$ and $\int_{\mathbb{R}^m} f(s,\cdot)\,\mathrm{d}s$ are Bochner integrable and

$$\int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^n} f(s,t)\,\mathrm{d}t\right)\mathrm{d}s = \int_{\mathbb{R}^m\times\mathbb{R}^n} f(s,t)\,\mathrm{d}(s,t) = \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^m} f(s,t)\,\mathrm{d}s\right)\mathrm{d}t.$$

Proof. The key observation is that the scalar-valued version of Fubini's theorem together with Example A.4 (b) already yields the claim in the special case that $f = u \cdot \mathbf{1}_E$ with $u \in X$ and $E \subseteq \mathbb{R}^m \times \mathbb{R}^n$ a set with finite Lebesgue measure. The claim for simple functions f then follows by linearity.

In the general case let (f_j) be a sequence of simple functions such that $\|f_j\|_X \le 2\|f\|_X$ everywhere and $f_j \to f$ in the limit as $j \to \infty$ a.e. on $\mathbb{R}^m \times \mathbb{R}^n$. (Such a sequence was constructed in the proof of Bochner's theorem). Hence, almost every $s \in \mathbb{R}^m$ has the property that

- all functions $f_j(s,\cdot)$ are measurable,

- we have $f_j(s,\cdot) \to f(s,\cdot)$ a.e. on $\mathbb{R}^n$
- and $\|f_j(s,\cdot)\|_X \leq 2\|f(s,\cdot)\|_X$ a.e. on $\mathbb{R}^n$, where, taking into account the scalar-valued Fubini's theorem, $\|f(s,\cdot)\|_X$ is integrable.

For s as above we conclude from Example A.4 (c) that $f(s,\cdot)$ is measurable and Bochner integrability follows from Bochner's theorem. This proves (a). Moreover, we find by dominated convergence that

$$\lim_{j\to\infty} \int_{\mathbb{R}^m} f_j(s,t)\,\mathrm{d}t = \int_{\mathbb{R}^m} f(s,t)\,\mathrm{d}t.$$

Since this happens for almost every s, we conclude that $\int_{\mathbb{R}^m} f(\cdot,t)\,\mathrm{d}t$ is measurable. This function is also bounded by $2\int_{\mathbb{R}^n} \|f(\cdot,t)\|_X\,\mathrm{d}t$, which is integrable once again by the scalar-valued Fubini's theorem. Dominated convergence now yields that

$$\begin{aligned}\int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^n} f(s,t)\,\mathrm{d}t\right)\mathrm{d}s &= \lim_{j\to\infty}\int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^n} f_j(s,t)\,\mathrm{d}t\right)\mathrm{d}s \\ &= \lim_{j\to\infty}\int_{\mathbb{R}^m\times\mathbb{R}^n} f_j(s,t)\,\mathrm{d}(s,t) = \int_{\mathbb{R}^m\times\mathbb{R}^n} f(s,t)\,\mathrm{d}(s,t).\end{aligned}$$

We have proved the first half of (c). The second half and (b) follow by switching the roles of s and t. □

A.3. Vector-valued holomorphic functions

Holomorphic functions are often considered as the most beautiful of all functions. To the long list of their remarkable properties we shall add in this section that their vector-valued theory is almost as easy as the scalar one.

Definition A.19. *Let $\Omega \subseteq \mathbb{C}$ be a non-empty, open set. A function $f\colon \Omega \to X$ is*

(a) holomorphic *if*

$$f'(z_0) \coloneqq \lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0} \tag{A.4}$$

exists for all $z_0 \in \Omega$.

(b) weakly holomorphic *if $\langle \phi, f\rangle_{X',X} : \Omega \to \mathbb{C}$ is holomorphic for every $\phi \in X'$.*

Example A.20. Given $z_0 \in \mathbb{C}$ and $(u_k) \subseteq X$, consider the power series

$$f(z) \coloneqq \sum_{k=0}^{\infty} (z-z_0)^k u_k$$

with radius of convergence $\rho := (\limsup_{k\to\infty} \sqrt[k]{\|u_k\|_X})^{-1}$. Exactly as in the scalar-valued setting we obtain that f is holomorphic in $B(z_0, \rho)$ and that f' can be computed term-by-term, leading to the formulæ $f^{(k)}(z_0) = k!u_k$ for the iterated complex derivatives.

Holomorphic functions are continuous, hence in particular bounded on compact subsets of Ω, and weakly holomorphic. In contrast to measurability, no subtleties concerning separability need to be taken into account when comparing the two notions of holomorphy.

Theorem A.21. *Let $\Omega \subseteq \mathbb{C}$ be a non-empty, open set and $f\colon \Omega \to X$ a function. Then f is holomorphic if and only if it is weakly holomorphic.*

Proof. Above, we already discussed the direct part. Now, we assume that f is weakly holomorphic and fix $z_0 \in \Omega$. Since Ω is open, we can also fix a radius $r > 0$ with $\overline{B(z_0, 2r)} \subseteq \Omega$ and since weakly bounded sets are bounded, $C := \sup_{|z-z_0|\le r} \|f(z)\|_X$ is finite. For $z, z' \in B(z_0, r) \setminus \{z_0\}$ we set

$$F(z, z') := \frac{f(z) - f(z_0)}{z - z_0} - \frac{f(z') - f(z_0)}{z' - z_0}.$$

We are going to prove the bound

$$\|F(z, z')\|_X \le \frac{C|z - z'|}{r^2}, \tag{A.5}$$

which, by the Cauchy criterion, implies that the limit in (A.4) exists. To this end, take any $\phi \in X'$ normalized to $\|\phi\|_{X'} = 1$. By means of the scalar-valued Cauchy integral formula we can write

$$\begin{aligned}
&\langle \phi, F(z, z')\rangle_{X',X} \\
&\quad= \frac{1}{z - z_0} \cdot \frac{1}{2\pi \mathrm{i}} \int_{|\lambda - z_0| = 2r} \langle \phi, f(\lambda)\rangle_{X',X} \left(\frac{1}{\lambda - z} - \frac{1}{\lambda - z_0} \right) \mathrm{d}\lambda \\
&\quad\quad - \frac{1}{z' - z_0} \cdot \frac{1}{2\pi \mathrm{i}} \int_{|\lambda - z_0| = 2r} \langle \phi, f(\lambda)\rangle_{X',X} \left(\frac{1}{\lambda - z'} - \frac{1}{\lambda - z_0} \right) \mathrm{d}\lambda \\
&\quad= \frac{1}{2\pi \mathrm{i}} \int_{|\lambda - z_0| = 2r} \langle \phi, f(\lambda)\rangle_{X',X} \left(\frac{1}{(\lambda - z)(\lambda - z_0)} - \frac{1}{(\lambda - z')(\lambda - z_0)} \right) \mathrm{d}\lambda \\
&\quad= \frac{1}{2\pi \mathrm{i}} \int_{|\lambda - z_0| = 2r} \langle \phi, f(\lambda)\rangle_{X',X} \left(\frac{z - z'}{(\lambda - z)(\lambda - z')(\lambda - z_0)} \right) \mathrm{d}\lambda.
\end{aligned}$$

Since $z, z' \in B(z_0, r)$, the integrand is bounded in norm by ${}^{C|z-z'|}/_{2r^3}$ and we obtain

$$|\langle \phi, F(z, z')\rangle_{X',X}| \le \frac{C|z - z'|}{r^2}.$$

By virtue of the Hahn–Banach theorem, this bound implies (A.5) and the proof is complete. ◻

Suppose that $f\colon \Omega \to X$ is holomorphic and $\gamma\colon I \to \Omega$ is a continuously differentiable function on a compact interval I, henceforth called *compact* C^1*-path* in Ω. We define the *path integral*

$$\int_\gamma f \,\mathrm{d}z := \int_I f(\gamma(t)) \cdot \gamma'(t) \,\mathrm{d}t,$$

now as a Bochner integral in X, keeping in mind that by Example A.12 a continuous function on a compact interval is Bochner integrable. The Hahn–Banach theorem allows to transfer identities for path integrals from the scalar- to the vector-valued setting. You should keep that strategy in mind. Here are two examples — more will pop up during the lectures.

Example A.22 (Cauchy integral theorem)**.** Suppose that $\Omega \subseteq \mathbb{C}$ is non-empty and open and γ is a compact C^1-path in Ω. Pick your favourite geometric assumption that guarantees a Cauchy integral theorem of the form

$$\int_\gamma f(z) \,\mathrm{d}z = 0$$

for all *scalar-valued* holomorphic functions f on Ω. Then the same is true in the vector-valued setting. Indeed, if $f\colon \Omega \to X$ is holomorphic and $\phi \in X'$ is any functional, then the scalar-valued theory yields

$$0 = \int_\gamma \langle \phi, f(z) \rangle_{X',X} \,\mathrm{d}z \overset{\text{A.13}}{=} \left\langle \phi, \int_\gamma f(z) \,\mathrm{d}z \right\rangle_{X',X}$$

and hence $\int_\gamma f(z) \,\mathrm{d}z = 0$ by the Hahn–Banach theorem.

Example A.23 (Cauchy integral formula)**.** Exactly as in the previous example you can also pick your favourite geometric assumption that guarantees a Cauchy integral formula of the form

$$f^{(k)}(z_0) = \frac{k!}{2\pi \mathrm{i}} \int_\gamma \frac{f(z)}{(z - z_0)^{k+1}} \,\mathrm{d}z$$

for a fixed $z_0 \in \Omega$ and all scalar-valued holomorphic functions f on Ω and conclude that the same is true in the vector-valued setting.

Bibliography

[AA11] Pascal Auscher and Andreas Axelsson. Weighted maximal regularity estimates and solvability of non-smooth elliptic systems I. *Invent. Math.*, 184(1):47–115, 2011.

[ABHN01] Wolfgang Arendt, Charles J. K. Batty, Matthias Hieber, and Frank Neubrander. *Vector-valued Laplace transforms and Cauchy problems*. Birkhäuser Verlag, Basel, 2001.

[ABHR15] Pascal Auscher, Nadine Badr, Robert Haller-Dintelmann, and Joachim Rehberg. The square root problem for second-order, divergence form operators with mixed boundary conditions on L^p. *J. Evol. Equ.*, 15(1):165–208, 2015.

[AE09] Herbert Amann and Joachim Escher. *Analysis III*. Birkhäuser Verlag, Basel, 2009.

[AE23] Pascal Auscher and Moritz Egert. *Boundary value problems and Hardy spaces for elliptic systems with block structure*, volume 346 of *Progress in Mathematics*. Birkhäuser/Springer, Cham, 2023.

[AEN20] Pascal Auscher, Moritz Egert, and Kaj Nyström. L^2 well-posedness of boundary value problems for parabolic systems with measurable coefficients. *J. Eur. Math. Soc. (JEMS)*, 22(9):2943–3058, 2020.

[AGP18] Josefina Álvarez and Martha Guzmán-Partida. The $\mathrm{T}(1)$ theorem revisited. *Surv. Math. Appl.*, 13:41–94, 2018.

[AH96] David R. Adams and Lars I. Hedberg. *Function spaces and potential theory*. Springer-Verlag, Berlin, 1996.

[AHL+02] Pascal Auscher, Steve Hofmann, Michael Lacey, Alan McIntosh, and Philippe Tchamitchian. The solution of the Kato square root problem for second order elliptic operators on $\mathbb{R}^n$. *Ann. of Math. (2)*, 156(2):633–654, 2002.

[AHLT01] Pascal Auscher, Steve Hofmann, John L. Lewis, and Philippe Tchamitchian. Extrapolation of Carleson measures and the analyticity of Kato's square-root operators. *Acta Math.*, 187(2):161–190, 2001.

[AKM06] Andreas Axelsson, Stephen Keith, and Alan McIntosh. Quadratic estimates and functional calculi of perturbed Dirac operators. *Invent. Math.*, 163(3):455–497, 2006.

[Alt16] Hans W. Alt. *Linear Functional Analysis*. Springer-Verlag London, Ltd., London, 2016.

[AT98] Pascal Auscher and Philippe Tchamitchian. Square root problem for divergence operators and related topics. *Astérisque*, (249):viii+172, 1998.

[Aus23] Pascal Auscher. Impact of Yves Meyer's work on Kato's conjecture. *Preprint,* `https://arxiv.org/abs/2311.01068`, 2023.

[Bal60] Alampallam V. Balakrishnan. Fractional powers of closed operators and the semigroups generated by them. *Pacific J. Math.*, 10:419–437, 1960.

[BEH20] Sebastian Bechtel, Moritz Egert, and Robert Haller-Dintelmann. The Kato square root problem on locally uniform domains. *Adv. Math.*, 375:107410, 37, 2020.

[Car62] Lennart Carleson. Interpolations by bounded analytic functions and the corona problem. *Ann. of Math. (2)*, 76:547–559, 1962.

[CD23] Andrea Carbonaro and Oliver Dragičević. On semigroup maximal operators associated with divergence-form operators with complex coefficients. *Preprint,* `https://arxiv.org/abs/2207.11045`, 2023.

[CDMY96] Michael Cowling, Ian Doust, Alan McIntosh, and Atsushi Yagi. Banach space operators with a bounded H^∞ functional calculus. *J. Austral. Math. Soc. Ser. A*, 60(1):51–89, 1996.

[CJ87] Michael Christ and Jean-Lin Journé. Polynomial growth estimates for multilinear singular integral operators. *Acta Math.*, 159(1-2):51–80, 1987.

[CMM82] Ronald R. Coifman, Alan McIntosh, and Yves Meyer. L'intégrale de Cauchy définit un opérateur borné sur L^2 pour les courbes lipschitziennes. *Ann. of Math. (2)*, 116(2):361–387, 1982.

[CRS10] Luis A. Caffarelli, Jean-Michel Roquejoffre, and Ovidiu Savin. Nonlocal minimal surfaces. *Comm. Pure Appl. Math.*, 63(9):1111–1144, 2010.

[Dav95] E. Brian Davies. Uniformly elliptic operators with measurable coefficients. *J. Funct. Anal.*, 132(1):141–169, 1995.

[DJ84] Guy R. David and Jean-Lin Journé. A boundedness criterion for generalized Calderón-Zygmund operators. *Ann. of Math. (2)*, 120(2):371–397, 1984.

[DJS85] Guy R. David, Jean-Lin Journé, and Stephen W. Semmes. Opérateurs de Calderón-Zygmund, fonctions para-accrétives et interpolation. *Rev. Mat. Iberoamericana*, 1(4):1–56, 1985.

[DNPV12] Eleonora Di Nezza, Giampiero Palatucci, and Enrico Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.*, 136(5):521–573, 2012.

[DS58] Nelson Dunford and Jacob T. Schwartz. *Linear Operators. I. General Theory*, volume 7 of *Pure and Applied Mathematics*. Interscience Publishers, Inc., New York; Interscience Publishers Ltd., London, 1958. With the assistance of W. G. Bade and R. G. Bartle.

[Eva10] Lawrence C. Evans. *Partial Differential Equations*. American Mathematical Society, Providence, RI, 2010.

[Gaf59] Matthew P. Gaffney. The conservation property of the heat equation on Riemannian manifolds. *Comm. Pure Appl. Math.*, 12:1–11, 1959.

[Gra14] Loukas Grafakos. *Classical Fourier Analysis*. Springer, New York, 2014.

[Haa06] Markus Haase. *The functional calculus for sectorial operators*. Birkhäuser Verlag, Basel, 2006.

[HKMP15] Steve Hofmann, Carlos Kenig, Svitlana Mayboroda, and Jill Pipher. Square function/non-tangential maximal function estimates and the Dirichlet problem for non-symmetric elliptic operators. *J. Amer. Math. Soc.*, 28(2):483–529, 2015.

[HvNVW16] Tuomas Hytönen, Jan van Neerven, Mark Veraar, and Lutz Weis. *Analysis in Banach spaces. Vol. I.* Springer, Cham, 2016.

[Kat61] Tosio Kato. Fractional powers of dissipative operators. *J. Math. Soc. Japan*, 13:246–274, 1961.

[Kat95] Tosio Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[Ken94] Carlos E. Kenig. *Harmonic analysis techniques for second order elliptic boundary value problems*, volume 83 of *CBMS Regional Conference Series in Mathematics*. Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1994.

[McI72] Alan McIntosh. On the comparability of $A^{1/2}$ and $A^{*1/2}$. *Proc. Amer. Math. Soc.*, 32:430–434, 1972.

[McI86] Alan McIntosh. Operators which have an H_∞ functional calculus. In *Miniconference on operator theory and partial differential equations (North Ryde, 1986)*, volume 14 of *Proc. Centre Math. Anal. Austral. Nat. Univ.*, pages 210–231. Austral. Nat. Univ., Canberra, 1986.

[Mor12] Andrew J. Morris. The Kato square root problem on submanifolds. *J. Lond. Math. Soc. (2)*, 86(3):879–910, 2012.

[MY90] Alan McIntosh and Atsushi Yagi. Operators of type ω without a bounded H_∞ functional calculus. In *Miniconference on Operators in Analysis (Sydney, 1989)*, volume 24 of *Proc. Centre Math. Anal. Austral. Nat. Univ.*, pages 159–172. Austral. Nat. Univ., Canberra, 1990.

[Ros13] Andreas Rosén. Layer potentials beyond singular integral operators. *Publ. Mat.*, 57(2):429–454, 2013.

[Rud83] Walter Rudin. Well-distributed measurable sets. *Amer. Math. Monthly*, 90(1):41–42, 1983.

[Sem90] Stephen Semmes. Square function estimates and the $T(b)$ theorem. *Proc. Amer. Math. Soc.*, 110(3):721–726, 1990.

[vN51] Johann von Neumann. Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes. *Math. Nachr.*, 4:258–281, 1951.

Index