The Helmholtz equation

Constant coefficients and periodic media

Ben Schweizer

Lecture Notes, Summer 2026

This is the machine-translated version of the german original

Contents

1. Introduction

Wave equation

The second-order linear wave equation with a constant $c > 0$ is

$$\partial_t^2 u = c^2 \Delta u \,. \tag{1.1}$$

It is used as a simple model in many physical applications, for example in the description of elastic materials, the propagation of sound waves or electromagnetic waves.

The one-dimensional case is a model for the vibrations of a string. If $u(x,t)$ describes the displacement of the string from the rest position at time $t > 0$ at a reference point $x \in \mathbb{R}$, then $\Delta u(x,t) = \partial_x^2 u(x,t)$ is a measure of the curvature at that point and thus a measure of the restoring force. Newton's law therefore yields the one-dimensional version for some $c > 0$,

$$\partial_t^2 u = c^2 \, \partial_x^2 u \,. \tag{1.2}$$

In the whole-space problem, we seek $u : \mathbb{R} \times [0,\infty) \to \mathbb{R}$. In analogy to ordinary second-order differential equations, we provide initial conditions for the values and the first time derivatives,

$$u(.,0) = u_0, \quad \partial_t u(.,0) = u_1 \,. \tag{1.3}$$

Two solutions of the equation (1.2) can be given in elementary terms. For arbitrary functions $\varphi, \psi \in C^2(\mathbb{R},\mathbb{R})$, the functions

$$u(x,t) = \varphi(x + ct) \quad \text{and} \quad u(x,t) = \psi(x - ct) \tag{1.4}$$

satisfy the wave equation. This can be verified by differentiating twice. By superposition of the elementary solutions, both initial conditions in (1.3) can be satisfied.

Helmholtz equation

A separation ansatz $u(x,t) = v(x)w(t)$ is often useful for wave equations. We use the ansatz corresponding to an oscillation with a fixed frequency $\omega > 0$ and set $w(t) = e^{-i\omega t}$. If $u(x,t) = v(x)w(t)$ solves the wave equation $\partial_t^2 u = c^2 \Delta u$, then v solves the Helmholtz equation

$$-c^2 \, \Delta v = \omega^2 v \quad \text{in } \Omega \,. \tag{1.5}$$

Conversely, each solution v of the Helmholtz equation (1.5) yields a solution $u(x,t) = v(x)e^{-i\omega t}$ of the wave equation (initial values, however, are generally not satisfied).

One can also derive the Helmholtz equation with a Fourier transform: Let u be a solution of the wave equation and $\hat{u} = \hat{u}(x, \omega)$ the (in time) Fourier transform of u. Then every time derivative of u translates into a multiplication of $\hat{u}$ by the factor $i\omega$. The function $\hat{u}(\cdot, \omega) : \Omega \to \mathbb{C}$ is therefore a solution of the Helmholtz equation (1.5). This can also be reversed: If we can solve the Helmholtz equation for $\hat{u}(\cdot, \omega)$ for each $\omega \in \mathbb{R}$, we find a function $\hat{u} = \hat{u}(x, \omega)$ whose inverse Fourier transform is a solution of the wave equation. However, initial values cannot be specified in this way.

If we consider a wave equation with sources $F(x,t) = f(x)e^{-i\omega t}$, i.e. $\partial_t^2 u = c^2 \Delta u + F$, we find a solution using the ansatz $u(x,t) = v(x)e^{-i\omega t}$, if v satisfies the Helmholtz equation with the right-hand side f,

$$-c^2 \Delta v - \omega^2 v = f \quad \text{ in } \Omega \,. \tag{1.6}$$

First examples. We first look for solutions of the form $v(x) = e^{ikx}$ or $v(x) = \sin(kx)$. In the whole space, this makes sense because the Fourier transform allows us to represent an arbitrary function of the spatial variables x as a superposition of such functions e^{ikx}.

In fact, the functions $v(x) = \sin(kx)$ and $v(x) = \cos(kx)$ are solutions of the Helmholtz equation (1.5) in one dimension and on the whole space, i.e. for $\Omega = \mathbb{R}$. With the *wavenumber* k we have to satisfy $c^2k^2 = \omega^2$, i.e. $k = \pm\omega/c$. It makes sense to keep an eye on the physical units: The unit of ω is Hertz, i.e. $1/s$. The unit of k is $1/m$. In fact, the expressions kx and ωt should always be free of units, because we use them in the sine function, for example. For c, the unit is m/s, i.e. the unit for speeds.

Two other solutions to the same equation in the same situation are $v(x) = e^{ikx}$ and $v(x) = e^{-ikx}$. The complex solutions can be constructed from the real solutions by superposition and vice versa.

If we want to interpret a solution physically, we always have to construct a real function. Here, interesting solutions of the wave equation result from the above examples. If we want to use $v(x) = \sin(kx)$ to construct a solution of the wave equation, we consider $u(x,t) = v(x)e^{-i\omega t}$. If we want to extract a physically relevant function from this, we can, for example, consider the real part, i.e.

$$U(x,t) := \operatorname{Re} u(x,t) = \operatorname{Re} v(x)e^{-i\omega t} = \sin(kx)\cos(\omega t)\,.$$

This describes a standing wave; the "antinodes" are always at the points $\pi/(2k) + \mathbb{Z}\,\pi/k$. The frequency of the oscillation is ω. The spatial structure is determined by k; recall the relation $c^2k^2 = \omega^2$.

If we perform the same process with the solution $v(x) = e^{ikx}$ of the Helmholtz equation, we get $u(x,t) = v(x)e^{-i\omega t} = e^{i(kx-\omega t)} = e^{ik(x-ct)}$ as a solution of the wave equation. Compare this solution with (1.4). A physically relevant function would then be, for example,

$$U(x,t) := \operatorname{Re} u(x,t) = \operatorname{Re} e^{ik(x-ct)} = \cos(k(x-ct))\,.$$

This is a wave moving to the right with velocity c. The solution $v(x) = e^{-ikx}$ of the Helmholtz equation yields a wave moving to the left.

If we consider the wave equation in different materials, we typically obtain equations of the form

$$\rho\partial_t^2 u = \nabla \cdot (a\nabla u)\,, \tag{1.7}$$

where $\rho = \rho(x)$ describes a local density and $a = a(x)$ describes a stiffness of the material. Accordingly, coefficients are obtained in the Helmholtz equation (1.6). The time-homogeneous approach can still be carried out because the material parameters do not depend on time. The result is

$$-\nabla \cdot (a\nabla v) - \omega^2 \rho v = f \quad \text{in } \Omega\,. \tag{1.8}$$

Higher dimension, constant coefficients. Solutions in the whole space $\mathbb{R}^n$ can be constructed by letting solutions depend on only one of the directions. For $n = 3$, for example, $v(x_1, x_2, x_3) = \sin(kx_1)$ is such a planar solution of the Helmholtz equation (1.5); it corresponds to a standing wave of the wave equation. The Helmholtz solution $v(x_1, x_2, x_3) = e^{ikx_1}$ with the time dependence $e^{-i\omega t}$ corresponds to a planar traveling wave in the x_1 direction.

A generalization is to use any wave vector $k \in \mathbb{R}^n$ and look for the solution in the form $v(x) = e^{ik\cdot x}$. The initial condition $\partial_t u|_{t=0} = 0$ is satisfied by $u(x,t) = e^{ik\cdot x}\cos(\omega t)$. For $\omega = c|k|$, u is indeed a solution of the wave equation. By combining such solutions for all $k \in \mathbb{R}^n$, one obtains a solution of the wave equation in $\mathbb{R}^n$ for $u_1 = 0$ and arbitrary u_0

$$u(x,t) = \int_{\mathbb{R}^n} \hat{u}_0(k) e^{ik\cdot x} \cos(c|k|t)\, dk$$

This follows from the theory of the Fourier transform, which states that any u_0 can be represented as a superposition of functions $e^{ik\cdot x}$. This shows that the analysis of $e^{ik\cdot x}$ waves actually yields a lot of information.

The Helmholtz equation is particularly interesting because of resonance phenomena. The simplest example is the Helmholtz resonator; see [11]. Interesting phenomena can be observed in periodic media; we mention [12, 8] here for mathematics.

Exercises

Exercise 1.1 (One-dimensional wave equation)**.** *We consider the wave equation $\partial_t^2 u = c^2\,\Delta u$ in one spatial dimension.*

(a) For any smooth initial data $u(.,0) = u_0$ and $\partial_t u(.,0) = u_1$, find a solution of the form $u(x,t) = \varphi(x+ct) + \psi(x-ct)$ with appropriate φ and ψ.

(b) Show that also for $\varphi \in L^1(\mathbb{R},\mathbb{R})$ the function $u(x,t) = \varphi(x+ct)$ is a distributional solution of the equation in $\mathbb{R} \times (0,\infty)$.

Exercise 1.2 (Wave equation on a bounded domain)**.** *For the domain $\Omega \subset \mathbb{R}^n$, we choose the eigenfunctions $(\psi_k)_k$ of the Laplace operator (Dirichlet boundary data) as the basis of $L^2(\Omega)$. Use it to specify the solution of wave equation on Ω for initial data $u(.,0) = u_0$ and $\partial_t u(.,0) = u_1$ and homogeneous Dirichlet boundary data.*

Part I.

Mathematical tools

2. Fredholm operators and spectral theorem

2.1. Features of Fredholm operators

This section is similar to Section 7 of my script "Nonlinear Analysis" (online). Therefore: For further connections see there.

Fredholm operators are in a way the counterpart to compact operators. While the compact operators are close to zero except for a finite number of components, the Fredholm operators are essentially close to the identity. We consider continuous linear mappings $L : X \to Y$ between Banach spaces, i.e. elements $L \in \mathcal{L}(X, Y)$. We use $\ker(L)$ to denote the operator's kernel, $\ker(L) := \{x \in X | Lx = 0\}$, and $R(L) := L(X) = \{Lx | x \in X\}$ to refer to the operator's image.

Definition 2.1 (Fredholm operator)**.** *An operator $L \in \mathcal{L}(X, Y)$ is called a Fredholm operator if $R(L)$ is closed and*

$$\dim \ker(L) < \infty \qquad \textit{and} \qquad \operatorname{codim} R(L) < \infty\,.$$

The second condition is that there is a decomposition $Y = Y_0 \oplus R(L)$ with finite-dimensional Y_0. The codimension $\operatorname{codim} R(L)$ *is the minimum dimension of Y_0. The integer* $\dim \ker(L) - \operatorname{codim} R(L)$ *is called the index of the Fredholm operator.*

An important result of functional analysis is the following closed complement theorem (see, e.g., [1], Theorem 7.3).

Theorem 2.2 (Closed complement theorem)**.** *Let X be a Banach space, $Y \subset X$ a closed subspace, and $Z \subset X$ a subspace with $X = Y \oplus Z$. Then the following are equivalent:*

(a) There is a continuous projection onto Y with kernel Z, i.e. an element $P \in \mathcal{L}(X)$ with $P \circ P = P$, $R(P) = Y$ and $\ker(P) = Z$.

(b) Z is closed.

One direction in this theorem is trivial, namely (a)$\Rightarrow$(b): If P is continuous, then $Z = P^{-1}(\{0\})$ is closed as the preimage of a closed set.

Regarding the assumption on Y: With a projection P as in (a), the projection $Q = \mathrm{id} - P$ is also continuous, so $Y = \ker Q$ is also closed. Thus, a projection as in (a) can actually only exist for closed subspaces Y.

The interesting part of the theorem is: "If $Y \subset X$ is a closed subspace with a closed complement, then there is a continuous projection onto Y."

For a direct sum, the projections are something very concrete: any element $x \in X$ can be uniquely written as $x = y + z$ with $y \in Y$ and $z \in Z$. The projection P is given by $P(x) = y$.

Comments on Definition 2.1: (i) Let X be a Banach space and X_0 a finite-dimensional subspace. Then there is a continuous projection onto X_0. In particular, X_0 has a closed complement.

Proof of (i): Let $X_0 = \mathrm{span}(e_1, ..., e_n)$ and $\lambda_k : X_0 \to \mathbb{R}$ be the linear functionals with $\lambda_k(e_l) = \delta_{kl}$. By the Hahn–Banach theorem, extend λ_k to continuous linear functionals on X. Then a continuous linear projection from X to X_0 is defined by

$$Px := \sum_{k=1}^{n} \lambda_k(x)\ e_k\,.$$

In particular, for Fredholm operators L and $X_0 = \ker(L)$, the following holds: we can always choose the $X = X_0 \oplus X_1$ decomposition so that the X_1 space is closed.

(ii) For a Fredholm operator L, let $X = X_0 \oplus X_1$ be a decomposition of X with $X_0 = \ker(L)$ and X_1 closed. Then the operator $L_1 := L|_{X_1} : X_1 \to R(L)$ is continuously invertible. In fact, according to the definition of subspaces, the operator is bijective. Because because the spaces are closed, its continuous invertibility follows from the inverse mapping theorem.

(iii) The closedness of the image $R(L)$ does not have to be required in the definition of the Fredholm operator, it follows from the other properties. This is also known as the theorem of Kato.

Proof of (iii): Let $X = X_0 \oplus X_1$, where $X_0 = \ker L$ is finite-dimensional and X_1 is closed. In addition, suppose that $Y = Y_0 \oplus Y_1$ with $Y_1 = R(L)$ and finite-dimensional Y_0. Let's look at the extended operator $\tilde{L} : X_1 \times Y_0 \to Y$, $(x_1, y_0) \mapsto Lx_1 + y_0$. This operator is linear and continuous between Banach spaces. Also, the operator is bijective: $0 = Lx_1 + y_0 \in Y_1 \oplus Y_0$ implies $Lx_1 = 0$ and $y_0 = 0$, and therefore also $x_1 \in \ker(L) \cap X_1 = \{0\}$. Conversely, any element $y = y_1 + y_0 \in Y_1 \oplus Y_0$ has a preimage under $\tilde{L}$; it is given by y_0 and a preimage of $y_1 \in Y_1 = R(L)$. As a bijective linear continuous operator between Banach spaces, $\tilde{L}$ has a continuous inverse $B := \tilde{L}^{-1} : Y \to X_1 \times Y_0$. So

$$R(L) = L(X_1) = \tilde{L}(X_1 \times \{0\}) = B^{-1}(X_1 \times \{0\})$$

is closed.

The following theorem provides a first connection between Fredholm operators and compact operators.

Lemma 2.3 (A Fredholm operator is the identity plus a compact operator)**.** *Let $L : X \to Y$ be a Fredholmoperator with index 0. Then there is a Banach space isomorphism $I : Y \to X$, so that $I \circ L$ is of the form $I \circ L := id + K$ with compact $K \in \mathcal{L}(X)$. Moreover, K can be chosen to have finite-dimensional range.*

Proof. We use $\tilde{L} : X_1 \times Y_0 \to Y$ again, $(x_1, y_0) \mapsto Lx_1 + y_0$ with the continuous inverse $B := \tilde{L}^{-1} : Y \to X_1 \times Y_0$. The first component of the operator $\Phi := (\Phi_1, \Phi_2) := B \circ L : X \to X_1 \times Y_0$, understood as $\Phi_1|_{X_1} : X_1 \to X_1$, is the identity.

The finite-dimensional subspaces X_0 and Y_0 have the same dimension and are therefore isomorphic. With such an isomorphism, $B : Y \to X_1 \times Y_0$ can be composed with an isomorphism on the Y_0 component of the range to obtain a Banach space isomorphism $I : Y \to X_1 \times X_0$. This satisfies what is desired. □

Fredholm alternative

The property of Fredholm operators from Lemma 2.3 can essentially be reversed, so it yields a characterization of Fredholm operators with index 0. Again, the statement only holds up to an isomorphism.

Theorem 2.4 (Identity plus compact is Fredholm)**.** *If $K \in \mathcal{L}(X, X)$ is a compact operator, then $L := id + K$ is a Fredholm operator with index* 0*.*

Proof. See Alt [1], Theorem 8.15. We do not prove the theorem here; for self-adjoint operators, it is a corollary of the spectral theorem 2.8. □

The theorem yields the famous Fredholm alternative as a corollary.

Corollary 2.5 (Fredholm alternative)**.** *Let $\lambda \neq 0$ and let $K \in \mathcal{L}(X, X)$ be compact. Then the following Fredholm alternative holds*

> *Either the equation $(\lambda - K)x = 0$ has a solution $x \neq 0$,*
> *or the operator $(\lambda - K) : X \to X$ is invertible.*

Remark. The alternative also yields: Either the equation $(\lambda - K)(x) = 0$ has a nontrivial solution x or the equation $(\lambda - K)(x) = f$ has a unique solution $x \in X$ for each $f \in X$.

Another equivalent formulation is:

$$\nexists (\lambda - K)^{-1} : X \to X \quad \Longleftrightarrow \quad \exists x \neq 0 : \ (\lambda - K)(x) = 0 \tag{2.1}$$

In this equivalence, the direction "$\Leftarrow$" is trivial; the interesting direction is "$\Rightarrow$".

With this formulation you can see the relationship with a spectral statement: If λ is a spectral value of K, i.e. $(\lambda - K)^{-1}$ does not exist, then λ must be an eigenvalue of K, i.e. an $x \neq 0$ must exist with $(\lambda - K)(x) = 0$.

Proof of Corollary 2.5. According to Theorem 2.4, the operator $L := \lambda - K = \lambda\,\mathrm{id} - K$ is a Fredholm operator with index 0 (a non-zero factor does not change Fredholm properties). If the operator $L = \lambda - K$ is invertible, then the only preimage of zero is zero. Conversely, if $\ker(L) = \{0\}$, then $R(L) = X$ holds according to the definition of the index of Fredholm operators. □

Based on previous considerations about Fredholm operators, the following lemma is no surprise.

Lemma 2.6 (Fredholm plus compact is Fredholm)**.** *Let $L : X \to Y$ be a Fredholm operator and*

$$K : X \to Y \text{ compact,}$$

both linear and continuous. Then the operator

$$L + K : X \to Y$$

is a Fredholm operator. The index of $L + K$ is the same as that of L.

Proof. By adding a finite-dimensional component to one of the Banach spaces X or Y, we can modify any Fredholm operator so that it has index 0. We can therefore assume in the following without loss of generality that L has index 0.

We have to show that the operator $L+K$ is also Fredholm with index 0. Lemma 2.3 yields the existence of an isomorphism $I : Y \to X$, so that $I \circ L = \mathrm{id} + K_1$ with a compact operator $K_1 : X \to X$. The operator $I \circ K : X \to X$ is compact again as a composition of a continuous operator with a compact operator. According to Theorem 2.4, $I \circ (L + K) = I \circ L + I \circ K = \mathrm{id} + K_1 + I \circ K$ is therefore a Fredholm operator with index 0. The isomorphism I does not change the Fredholm properties, so we get the result for $L + K$. □

Example 2.7. *On a bounded domain $\Omega \subset \mathbb{R}^n$, we can consider:*

$$L = \Delta : X = H^2(\Omega) \cap \{u \,|\, \partial_n u|_{\partial\Omega} = 0\} \to Y = L^2(\Omega)\,. \tag{2.2}$$

The image of L is not quite Y, because all $f = Lu$ have the property $\int_\Omega f = 0$. But: $Y = R(L) \oplus \{u = c\}$, the range has a one-dimensional complement.

The kernel of L is nontrivial because constant functions are contained in X, but are mapped to $0 \in Y$. Again, $\ker(L) = \{u = c\}$ is one-dimensional.

Result: The operator L from (2.2) is not invertible, but it is a Fredholm operator with index 0. To put it bluntly: There are finitely many exceptional dimensions in the domain and codomain; apart from these, L is invertible.

2.2. Spectral theorem for compact operators

This section contains material that also appears in Chapter 8.3 of [10]. This is where the tools that we use here are provided.

We consider linear operators $K : X \to X$ on a Banach space X and are interested in the spectrum

$$\sigma(K) := \{\lambda \in \mathbb{C} \,|\, \lambda\,\mathrm{id} - K \text{ is not continuously invertible}\}\,. \tag{2.3}$$

For this lecture, the simplest spectral theorem is sufficient: K is a compact self-adjoint operator on a separable Hilbert space X.

As a reminder, an operator is compact if for each bounded set in the preimage, the closure of the image is compact. The operator K is called self-adjoint if $\langle Ku, v\rangle_X = \langle u, Kv\rangle_X$ for all elements $u, v \in X$.

Theorem 2.8 (Spectral theorem for self-adjoint compact operators)**.** *Let X be an infinite-dimensional separable real or complex Hilbert space, $K : X \to X$ a linear compact self-adjoint operator. Then there exists a countable family of eigenvalues $(\lambda_k)_{k\in\mathbb{N}_*}$ and eigenvectors $(u_k)_{k\in\mathbb{N}_*}$,*

$$Ku_k = \lambda_k u_k, \qquad u_k \in X,\ \lambda_k \in \mathbb{R}\,,$$

such that $(u_k)_{k\in\mathbb{N}_*}$ forms an orthonormal basis of X: We have $\langle u_l, u_k\rangle = \delta_{kl}$ for all $k, l \in \mathbb{N}_*$ and, for real Hilbert spaces X,

$$\forall u \in X\ \exists (b_k)_{k\in\mathbb{N}_*},\ b_k \in \mathbb{R}\ : \qquad u = \sum_{k=1}^{\infty} b_k u_k$$

with convergence of the series in X. The spectrum is $\sigma(K) = \{0\} \cup \{\lambda_1, \lambda_2, \ldots\}$.

The same theorem holds in finite-dimensional spaces, but then the index set is finite and 0 is not necessarily in the spectrum of K. The last two statements also hold in the complex case, except that $b_k \in \mathbb{C}$ in general.

In preparation for the proof, we find that in the case of a complex Hilbert space, because of the self-adjoint nature of K and the (Hermitian) symmetry of the inner product, the equation

$$\langle Ku, u\rangle = \langle u, Ku\rangle = \overline{\langle Ku, u\rangle} \tag{2.4}$$

holds. Thus, $\langle Ku, u\rangle \in \mathbb{R}$ for all $u \in X$. For an eigenvector u with $Ku = \lambda u$, we have

$$\lambda \|u\|^2 = \langle \lambda u, u\rangle = \langle Ku, u\rangle \in \mathbb{R}\,, \tag{2.5}$$

so all eigenvalues are real. Orthogonality: For eigenvectors u and v corresponding to eigenvalues $\lambda, \mu \in \mathbb{R}$, we have

$$\lambda\langle u, v\rangle = \langle Ku, v\rangle = \langle u, Kv\rangle = \bar{\mu}\langle u, v\rangle = \mu\langle u, v\rangle\,. \tag{2.6}$$

If $\lambda \neq \mu$, we therefore have $\langle u, v\rangle = 0$; eigenvectors corresponding to different eigenvalues are orthogonal.

The idea of the following proof is: Minimize a functional on spheres in a Hilbert space, this yields the eigenvectors. Because of the orthogonality of the eigenspaces, we can search for the next eigenvector in the orthogonal complement of the previous eigenspaces.

Proof of the spectral theorem 2.8. We use variational methods.

Step 1: Variational problem with constraints and direct method. We consider the real functional

$$A : X \to \mathbb{R}, \qquad A(u) := \langle Ku, u\rangle\,. \tag{2.7}$$

We consider the functional A on the unit sphere, i.e. we investigate the constraint $\mathcal{N} = \{u \in X | \, \|u\| = 1\}$. We are looking for positive maxima and negative minima of A on $\mathcal{N}$. The functional A is bounded up and down because of the continuity of K. The space X is reflexive because it is a Hilbert space. We use the direct method. Due to the $\mathcal{N}$ constraint, all minimizing sequences are automatically bounded.

We claim that A is weakly sequentially continuous. To do this, we consider any weakly convergent sequence $u_j \rightharpoonup u$ in X. Then the sequence $(u_j)_j$ is bounded in

X. Because of the compactness of K, we find a subsequence and a limit f with $Ku_j \to f$ in X along the subsequence. We claim that $Ku = f$. We obtain this from the following calculation for an arbitrary test vector $\varphi \in X$ for $j \to \infty$,

$$\langle Ku, \varphi \rangle = \langle u, K\varphi \rangle \longleftarrow \langle u_j, K\varphi \rangle = \langle Ku_j, \varphi \rangle \to \langle f, \varphi \rangle .$$

The strong convergence $Ku_j \to f$ in X implies the convergence of $A(u_j)$ because of

$$A(u_j) = \langle u_j, Ku_j \rangle \to \langle u, f \rangle = \langle u, Ku \rangle = A(u) .$$

The lemma without a name implies that the original sequence $A(u_j)$ also converges. This shows the weak continuity of A.

Step 2: Positive maxima and negative minima. It is not true that the unit sphere $\mathcal{N}$ is weakly sequentially closed. We will still find maxima and minima with the direct method.

Consider a minimizing sequence $u_j \in \mathcal{N}$ with $A(u_j) \searrow \inf_{v \in \mathcal{N}} A(v) =: \alpha$ and assume that $\alpha < 0$. Passing to a subsequence, we find a weak limit, i.e. $u_j \rightharpoonup u$ in X, with $\|u\| \leq 1$ because of the weak lower semicontinuity of the norm. We normalize the weak limit u and consider $\tilde{u} = u/\|u\|$ (the weak continuity of A implies $A(u) = \alpha < 0$ and therefore $u \neq 0$). The weak continuity of A and $\alpha < 0$ yields

$$A(\tilde{u}) = \frac{1}{\|u\|^2} A(u) = \frac{1}{\|u\|^2} \alpha \leq \alpha .$$

The left-hand side is greater than or equal to α according to the definition of α; therefore, equality holds throughout and, in particular, $\|u\| = 1$, i.e. $u \in \mathcal{N}$. Thus, we can find negative minima and analogously positive maxima with the direct method. For a minimizer u, the theorem on constrained minimization yields some $\lambda \in \mathbb{R}$ such that

$$Ku = \lambda u, \qquad \|u\| = 1$$

Then $0 > \alpha = A(u) = \langle u, Ku \rangle = \lambda \|u\|^2 = \lambda$, so we have constructed an eigenvector corresponding to the negative eigenvalue $\lambda = \alpha$. Analogously, a maximizer of A, we get the positive eigenvalue $\lambda = \alpha$ when $\alpha = \sup_{v \in \mathcal{N}} A(v)$ is positive.

Step 3: Dimension of eigenspaces and orthogonal complements. We claim that for any $\lambda \neq 0$, the eigenspace is finite-dimensional

$$X_\lambda := \{u \in X \mid Ku = \lambda u\} = \ker(\lambda \, \mathrm{id} - K)$$

. We prove this by showing that the unit sphere in X_λ is compact. To do this, let u_j be any sequence in X_λ with $\|u_j\| \leq 1$ for all $j \in \mathbb{N}$. Because of the compactness of K, we find a subsequence and $f \in X$ with $Ku_j \to f$ in X. This implies the strong convergence of $u_j = \lambda^{-1} K u_j \to \lambda^{-1} f$ along the subsequence. This shows the compactness of the sphere, X_λ is finite-dimensional.

We now consider for finitely many eigenvalues $0 \neq \lambda_k \in \mathbb{R}$, $k = 1, ..., m$, the corresponding eigenspaces X_{λ_k} and the span of these spaces, $Y_m := \mathrm{span}\{X_{\lambda_1}, ..., X_{\lambda_m}\}$. We refer to the orthogonal complement of Y_m in X as $Y_m^\perp$. This space is closed and is therefore a Hilbert space. We can restrict the operator K to $Y_m^\perp$; we claim that this gives an operator $K|_{Y_m^\perp} : Y_m^\perp \to Y_m^\perp$. The relation to be proven is $Ky \in Y_m^\perp$ for $y \in Y_m^\perp$. For any $\varphi \in X_{\lambda_k}$, this follows from the calculation:

$$\langle Ky, \varphi\rangle = \langle y, K\varphi\rangle = \lambda_k \langle y, \varphi\rangle = 0\,.$$

We now proceed inductively. For eigenvalues $0 \neq \lambda_k \in \mathbb{R}$, $k = 1, ..., m$, we consider $\alpha_+ := \sup_{v\in\mathcal{N}\cap Y_m^\perp} A(v)$ and $\alpha_- := \inf_{v\in\mathcal{N}\cap Y_m^\perp} A(v)$. In the case of $|\alpha_+| \geq |\alpha_-|$, we are looking for a maximum of $A(v)$ in $\mathcal{N} \cap Y_m^\perp$, in the opposite case a minimum. We get a new eigenvalue λ_{m+1} (unless minimum and maximum vanish; in which case all eigenvalues are found in finitely many steps). The construction yields $|\lambda_{m+1}| \leq |\lambda_m|$, where equality only occurs if $\lambda_{m+1} = -\lambda_m$.

Step 4: The sequence of eigenvalues. We claim that the sequence of eigenvalues found in this way satisfies

$$\lambda_m \to 0 \quad \text{for} \;\; m \to \infty$$

. This follows immediately by a contradiction argument: Suppose that, for a subsequence, $\lambda_m \to \lambda \neq 0$, where all λ_m are different. Then we can choose normalized eigenvectors u_m with $Ku_m = \lambda_m u_m$. Because of the compactness of K and $\lambda \neq 0$, we have $u_m = \lambda_m^{-1} K u_m \to \lambda^{-1} f$. On the other hand, for $m \neq n$, orthogonality gives $\|u_m - u_n\|^2 = \|u_m\|^2 + \|u_n\|^2 = 2$, a contradiction. In particular, we have shown that there are only countably many eigenvalues.

We claim that the following norm estimate also holds: For every $\varepsilon > 0$ there is an index $m \in \mathbb{N}$ such that, for $Y_m = \operatorname{span}\{X_{\lambda_1}, ..., X_{\lambda_m}\}$, the following holds:

$$K|_{Y_m^\perp} : Y_m^\perp \to Y_m^\perp \quad \text{has norm} \quad \left\|K|_{Y_m^\perp}\right\| < \varepsilon\,. \tag{2.8}$$

In the case of $\alpha_+ > 0$, the number $\alpha_+ := \sup_{u\in\mathcal{N}\cap Y_m^\perp} A(u)$ is an eigenvalue of $K|_{Y_m^\perp}$ after step 2, i.e. also an eigenvalue of K. By construction, $\alpha_+ \leq |\lambda_m|$, and analogously $\alpha_- \geq -|\lambda_m|$ for the infimum. We exploit the self-adjoint nature of K to establish a parallelogram identity (noted here for real Hilbert spaces):

$$4\langle Ku, v\rangle = \langle K(u+v), (u+v)\rangle - \langle K(u-v), (u-v)\rangle = A(u+v) - A(u-v)\,. \tag{2.9}$$

We consider a normalized vector $u \in Y_m^\perp$. We decompose any test vector $v \in X$ as $v = v_\parallel + v_\perp \in Y_m \oplus Y_m^\perp$, so that $\langle Ku, v\rangle = \langle u, Kv_\parallel\rangle + \langle Ku, v_\perp\rangle = \langle Ku, v_\perp\rangle$. We get the inequality $\|Ku\| \leq C \sup_{\varphi\in\mathcal{N}\cap Y_m^\perp} |A(\varphi)| \leq C|\lambda_m|$ from (2.9), and thus (2.8).

We now want to construct an orthonormal basis of X from eigenvectors. In each of the finite-dimensional eigenspaces X_λ, $\lambda \neq 0$, we choose an orthonormal basis. Because of the orthogonality of different eigenspaces, we get $\langle u_k, u_l\rangle = \delta_{kl}$ for all k, l by normalization. Finally, we can choose an orthonormal basis of the (generally infinite-dimensional) space $X_0 = \ker K \subset X$, since X is separable. These eigenvectors corresponding to the eigenvalue 0 are in turn orthogonal to the previously constructed eigenvectors.

In the formulation of the theorem, we used a sequence of eigenvectors and did not distinguish between the kernel X_0 and its complement. In order to obtain eigenvalues as in the theorem, the nontrivial λ_k found here may have to be partially repeated (according to the dimension of the corresponding eigenspace) and eigenvalues 0 must be inserted into the sequence (for example, the even indices k could be used to traverse the basis of X_0 and the odd indices k the bases of all other eigenspaces). The index k now refers to this new count of eigenvalues, the orthonormality $\langle u_k, u_l\rangle = \delta_{kl}$ is preserved.

We note at this point that $0 \in \sigma(K)$ necessarily holds (although the triviality of kernel $X_0 = \{0\}$ is possible). In fact, we could otherwise write the identity on X as $\mathrm{id} = K^{-1} \circ K$; the identity would then be a compact mapping. This is only the case in finite-dimensional space.

Step 5: Completeness. We now have to show the completeness of the constructed family, i.e. for any $v \in X$ the existence of coefficients $b_k \in \mathbb{R}$ with $v = \sum_k b_k u_k$ (convergence in X). With the help of our normalized basis vectors u_k, we can obtain the coefficients b_k,

$$b_k := \langle v, u_k\rangle, \qquad P_M v := \sum_{k=1}^{M} b_k u_k, \qquad e^M := v - P_M v.$$

The expression $P_M v = \sum_{k=1}^{M} b_k u_k$ yields the projection of v onto the span of the first M eigenvectors; after choosing the coefficients, especially

$$\|P_M v\|_X^2 = \sum_{k=1}^{M} |b_k|^2 = \langle P_M v, v\rangle \le \|P_M v\|_X \, \|v\|_X \,,$$

we also have $\|P_M v\|_X \le \|v\|_X$. We need to show the convergence of the error, $e^M \to 0$ in X. Since we have included a basis of the kernel in the $(u_k)_k$ family, it is sufficient in the following to consider elements of $v \in X_0^\perp$ in the complement of the kernel.

The error $e^M = v - P_M v$ satisfies $e^M \in Y_M^\perp$. Since the projection is bounded, $(e^M)_M$ is a bounded sequence. We choose a weakly convergent subsequence $e^M \rightharpoonup e$ in X. Due to the compactness of K, we have $Ke^M \to Ke$ in X. Moreover, by (2.8),

$$\|Ke^M\|_X \le \left\| K|_{Y_M^\perp} \right\| \, \|e^M\|_X \to 0\,,$$

$Ke = 0$. The relation $e^M \in X_0^\perp$ for all M implies $e \in X_0^\perp$ (weak closedness of closed subspaces), so $e = 0$ must hold. Thus, the errors e^M converge at least weakly to 0 in X.

The strong convergence of errors follows with the calculation

$$\|e^M\|_X^2 = \left\langle e^M, v - \sum_{k=1}^{M} b_k u_k \right\rangle = \langle e^M, v\rangle - \sum_{k=1}^{M} b_k \langle e^M, u_k\rangle = \langle e^M, v\rangle \to 0$$

for $M \to \infty$. This shows the completeness of the u_k eigenvectors.

We still have to show $\sigma(K) = \{0\} \cup \{\lambda_1, \lambda_2, \ldots\}$, although the inclusion "$\supset$" has already been shown. We use a simple spectral statement: For all $\mu \in \sigma(K)$, the relation $|\mu| \le \|K\|$ holds. This follows immediately with the Neumann series: For all $\lambda \in \mathbb{C}$ with $|\lambda| > \|K\|$, the Neumann series converges and yields the inverse of the $\lambda - K$ operator. In our construction above, this means that any $0 \ne \lambda \in \mathbb{C}$ for sufficiently large M is not in the spectrum of the operator $K|_{Y_M^\perp}$. This yields "$\subset$". $\square$

Exercises

Exercise 2.1 (Decomposition with a non-closed subspace)**.** *We consider the Banach space* $X = \ell^2(\mathbb{N})$ *and the (one-dimensional and therefore closed) subspace* $Y_0 = \mathrm{span}\{e_1\}$*. For a non-continuous linear functional* $\lambda : X \to \mathbb{R}$*, consider*

$$Y := \{x = (x_1, x_2, x_3, ...) \in X \mid x_1 = \lambda((x_2, x_3, ...))\} .$$

Show that (a) Y *is a linear subspace of* X*, (b)* Y *is not closed in* X*, and (c)* $X = Y_0 \oplus Y$*.*

Note: For Fredholm operators L*,* $R(L)$ *is automatically closed (theorem of Kato). The exercise shows that this does not follow solely from the fact that there is a finite-dimensional complement.*

Exercise 2.2 (Spectral theorem for non-normal operators)**.** *Formulate a spectral theorem for compact operators* $K : X \to X$ *on a Hilbert space* X *without the requirement that* K *is normal (keyword: theorem of Riesz–Schauder). Use matrices to think about which statements of the spectral theorem are not universally valid. Use the material of the lecture to think about which statements of the spectral theorem remain valid.*

Exercise 2.3 (Special spectral statements)**.** *Let* X *be a Hilbert space and* $K : X \to X$ *compact and normal. Show that* K *has an eigenvalue* $\lambda \in \mathbb{C}$ *with* $|\lambda| = \|K\|$*. This number corresponds to the spectral radius.*

Let X *be a Banach space. Show that every nilpotent operator* $T : X \to X$ *(i.e.* $T^k = 0$ *for some* $k \in \mathbb{N}$*) has the spectrum* $\sigma(T) = \{0\}$*.*

Exercise 2.4 (Spectral theorem and compactness)**.** *Let* $K : X \to X$ *be a normal operator on a Hilbert space. Assume further that* 0 *is the only accumulation point of* $\sigma(K)$ *and for all* $\lambda \in \sigma(K) \setminus \{0\}$*, the generalized eigenspace* $\bigcup_{j\in\mathbb{N}} \ker((\lambda - K)^j)$ *is finite-dimensional. Show that* K *is compact. Show that without the assumption of normality the statement is false.*

3. Laplace–Beltrami operator and spherical harmonics

3.1. The Laplace–Beltrami operator

This section contains material that is included in Chapter 8.3 of [10].

In this section we consider the manifold $S := S^{n-1} = \partial B_1(0) \subset \mathbb{R}^n$. However, the concepts presented can be generalized to arbitrary submanifolds.

The tangential gradient

We first define the tangential gradient. Let $u : S \to \mathbb{R}$ be a continuously differentiable mapping. We denote the tangent space to S at the point x by T_xS; it can be identified with an $(n-1)$-dimensional subspace of $\mathbb{R}^n$, namely the span of $n-1$ tangent vectors $\tau_1, ..., \tau_{n-1}$. Tangential vectors $\tau \in T_xS$ are those vectors $\tau \in \mathbb{R}^n$ for which there are differentiable paths $\gamma : (-\varepsilon, \varepsilon) \to S$ with $\gamma(0) = x$ and $\gamma'(0) = \tau$. For the tangent space T_xS, we use $P_x : \mathbb{R}^n \to \mathbb{R}^n$ to denote the orthogonal projection onto T_xS.

An elementary definition of the tangential gradient can be given with extensions of u. Let $\tilde{u} : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable extension of u, i.e. $\tilde{u} = u$ on S (it is sufficient if $\tilde{u}$ is defined only in a neighborhood of S; you can choose $\tilde{u}(x) = u(\frac{x}{|x|})$ on the sphere, for example). With the projection P_x onto the tangent space (given on the sphere by $P_x(\xi) = \xi - \langle \xi, \frac{x}{|x|}\rangle \frac{x}{|x|}$), we define

$$\nabla_T u(x) := P_x(\nabla \tilde{u}(x)) \,. \tag{3.1}$$

If we are not given a surrounding space, another definition is necessary. We then start from the derivative $Du(x) : T_xS \to \mathbb{R}$; this can generally be defined via the derivatives of $u \circ \gamma$ for paths γ. With the help of an inner product $\langle ., . \rangle$ on the tangent space and Riesz's representation theorem, we now represent $Du(x)$ by a vector $g \in T_xS$:

$$\text{if } Du(x)(\tau) = \langle g, \tau \rangle \ \ \forall \tau \in \mathbb{R}^{n-1}\,, \qquad \text{then define } \nabla_T u(x) := g. \tag{3.2}$$

We still want to see that the two definitions of $\nabla_T u(x)$ actually give the same result. We use a basis $\tau_1, ..., \tau_{n-1}$ of T_xS, and we refer to the associated paths as $\gamma_1, ..., \gamma_{n-1}$. By the chain rule,

$$\langle g, \tau_j \rangle = Du(x)(\tau_j) = \frac{d}{d\varepsilon}(u \circ \gamma_j)\Big|_{\varepsilon=0} = \frac{d}{d\varepsilon}(\tilde{u} \circ \gamma_j)\Big|_{\varepsilon=0} = \nabla \tilde{u}(x) \cdot \tau_j \,.$$

The normal component of g vanishes because of the requirement $g \in T_xS$. Therefore, the two definitions (3.1) and (3.2) of the tangential gradient match.

$H^1(S)$ and Laplace–Beltrami operator

For functions $u \in C^1(S, \mathbb{R})$, we can consider the norm

$$\|u\|^2_{H^1(S)} := \int_S |u|^2 + |\nabla_T u|^2 \tag{3.3}$$

and define the space $H^1(S)$ as the completion of $C^\infty(S, \mathbb{R})$ with respect to the norm $\|.\|_{H^1(S)}$. On the space $H^1(S)$, we then define a bilinear form by

$$a : H^1(S) \times H^1(S) \to \mathbb{R},\ a(u,v) := \int_S \nabla_T u \cdot \nabla_T v\,. \tag{3.4}$$

For each $u \in H^1(S)$ a linear form $\lambda \in H^{-1}(S) := H^1(S)'$ is given by $\lambda : H^1(S) \to \mathbb{R}$, $v \mapsto -a(u,v)$. This linear form is called $\Delta_S u$. If the linear form (due to the regularity of u) is even a continuous linear form on $L^2(S)$, then $\Delta_S u$ can be represented by an $L^2(S)$ function, its application to a test function v is then given by the $L^2(S)$ inner product. In this case, the following more suggestive notation is justified.

$$-\Delta_S : H^1(S) \to H^{-1}(S), \qquad \int_S (-\Delta_S u)\, v = \int_S \nabla_T u \cdot \nabla_T v\,. \tag{3.5}$$

The operator $\Delta_S : H^1(S) \to H^{-1}(S)$ is called the Laplace–Beltrami operator.

Connection to the full Laplace operator

We use spherical coordinates and write a point $x \in \mathbb{R}^n \setminus \{0\}$ as $x = r\xi$ with $r = |x|$ and $\xi \in S$. In a point $x \neq 0$, we use the unit vector in the radial direction, $e_r := \frac{x}{|x|}$. Our goal is to decompose the Laplace operator into a spherical part and a radial part. The radial part should contain only derivatives with respect to the radius r. We define

$$\begin{aligned} &\Delta_r : C^2((0,\infty),\mathbb{R}) \to C^0((0,\infty),\mathbb{R})\,, \quad w \mapsto \Delta_r w\,,\\ &\Delta_r w(r) = \frac{1}{r^{n-1}}\partial_r(r^{n-1}\partial_r w)(r)\,. \end{aligned} \tag{3.6}$$

Lemma 3.1 (Representation of the Laplace operator with the Laplace–Beltrami operator)**.** *For $u \in C^2(\mathbb{R}^n)$, the following identity holds at every point $x_0 \in \mathbb{R}^n \setminus \{0\}$:*

$$\Delta u = \Delta_r u + \frac{1}{r^2}\Delta_S u\,. \tag{3.7}$$

The formula (3.7) can be read as follows: In a point $x_0 = r_0\xi_0$, Δu is evaluated at the point x_0 on the left. On the right, in the first term, the operator Δ_r is applied to the function $r \mapsto u(r\xi_0)$ and the result is evaluated in r_0; in the second term, the Δ_S operator is applied to the $\xi \mapsto u(r_0\xi)$ function and the result is evaluated in ξ_0.

The formula (3.7) shows that the sum of all second derivatives can also be separated into second radial derivatives and second tangential derivatives with respect to the polar coordinates. However, due to the curvilinear coordinate system in both directions, the first derivatives appear, in the radial part relatively explicitly, in the tangential part rather implicitly via the Laplace-Beltrami operator.

Proof. To prove (3.7), an integration by parts with a test function $\varphi \in C_c^\infty(\mathbb{R}^n \setminus \{0\}, \mathbb{R})$ is sufficient. We calculate that both sides of (3.7) act on the test function φ in the same way. We use successively: (a) integration by parts in $\mathbb{R}^n$, (b) the formula for integration over spherical shells, (c) the decomposition of the gradient into a radial and a tangential part, $\nabla u = \partial_r u e_r + \nabla_T u$. In the second line of the calculation (and only there) $\nabla_T u$ denotes the tangential gradient on the sphere rS of radius r; this is calculated with the chain rule as $(\nabla_T u)(r\xi) = r^{-1}\nabla_T(u(r\,.))(\xi)$, which we exploit in (d). In addition, in (d) the transformation theorem with the volume element r^{n-1} is used, in (e) an integration by parts with respect to ∂_r on $(0,\infty)$ and the definition of Δ_S, in (f) another integration over spherical shells.

$$
\begin{aligned}
-\int_{\mathbb{R}^n} \Delta u \cdot \varphi &\overset{(a)}{=} \int_{\mathbb{R}^n} \nabla u \cdot \nabla \varphi \overset{(b)}{=} \int_0^\infty \int_{rS} \nabla u \cdot \nabla \varphi \, d\mathcal{H}^{n-1}\, dr \\
&\overset{(c)}{=} \int_0^\infty \int_{rS} \partial_r u \cdot \partial_r \varphi + \nabla_T u \cdot \nabla_T \varphi \, d\mathcal{H}^{n-1}\, dr \\
&\overset{(d)}{=} \int_0^\infty \int_S \partial_r u(r\xi) \cdot \partial_r \varphi(r\xi) d\mathcal{H}^{n-1}(\xi)\, r^{n-1}\, dr \\
&\quad + \int_0^\infty \int_S r^{-1}\nabla_T(u(r\,.))(\xi) r^{-1} \nabla_T(\varphi(r\,.))(\xi) d\mathcal{H}^{n-1}(\xi)\, r^{n-1}\, dr \\
&\overset{(e)}{=} -\int_0^\infty \int_S \frac{1}{r^{n-1}} \partial_r (r^{n-1}\partial_r u(.\,\xi))(r) \varphi(r\xi) d\mathcal{H}^{n-1}(\xi)\, r^{n-1}\, dr \\
&\quad - \int_0^\infty \int_S \frac{1}{r^2} \Delta_S u(r\,.)(\xi) \varphi(r\xi) d\mathcal{H}^{n-1}(\xi)\, r^{n-1}\, dr \\
&\overset{(f)}{=} -\int_{\mathbb{R}^n} \left(\Delta_r u + \frac{1}{r^2}\Delta_S u\right) \varphi\,.
\end{aligned}
$$

Since φ was arbitrary, the claimed formula (3.7) is proved. □

3.2. Spherical harmonics

The operator Δ_S is not invertible, because constant functions on the sphere are mapped to the constant function 0. We divide the constant functions out of the space and use the space $L_0^2(S) := \{u \in L^2(S) | \int_S u = 0\}$ with the norm of $L^2(S)$. Then the mapping into the dual space is

$$
\Delta_S : H^1(S) \cap L_0^2(S) \to \left[H^1(S) \cap L_0^2(S)\right]'
$$

invertible. This follows from Riesz's representation theorem with a from (3.4). The mapping $K := \Delta_S^{-1} : L_0^2(S) \to L_0^2(S)$ is compact and self-adjoint.

The spectral theorem 2.8 for self-adjoint compact operators yields: The space $L_0^2(S)$ is spanned by eigenvectors of K, i.e. by eigenvectors of Δ_S. There are $\mu_k \in \mathbb{R} \setminus \{0\}$, $k \geq 1$, and solutions $u_k \in L^2(S)$ with

$$
\Delta_S u_k = -\mu_k \, u_k\,.
$$

By adding the constant function $u_0 \equiv 1$ (a function with $\Delta_S u_0 = 0$, i.e. an eigenvector corresponding to $\mu_0 = 0$), we get a basis of $L^2(S)$. The eigenfunctions u_k are called spherical harmonics.

With the help of spherical harmonics, we can construct many harmonic functions: We use the separation ansatz $u(r,\xi) = w(r)u_k(\xi)$ and calculate

$$\Delta u(x) = (\Delta_r w)(r)u_k(\xi) + w(r)\frac{1}{r^2}\Delta_S u_k(\xi) = \left[\Delta_r w(r) - \frac{\mu_k}{r^2}w(r)\right] u_k(\xi).$$

The function u is harmonic if w solves the ordinary differential equation $\Delta_r w(r) = \mu_k r^{-2} w(r)$. For $q \in \mathbb{Z}$, we set $w(r) = r^q$ and calculate $\Delta_r w(r) = r^{-(n-1)}\partial_r(r^{n-1}\partial_r w)(r) = qr^{-(n-1)}\partial_r(r^{n-1+q-1}) = q(q+n-2)\, r^{-2}\, w(r)$. In fact, the eigenvalues μ_k of $-\Delta_S$ are of the form $\mu_k = q(q+n-2)$ and spherical harmonics provide harmonic functions with the corresponding factor r^q. This allows us to specify a harmonic extension in the form of a series for arbitrary boundary values on the sphere.

Eigenfunctions of $-\Delta$ on subsets of $\mathbb{R}^n$ can also be found with spherical harmonics. The equation $-\Delta u = \lambda u$ is solved by the function $u(r,\xi) = w(r)u_k(\xi)$ if w solves the equation

$$-\Delta_r w(r) + \frac{\mu_k}{r^2}w(r) = \lambda w(r)$$

. This equation is a form of Bessel's differential equation.

In fact, it can be shown that this method (based on a separation ansatz with spherical harmonics) provides a complete system of eigenfunctions of the Laplace operator. This explains the frequent occurrence of spherical harmonics in solutions of different equations. The best known are the eigenfunctions related to the Schrödinger operator: The s orbital of electron distributions in the atom corresponds to the constant spherical harmonic u_0, the dumbbell-shaped p orbital to the spherical harmonic with two opposing maxima.

Exercises

Exercise 3.1 (Spherical harmonics in two dimensions)**.** *Consider $n = 2$. Specify Δ_r and Δ_S and all spherical harmonics. For all spherical harmonics, specify the harmonic functions that have the same angular dependence. Make a connection to holomorphic functions.*

Exercise 3.2 (Laplace–Beltrami operator in three dimensions)**.** *We consider $n = 3$ and the coordinates*

$$x = \begin{pmatrix} r\cos(\varphi)\sin(\theta) \\ r\sin(\varphi)\sin(\theta) \\ r\cos(\theta) \end{pmatrix}.$$

Show that in three dimensions the Laplace–Beltrami operator satisfies the formula

$$\Delta_S = \frac{1}{\sin^2(\theta)}\partial_\varphi^2 + \frac{1}{\sin(\theta)}\partial_\theta(\sin(\theta)\partial_\theta)\,.$$

Part II.

Helmholtz equation with constant coefficients

4. Bounded domains

4.1. Spectrum of elliptic operators

The spectral theorem 2.8 makes a statement about bounded (even compact) operators K. Thus, it is not directly applicable to an elliptic operator, because it is not compact, not even bounded (as mapping from a space into itself). We obtain a spectral statement by applying the theorem to the inverse of the differential operator.

We consider the Helmholtz equation in a bounded domain $\Omega \subset \mathbb{R}^n$. Let the coefficients $a \in L^\infty(\Omega, \mathbb{R}^{n\times n})$ be symmetric and coercive, $b \in L^\infty(\Omega, \mathbb{R})$ with a positive lower bound. Consider the operator $L : H_0^1(\Omega) \to H^{-1}(\Omega)$,

$$L : u \mapsto -\nabla \cdot (a\nabla u) \ , \tag{4.1}$$

defined by

$$\langle Lu\rangle(\varphi) = \int (a(x)\nabla u(x)) \cdot \nabla\varphi(x)\, dx \qquad \forall \varphi \in H_0^1(\Omega)\,.$$

Under our assumptions (a and b positive), the operator $L + b$ is invertible; there is a solution operator $(L + b)^{-1} : H^{-1} \to H_0^1$, which is linear and continuous. In particular, after restricting the operator, we have

$$(L + b)^{-1} : L^2(\Omega) \to H_0^1(\Omega) \text{ linear and continuous.}$$

If we now use the compactness of the Rellich embedding, we get the operator

$$(L + b)^{-1} : L^2(\Omega) \to L^2(\Omega) \text{ compact.}$$

In order for the statement from the spectral theorem to be useful for us, we still have to incorporate the multiplication by the coefficient b in a suitable way. We look at

$$K := (L + b)^{-1} \circ (b\,\mathrm{id}) : L^2(\Omega) \to L^2(\Omega) \text{ compact.}$$

We equip the space $L^2(\Omega)$ with the inner product

$$(u, v) \mapsto \int_\Omega b(x)u(x)\, v(x)\, dx$$

. Because of the boundedness of b and the positive lower bound, this inner product induces an equivalent norm on $L^2(\Omega)$; we make $L^2(\Omega)$ a Hilbert space with a new inner product and write X for this space.

In X, the operator $(L + b)^{-1} \circ (b\,\mathrm{id})$ is self-adjoint. To see this, we calculate for arbitrary $f, g \in X$ and the associated “solutions” $u = K(f) = (L + b)^{-1}(b \cdot f)$ and $v = K(g) = (L + b)^{-1}(b \cdot g)$:

$$\langle Kf, g\rangle_X = \langle u, b^{-1}\,(L + b)(v)\rangle_X = \langle u, (L + b)(v)\rangle_{L^2}\,.$$

The symmetry of L and $b\,\mathrm{id}$ in L^2 yields that this is equal to $\langle (L+b)(u), v\rangle_{L^2}$, and this is (with the same calculation) $\langle f, Kg\rangle_X$. So the operator K is self-adjoint in X.

The spectral theorem 2.8 for self-adjoint compact operators yields: The space $L^2(\Omega)$ is spanned by the eigenvectors u_k of $K = (L+b)^{-1} \circ (b\,\mathrm{id}) : L^2(\Omega) \to L^2(\Omega)$. Thus, there are $\mu_k \in \mathbb{R}$ and solutions $u_k \in L^2(\Omega)$ with

$$(L+b)^{-1}(bu_k) = \mu_k\, u_k\,.$$

The operator is positive, so $\mu_k > 0$. We can apply the operator and get

$$\mu_k\,(L+b)u_k = bu_k\,.$$

By forming the inner product (in L^2) with u_k and using $\langle u_k, Lu_k\rangle > 0$, we get $0 < \mu_k < 1$. We get

$$Lu_k = \left(-1 + \frac{1}{\mu_k}\right) bu_k =: \lambda_k\, bu_k$$

with $\lambda_k = -1 + \mu_k^{-1} > 0$. The functions u_k are therefore eigenfunctions of the generalized eigenvalue problem $Lu_k = \lambda_k bu_k$. The eigenvalues are $\lambda_k = -1+\mu_k^{-1} > 0$ with $\lambda_k \to +\infty$ for $k \to \infty$. So we have found eigenfunctions for the differential operator with the help of the spectral theorem for compact self-adjoint operators.

4.2. Helmholtz equation in the bounded domain

In the bounded domain $\Omega \subset \mathbb{R}^n$, we consider the following inhomogeneous Helmholtz equation, a generalization of (1.6):

$$-\nabla \cdot (a\nabla u) - \omega^2 b\, u = f \quad \text{ in } \Omega\,. \tag{4.2}$$

Let the coefficients a and b be as in the previous section; in particular, both are positive. Let $\omega > 0$ be a frequency and $f \in L^2(\Omega)$ a right-hand side.

Fredholm alternative. We use the notation of the last section. We first insert a term into the equation and write it equivalently as

$$-\nabla \cdot (a\nabla u) + b\, u = (1+\omega^2)\, b\, u + f\,.$$

Now we apply the compact operator $(L+b)^{-1}$, which gives us u on the left. Rearranging gives

$$u - (1+\omega^2)\,(L+b)^{-1}(b\, u) = (L+b)^{-1}(f)\,.$$

This problem for u is of the form $(\mathrm{id} + \mathcal{K})(u) = g$ for a compact operator $\mathcal{K}$. Thus, the Fredholm alternative applies: Either there are solutions u for all f or there are non-trivial solutions to $f = 0$.

Additional statements with the spectral theorem. The two possibilities can be characterized more precisely. Which one is present depends on the frequency ω.

Option 1: We have $\omega^2 \neq \lambda_k$ for all $k \in \mathbb{N}$. The number ω^2 is therefore not a spectral value. Then the operator $L - \omega^2\, b$ is invertible. We get the unique solvability of the equation for all f.

Option 2: We have $\omega^2 = \lambda_k$ for some $k \in \mathbb{N}$. In this case, there is $u_k \neq 0$ with $Lu_k = \lambda_k b u_k$. This yields a non-trivial solution for (4.2) with $f = 0$. In particular, the problem of (4.2) cannot be solved uniquely. The Fredholm alternative yields an additional piece of information: There is actually a right-hand side f such that the equation is not solvable.

Exercises

Exercise 4.1 (Spectral theorem with mass operator)**.** *The spectral theorem yields eigenvalues λ_k and eigenvectors u_k with $Ku_k = \lambda_k u_k$. For a linear bounded self-adjoint invertible and positive operator $B : X \to X$, we want to search for eigenvalues and eigenvectors with the generalized relation $Ku_k = \lambda_k B u_k$. How can we adapt the proof of the spectral theorem to obtain the same statements with the "mass operator" as for $B = id$?*

5. Exterior domains

5.1. Rellich's lemma

For a spectral parameter $\lambda > 0$ and a radius $R > 0$, we consider solutions of the equation

$$(\Delta + \lambda)u = 0 \tag{5.1}$$

in the domain $\Omega := \mathbb{R}^n \setminus B_R(0)$. The dimension $n \in \mathbb{N}$ is arbitrary with $n \geq 2$.

Even if we set a boundary condition along $\partial B_R(0)$, we can't expect solutions to be unique; we also have to specify a "boundary condition for $|x| \to \infty$", a so-called radiation condition. With a radiation condition, uniqueness can be ensured.

Uniqueness can be inferred from Rellich's lemma. The lemma even shows the following: lemma: If the solution decays very quickly, then it must vanish.

Theorem 5.1 (Supercritical decay of Helmholtz solutions)**.** *If u is a solution of the Helmholtz equation* (5.1) *in the domain $\Omega := \mathbb{R}^n \setminus B_R(0)$ with satisfying the decay condition*

$$\int_{\partial B_\rho} |u|^2 \to 0 \qquad \textit{for } \rho \to \infty\,. \tag{5.2}$$

Then u vanishes identically.

Theorem 5.1 follows directly from Lemma 5.3 below, the famous Rellich lemma. The theorem is very surprising because we don't prescribe a boundary condition on $\partial B_R(0)$.

We mention here what is to be recalculated in Exercise 5.1: In the three-dimensional domain $\mathbb{R}^3 \setminus B_1(0)$, with $r = |x|$, the function

$$u(x) := \frac{e^{i\omega r}}{r} \tag{5.3}$$

a solution of the Helmholtz equation (5.1) for $\lambda = \omega^2$.

We notice that the solution in (5.3) does not satisfy the condition in (5.2). This is necessary by Theorem 5.1.

Another remark: Since we want to allow u from (5.3) as a solution in an exterior domain, we must never prescribe (5.2) as a boundary condition.

Remark 5.2 (Regularity of u)**.** *In Theorem 5.1 and Lemma 5.3, we do not specify a function space for the solution u. The statements are of course correct for $u \in C^2(\Omega)$. But it is enough to require that $u \in L^1_{loc}(\Omega)$ is a distributional solution of the equation. Then, because of elliptic regularity theory, u is automatically a smooth function.*

In Theorem 5.7, it must be ensured that boundary values are defined. In the Dirichlet problem, for example, one could demand: For each test function $\psi : \mathbb{R}^n \to \mathbb{R}$ of class C^1 with compact support, $\psi\, u \in H^1_0(\Omega)$.

Rellich's lemma: formulation and proof. We essentially use Franz Rellich's original formulation from 1943, formulated and proved in [9].

Lemma 5.3 (Rellich lemma)**.** *We look at parameters* $\lambda, R > 0$ *and a dimension* $n \in \mathbb{N}$ *with* $n \geq 2$*. A mapping* $u : \Omega = \mathbb{R}^n \setminus B_R(0) \to \mathbb{R}$ *is a solution of* (5.1)*. If* u *does not vanish identically, then constants* $r_1 > R$ *and* $c_R > 0$ *exist with*

$$\int_{B_\rho \setminus B_R} |u|^2 \geq c_R\, \rho \qquad \forall\ \rho \geq r_1\,. \tag{5.4}$$

Proof. Step 1: Reduction to a function v *of one variable.* Let $u : \Omega = \mathbb{R}^n \setminus B_R(0) \to \mathbb{R}$ be a non-trivial solution of (5.1). For a radius $r_0 > R$, consider the restriction of u to the sphere $\{x \in \mathbb{R}^n \mid |x| = r_0\}$. By choosing r_0 appropriately, we can assume that u does not vanish on this sphere. The spherical harmonics form a basis of $L^2(S^{n-1})$. Specifically, we find a spherical harmonic Y, normalized to have norm 1 in $L^2(S^{n-1})$, such that

$$\int_{|y|=1} u(r_0 y)\, Y(y)\, dS(y) \neq 0\,. \tag{5.5}$$

We define a function $v : [R, \infty) \to \mathbb{R}$ by

$$v(r) := \int_{|y|=1} u(ry)\, Y(y)\, dS(y)\,, \tag{5.6}$$

and (5.5) implies $v(r_0) \neq 0$.

Lemma 5.3 is proved once v, for some parameters $r_1 > r_0$ and $c_R > 0$ satisfies:

$$\int_R^\rho v(r)^2\, r^{n-1}\, dr \geq c_R\, \rho \qquad \forall \rho > r_1\,. \tag{5.7}$$

In fact, (5.7) implies inequality (5.4): For arbitrary $\rho > r_1$, using polar coordinates in (i), $\|Y\| = 1$, and Cauchy–Schwarz in (ii), we obtain

$$\begin{aligned} \int_{B_\rho \setminus B_R} |u|^2 &\overset{(i)}{=} \int_R^\rho \int_{|x|=1} |u(ry)|^2\, dS(y)\, r^{n-1}\, dr \\ &\overset{(ii)}{\geq} \int_R^\rho |v(r)|^2\, r^{n-1}\, dr \overset{(5.7)}{\geq} c_R\, \rho\,. \end{aligned}$$

So it is sufficient to show (5.7).

Step 2: An equation for v*.* We multiply equation (5.1) by the spherical harmonic Y and integrate over a sphere with radius $r > R$. We use the Laplace-Beltrami operator Δ_S and the fact that the Laplace operator has the form $\Delta = r^{-(n-1)}\partial_r(r^{n-1}\partial_r) + r^{-2}\Delta_S$ in spherical coordinates. If we consider Y as a r-independent function, we get:

$$\begin{aligned} 0 &= \int_{\partial B_r(0)} (\Delta + \lambda) u\ Y\, dS \\ &= \int_{\partial B_r(0)} r^{-(n-1)}\partial_r(r^{n-1}\partial_r u)\ Y + \int_{\partial B_r(0)} r^{-2}\Delta_S u\ Y + \int_{\partial B_r(0)} \lambda u\ Y\,. \end{aligned} \tag{5.8}$$

One must be careful how to read the term with the Laplace–Beltrami operator; we let Δ_S act only on functions of the unit sphere. Using the self-adjoint nature of Δ_S and the fact that Y is an eigenfunction of $-\Delta_S$ corresponding to the real eigenvalue $\mu > 0$, we get

$$\begin{aligned}\int_{\partial B_r(0)} r^{-2}\Delta_S u\; Y &= \int_{\partial B_1(0)} r^{-2}\Delta_S[u(r\,\cdot)]\; Y\; r^{n-1} = \int_{\partial B_1(0)} r^{-2}[u(r\,\cdot)]\;\Delta_S Y\; r^{n-1}\\ &= -\mu\int_{\partial B_1(0)} r^{-2}[u(r\,\cdot)]\; Y\; r^{n-1} = -\mu\int_{\partial B_r(0)} r^{-2}u\; Y = -\mu r^{-2}v(r)\, r^{n-1}\,.\end{aligned}$$

Likewise, one must also be careful with the first term in (5.8): So we calculate

$$\begin{aligned}\int_{\partial B_r(0)} r^{-(n-1)}\partial_r(r^{n-1}\partial_r u)\; Y &= \int_{\partial B_1(0)} \partial_r(r^{n-1}\partial_r u)(r\,\xi)\; Y(\xi)\, dS(\xi)\\ &= \partial_r\left(r^{n-1}\partial_r\right)\int_{\partial B_1(0)} u(r\;\cdot)\; Y = \partial_r\left(r^{n-1}\partial_r\right) v(r)\end{aligned}$$

Calculation (5.8), after division by r^{n-1}, yields

$$\begin{aligned}0 &= r^{-(n-1)}\partial_r(r^{n-1}\partial_r v)(r) + \left(\lambda - \frac{\mu}{r^2}\right) v(r)\\ &= \partial_r^2 v(r) + \frac{n-1}{r}\partial_r v(r) + \left(\lambda - \frac{\mu}{r^2}\right) v(r)\,.\end{aligned}$$

This is an ordinary differential equation for v, which now has to be analyzed.

In a way, there are three constants in this equation, namely λ, μ, and $n-1$. First, we notice that the factor λ can easily be scaled to 1: we switch from the old independent variable r to the new variable $s = r\sqrt{\lambda}$. This results in a factor λ in all terms except the term λv. For $V(s) := v(r)$, we get the equation with the factor 1 instead of λ:

$$\partial_s^2 V(s) + \frac{n-1}{s}\partial_s V(s) + \left(1 - \frac{\mu}{s^2}\right) V(s) = 0\,. \tag{5.9}$$

For the prefactor $n-1 = 1$, this is the Bessel differential equation with the parameter μ.

Step 3: Transformation of the equation for V into an equation for W. As we will see in a moment, the parameter $n-1$ can also be transformed in the equation. This also happens in the work of Rellich, who transforms the parameter $n-1$ to the parameter 1 to obtain a standard Bessel equation. He can then use the properties of the Bessel functions. Since we are aiming for a proof that is not based on properties of Bessel functions, we will deviate from the original proof here. We use a different transformation; we want the prefactor 0 instead of $n-1$.

We use the new unknown $W(s) := V(s)s^{(n-1)/2}$. The derivatives are

$$\begin{aligned}\partial_s W(s) &= \partial_s V(s)s^{(n-1)/2} + \frac{n-1}{2}V(s)s^{(n-3)/2}\,,\\ \partial_s^2 W(s) &= \partial_s^2 V(s)s^{(n-1)/2} + (n-1)\partial_s V(s)s^{(n-3)/2} + \frac{(n-1)(n-3)}{4}V(s)s^{(n-5)/2}\,.\end{aligned}$$

Using the fact that, except for the factor $s^{(n-1)/2}$, the first two terms in the last line match the first two terms in (5.9), we get a simple equation for W:

$$\begin{aligned}
&\partial_s^2 W(s) + W(s)\\
&\quad = -\left(1-\frac{\mu}{s^2}\right) V(s)s^{(n-1)/2} + \frac{(n-1)(n-3)}{4} V(s)s^{(n-5)/2} + V(s)s^{(n-1)/2}\\
&\quad = \frac{\mu}{s^2} V(s)s^{(n-1)/2} + \frac{(n-1)(n-3)}{4} V(s)s^{(n-5)/2} = \frac{\nu}{s^2} W(s)
\end{aligned}$$

with the coefficient $\nu = \mu + (n-1)(n-3)/4$. This is a linear differential equation for W; it can be considered as a perturbation of a harmonic oscillator. An important property is shown in Lemma 5.4, namely the asymptotic behavior of the solutions: $W(s) = \alpha\cos(s+\beta) + O(1/s)$ for $s \to \infty$ for suitable β and a $\alpha \neq 0$.[1]

The asymptotic behavior of W implies $V(s) = \alpha s^{-(n-1)/2}\cos(s+\beta) + O(s^{-(n+1)/2})$ and thus $V(s)^2 s^{n-1} = |\alpha|^2\cos^2(s+\beta) + O(s^{-1})$. This yields (5.7) and thus completes the proof. □

Lemma 5.4 (Asymptotic behavior of a perturbed harmonic oscillator)**.** *We consider intervals $(s_0,\infty) \subset \mathbb{R}$ for $s_0 > 0$. Let $W : [s_0,\infty) \to \mathbb{R}$ be a solution of*

$$(\partial_s^2 + 1)W(s) = g(s)W(s) \qquad \forall s \in (s_0,\infty) \tag{5.10}$$

for a function g with the estimate $|g(s)| \leq C_0 s^{-2}$. Then W is bounded and has the asymptotic behavior for suitable $\alpha, \beta, C_1 \in \mathbb{R}$

$$|W(s) - \alpha\cos(s+\beta)| \leq C_1 s^{-1}\,. \tag{5.11}$$

If W does not vanish identically, then $\alpha \neq 0$.

Proof. Variation of constants. The proof is based on the variation-of-constants formula. To write (5.10) as a system of first-order system, we set

$$X(s) := \begin{pmatrix} W(s) \\ W'(s) \end{pmatrix}, \quad A := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad F(s) := \begin{pmatrix} 0 \\ g(s)W(s) \end{pmatrix}.$$

This allows us to write (5.10) as

$$\partial_s X(s) + A\cdot X(s) = F(s)\,.$$

The fundamental matrix of the homogeneous equation is $S(t) = \exp(-At)$, the entries of this matrix are $\pm\sin(t)$ and $\cos(t)$. Variation of constants yields the formula

$$X(t) = S(t-t_0)X(t_0) + \int_{t_0}^t S(t-s)F(s)\,ds\,. \tag{5.12}$$

Boundedness. For arbitrary times $t_0 < t_1$, we set $C_X := \sup_{t_0\leq t\leq t_1}|X(t)|$ and $\eta = \int_{t_0}^\infty s^{-2}\,ds = 1/t_0$. Because of $|F(s)| \leq C_0 s^{-2} C_X$ and $|S(s)| \leq 1$, we get the estimate $|X(t) - S(t-t_0)X(t_0)| \leq C_0\eta C_X \leq C_X/2$ from (5.12), where the last

[1] At the appropriate point in his proof, Rellich writes: "As is well known...". This was certainly correct in his time, but nowadays the properties of Bessel functions are no longer widely known.

inequality holds for every $t_0 > 2C_0$. The estimate is independent of t_1. Since $S(t-t_0)X(t_0)$ is bounded independently of t_1, we obtain the following for $t_0 > 2C_0$:

$$C_X = \sup_{t_0 \le t \le t_1} |X(t)| \le |X(t_0)| + C_X/2\,.$$

We fix t_0 and obtain $C_X \le 2|X(t_0)|$ independently of t_1. The fact that C_X is independent of t_1 implies that $X(t)$ is bounded on $[t_0, \infty)$.

Decay and limit formula. With the same arguments, we can now improve the estimates. We choose a sequence $t_k \to \infty$ and use (5.12) with $t_0 = t_k$. The vector $X(t_k)$ determines $\alpha_k \in [0, \infty)$ and $\beta_k \in [0, 2\pi)$ in such a way that

$$S(s-t_k)X(t_k) = \begin{pmatrix} \alpha_k \cos(s+\beta_k) \\ -\alpha_k \sin(s+\beta_k) \end{pmatrix} \qquad \forall s > s_0\,.$$

Since α_k is bounded, we find a subsequence and limits such that $\alpha_k \to \alpha \ge 0$ and $\beta_k \to \beta \in [0, 2\pi]$. This implies that the first part of the solution $X(t)$ in (5.12) converges in the desired way: The first component satisfies $e_1 \cdot S(t-t_k)X(t_k) = \alpha_k \cos(t+\beta_k) \to \alpha\cos(t+\beta)$ for all t.

Using the boundedness of the solution, $|X(s)| \le C_X$ (with the global bound C_X), the other term satisfies, for each fixed t,

$$\varepsilon_k(t) := \int_{t_k}^{t} S(t-s)F(s)\,ds\,, \qquad |\varepsilon_k(t)| \le \int_t^{t_k} C_0 C_X s^{-2}\,ds \le C_0 C_X t^{-1}\,.$$

With (5.12) we write the solution as

$$\begin{aligned} W(t) = e_1 \cdot X(t) &= e_1 \cdot S(t-t_k)X(t_k) + e_1 \cdot \int_{t_k}^{t} S(t-s)F(s)\,ds \\ &= \alpha_k \cos(t+\beta_k) + e_1 \cdot \varepsilon_k(t)\,. \end{aligned}$$

Passing to the limit $k \to \infty$, we obtain (5.11). In addition, we get an integral formula for W: With $F(s) = g(s)W(s)\,e_2$, we have, for each t,

$$W(t) = \alpha\cos(t+\beta) + e_1 \cdot \int_{\infty}^{t} S(t-s)F(s)\,ds\,. \tag{5.13}$$

Non-trivial prefactor. It remains to show the last statement regarding $\alpha \neq 0$. We show: If (5.11) holds with $\alpha = 0$, then W is identically zero. We therefore assume $\alpha = 0$. We choose $t_0 \ge 2C_0$ and set $C_X := \sup_{t_0 \le t < \infty} |W(t)|$. Then (5.13) yields, for each $t \ge t_0$:

$$|W(t)| \le \int_t^{\infty} |S(t-s)|\,|F(s)|\,ds \le \int_t^{\infty} C_0 s^{-2} C_X\,ds \le t_0^{-1} C_0 C_X \le C_X/2\,.$$

Taking the supremum over $t \ge t_0$, we get $C_X \le C_X/2$. This implies $C_X = 0$ and thus $W \equiv 0$ on $[t_0, \infty)$. As a solution of a linear ordinary differential equation with vanishing initial data at t_0, W must vanish identically. $\square$

5.2. Sommerfeld condition and uniqueness

Exterior domains are open sets $\Omega \subset \mathbb{R}^n$ whose complement is a compact subset of $\mathbb{R}^n$. We always demand that the boundary is a Lipschitz boundary. A standard reference for this material is [2].

Definition 5.5 (Sommerfeld radiation condition)**.** *Let $\Omega \subset \mathbb{R}^n$ be an exterior domain and $u : \Omega \to \mathbb{C}$ a solution of the Helmholtz equation with parameter $\lambda = \omega^2$.*

The classical Sommerfeld radiation condition is

$$\lim_{r\to\infty} r^{(n-1)/2} \left(\partial_r u - i\omega u\right) = 0 \tag{5.14}$$

uniformly with respect to direction. This can be interpreted mathematically as follows: For every $\varepsilon > 0$, there exists a radius $R > 0$ such that

$$\sup_{|x|=r} r^{(n-1)/2} \left|\partial_r u(x) - i\omega u(x)\right| < \varepsilon \qquad \forall r \geq R\,. \tag{5.15}$$

For us, the following weaker integral condition will suffice:

$$\int_{\partial B_r} \left|\partial_r u(x) - i\omega u(x)\right|^2 \, dS(x) \to 0 \qquad \textit{for } r \to \infty\,. \tag{5.16}$$

Remark 5.6 (Interpretation of the Sommerfeld conditions)**.** *If the solution behaves like $e^{i\omega r}$ in the radial direction, we expect $\partial_r u$ to behave similarly to $i\omega\, u$. Such a solution can therefore satisfy Sommerfeld's radiation condition. For the solution u from* (5.3) *in three dimensions, we have*

$$\partial_r u(x) - i\omega u(x) = i\omega \frac{e^{i\omega r}}{r} - \frac{e^{i\omega r}}{r^2} - i\omega u(x) = -\frac{e^{i\omega r}}{r^2}\,.$$

After multiplication by $r^{(n-1)/2} = r$, this converges to 0.

The situation is different if the solution behaves in the radial direction like $e^{-i\omega r}$; for such a solution, $|\partial_r u(x) - i\omega u(x)| \sim r^{-1}$, and Sommerfeld's radiation condition is not satisfied.

If we look at the corresponding time-dependent solution, in the former case it is $e^{i\omega r}\, e^{-i\omega t} = e^{i\omega(r-t)}$, i.e. a wave that travels outward. The second case describes a wave that travels inward – such a wave is not physically relevant.

Uniqueness in the exterior domain. With the help of Rellich's lemma, we can prove uniqueness for solutions that satisfy the Sommerfeld condition.

Theorem 5.7 (Uniqueness in the exterior domain)**.** *Let $\Omega \subset \mathbb{R}^n$ be an exterior domain, let $R > 0$ be a radius with $\mathbb{R}^n \setminus B_R(0) \subset \Omega$. Let $u : \Omega \to \mathbb{C}$ be a solution of the Helmholtz equation with parameter $\lambda = \omega^2$. Suppose that $u = 0$ on $\partial\Omega$ and Sommerfeld's radiation condition* (5.16) *holds. Then $u \equiv 0$ on $\mathbb{R}^n \setminus B_R(0)$.*

Remark 5.8. *The statement of the theorem also holds for the Neumann boundary condition $\partial_\nu u = 0$ on $\partial\Omega$. It also holds if $u = 0$ is required on one part of the boundary and $\partial_\nu u = 0$ on the rest of the boundary. The proof remains unchanged.*

Proof. The proof is based on two calculations. One is an evaluation of the Sommerfeld condition. For arbitrary differentiable functions u, we can calculate:

$$|\partial_r u - i\omega u|^2 = |\partial_r u|^2 + \omega^2 \; |u|^2 - 2\omega \operatorname{Im}\left(\partial_r u \, \bar{u}\right) \, . \tag{5.17}$$

If we integrate this identity over a sphere $S_r := \{x \in \mathbb{R}^n \mid \|x\| = r\}$, we obtain

$$\int_{S_r} |\partial_r u - i\omega u|^2 = \int_{S_r} |\partial_r u|^2 + \omega^2 \int_{S_r} |u|^2 - 2\omega \int_{S_r} \operatorname{Im}\left(\partial_r u \; \bar{u}\right) \, . \tag{5.18}$$

The radiation condition (5.16) states that the left-hand side converges to 0 for $r \to \infty$. Thus, the right-hand side also converges to 0, even though it contains two non-negative terms.

The second calculation yields a statement about the term with an indefinite sign, the integral of $\operatorname{Im}(\partial_r u \; \bar{u})$. In fact, this term arises as a boundary integral when testing the Helmholtz equation with the solution. We integrate over $\Omega_r := \Omega \cap B_r(0)$ for $r \geq R$. In the integration by parts, there is no boundary term over $\partial\Omega$ because u (or $\partial_\nu u$) vanishes there.

$$0 = \int_{\Omega_r} \left(-\Delta u - \omega^2 \, u\right) \; \bar{u} = \int_{\Omega_r} \left\{|\nabla u|^2 - \omega^2 \, |u|^2\right\} - \int_{S_r} \partial_r u \, \bar{u} \, .$$

The volume integral is real. By forming the imaginary part of the equation, we get

$$\int_{S_r} \operatorname{Im}\left(\partial_r u \, \bar{u}\right) = 0 \, . \tag{5.19}$$

So the last term in (5.18) vanishes. Since the left-hand side vanishes in the limit, both non-negative integrals must vanish in the limit, in particular the sphere integral of $|u|^2$. Theorem 5.1 yields $u \equiv 0$ on $\mathbb{R}^n \setminus B_R(0)$. □

For many exterior domains Ω, the theorem actually yields $u \equiv 0$ on Ω. Since we do not want to go into this question in depth here, we'll use the following ad-hoc definition, where "E" stands for "uniqueness":

> For $\lambda > 0$, an exterior domain Ω is called an *E-domain* if the following holds: Every solution of $-\Delta u = \lambda u$ with $u \equiv 0$ on every set $\mathbb{R}^n \setminus B_R(0) \subset \Omega$ vanishes identically on Ω.

The definition is designed so that, for E-domains, our theorem yields $u \equiv 0$ on Ω.

The definition implies that every set of the form $\Omega = \mathbb{R}^n \setminus \bar{B}_R(0)$ is an E-domain.

The principle of unique continuation implies that every simply connected exterior domain Ω is an E-domain.

The energy flux. An important quantity in the uniqueness proof was the energy flux. For a direction $\nu \in \mathbb{R}^n$, $|\nu| = 1$, we consider the energy flux density

$$F_\nu(x) := \operatorname{Im}\left(\partial_\nu u(x) \, \bar{u}(x)\right) \, . \tag{5.20}$$

For solutions u of the homogeneous equation, the total energy flux is the same across all surfaces. More precisely: Let $\Sigma \subset \Omega$ be a domain whose boundary consists of two closed surfaces, $\partial\Sigma = \Gamma_1 \cup \Gamma_2$ with a "inner" surface $\Gamma_1 \subset \Omega$ and a disjoint "outer" surface $\Gamma_2 \subset \Omega$. Then, with the normal vector ν_2 on Γ_2 (outer normal vector of Σ) and the normal vector ν_1 on Γ_1 (inner normal vector of Σ):

$$\int_{\Gamma_2} \operatorname{Im}\left(\partial_{\nu_2} u\, \bar{u}\right) = \int_{\Gamma_1} \operatorname{Im}\left(\partial_{\nu_1} u\, \bar{u}\right) . \tag{5.21}$$

The flux identity (5.21) is obtained by multiplying the equation $-\Delta u = \lambda u$ by $\bar{u}$, integrating the result over Σ, integrating by parts, and then taking the imaginary part.

This allows us to interpret the uniqueness proof as follows: The boundary condition $u = 0$ on $\partial\Omega$ implies that the energy flux through $\partial\Omega$ vanishes. Because of (5.21), the energy flux through each spherical shell S_r must vanish. This is used for uniqueness, see (5.19).

The scattering problem. For us, the most important solution of the Helmholtz equation in the exterior domain is the fundamental solution in (5.3). There is an even simpler type of solution in the whole space that we have not considered before, namely plane waves. For a vector $k \in \mathbb{R}^n$ (the wave vector) we consider

$$u^{\mathrm{inc}}(x) := e^{ik\cdot x} . \tag{5.22}$$

To obtain the physically relevant field, it is often sufficient to consider the real part, i.e. $\mathrm{Re} u^{\mathrm{inc}}(x) = \cos(k \cdot x)$. Because of $\Delta u^{\mathrm{inc}} = -|k|^2\, u^{\mathrm{inc}}$, u^{inc} is a solution of the Helmholtz equation with $\lambda = \omega^2$ if the magnitude of the wave vector is equal to the temporal frequency, $|k| = \omega$.

The solution u^{inc} does not satisfy Sommerfeld's radiation condition. The following situation is physically relevant. An object, described by a bounded closed set $\Sigma \subset \mathbb{R}^3$, is "illuminated" or "insonified" by the wave u^{inc}. We can then ask ourselves which parts of the irradiation are reflected or scattered to the side or whether there is a shadow "behind the object", i.e. a reduced field. This is the so-called scattering problem.

Mathematically, it is easy to describe: We assume that on the exterior domain $\Omega := \mathbb{R}^3 \setminus \Sigma$, the entire field is given by a function $u : \Omega \to \mathbb{C}$; this is a superposition of the plane wave and an outgoing wave,

$$u = u^{\mathrm{inc}} + u^{\mathrm{rad}} , \qquad u(x) = e^{ik\cdot x} + u^{\mathrm{rad}} . \tag{5.23}$$

If the physical setting requires a Dirichlet boundary condition for u, we have to solve the following system for u^{rad} (we recall $|k| = \omega$):

$$-\Delta u^{\mathrm{rad}} = \omega^2\, u^{\mathrm{rad}} \quad \text{in } \Omega , \tag{5.24}$$

$$u^{\mathrm{rad}}(x) = -e^{ik\cdot x} \quad \text{for } x \in \partial\Sigma , \tag{5.25}$$

$$u^{\mathrm{rad}}(x) \text{ satisfies Sommerfeld's radiation condition.} \tag{5.26}$$

For this system, we have obtained a uniqueness statement in Theorem 5.7. In the next section, we will also prove its existence.

5.3. Representation formulas and existence

While the previous considerations were dimension-independent, from now on we will restrict ourselves to three spatial dimensions because we use the fundamental solution $\Phi(x,y) := e^{i\omega|x-y|}/(4\pi|x-y|)$. We start from the representation formula on bounded domains G. For any $x \in G$, the following holds; see Exercise 5.4:

$$u(x) = \int_{\partial G} \{\partial_\nu u(y)\Phi(x,y) - u(y)\partial_\nu \Phi(x,y)\}\, dS(y)\,. \tag{5.27}$$

In the following, we consider the fundamental solution $\Phi(x,y)$ for a fixed $x \in \Omega$ and for large $|y| = r$. The explicit expression for $\Phi(x,y)$ can be approximated: we use $\nu = y/r$ and compute $|x-y|^2 = |y|^2 - 2x\cdot y + O(1)$ and therefore, using the Taylor formula, $|x-y| = |y| - x\cdot\nu + O(1/r)$. For the fundamental solution we get

$$\frac{e^{i\omega|x-y|}}{|x-y|} = \frac{e^{i\omega|y|}}{|y|}\left[e^{-i\omega x\cdot\nu} + O\left(\frac{1}{r}\right)\right]. \tag{5.28}$$

If ∇_y denotes the gradient in the variable y, then we get for the derivative

$$\begin{aligned}\nu\cdot\nabla_y\left(\frac{e^{i\omega|x-y|}}{|x-y|}\right) &= -i\omega\left(\frac{e^{i\omega|x-y|}}{|x-y|}\,\nu\cdot\frac{x-y}{|x-y|}\right) - \left(\frac{e^{i\omega|x-y|}}{|x-y|^2}\,\nu\cdot\frac{x-y}{|x-y|}\right)\\ &= i\omega\frac{e^{i\omega|y|}}{|y|}\left[e^{-i\omega x\cdot\nu} + O\left(\frac{1}{r}\right)\right].\end{aligned} \tag{5.29}$$

In particular, we have also verified that the fundamental solution satisfies the classical Sommerfeld condition in (5.14): $|y|\,|\partial_\nu\Phi(x,y) - i\omega\Phi(x,y)| \to 0$ for $|y| = r \to \infty$.

Existence in the exterior domain. We use the representation formulas to derive an existence theorem. Recall that we introduced E-domains on page 32 and that every simply connected exterior domain is an E-domain.

Theorem 5.9 (Existence in the exterior domain)**.** *Let $\Omega \subset \mathbb{R}^3$ be an E-domain, let $\omega > 0$, and let $R > 0$ be a radius with $\mathbb{R}^3 \setminus B_R(0) \subset \Omega$. Let $f : \Omega \to \mathbb{C}$ belong to $L^2(\Omega)$ and have support in $B_R(0)$. Then there is a solution $u : \Omega \to \mathbb{C}$ of the Helmholtz equation $-\Delta u - \omega^2 u = f$ satisfying the boundary condition $u = 0$ on $\partial\Omega$ and the radiation condition* (5.16)*. For $|x| > R$, the solution has the representation*

$$u(x) = \int_{\partial B_R(0)} \{\partial_\nu u(y)\Phi(x,y) - u(y)\partial_\nu\Phi(x,y)\}\, dS(y)\,. \tag{5.30}$$

Note: Another boundary condition (homogeneous Neumann condition or an inhomogeneous condition) can be imposed on $\partial\Omega$. The statement and proof remain unchanged.

Proof. Step 1: Approximate solutions. For a (large) radius $r \geq R$, consider solutions $u_r : \Omega_r \to \mathbb{C}$ on a bounded domain, namely on $\Omega_r := \Omega \cap B_r(0)$. We look at the problem

$$-\Delta u_r - \omega^2 u_r = f \quad \text{in } \Omega_r\,, \tag{5.31}$$

$$u_r = 0 \quad \text{on } \partial\Omega\,, \tag{5.32}$$

$$\partial_\nu u_r = i\omega u_r \quad \text{on } \partial B_r(0)\,. \tag{5.33}$$

For solvability: As a function space we use $X := \{u \in H^1(\Omega_r) \,|\, u|_{\partial\Omega} = 0\}$, on which we define a sesquilinear form by

$$a(u,v) := \int_{\Omega_r} \nabla u \cdot \nabla \bar{v} - \omega^2 \int_{\Omega_r} u\,\bar{v} + i\omega \int_{\partial B_r(0)} u\,\bar{v}\,. \tag{5.34}$$

An element $u \in X$ is a solution of (5.31)–(5.33) if

$$a(u,v) = \int_{\Omega_r} f\,\bar{v} \qquad \forall v \in X\,. \tag{5.35}$$

The problem (5.35) on the bounded open set Ω_r has the Fredholm property. This can be shown as in Section 4.2; we explore this in Exercise 5.5. In the next steps of the proof, we assume that these problems can be solved uniquely.

> Assumption E: Existence. There is a subsequence $r \to \infty$ such that the problems (5.35) are uniquely solvable for all r in the sequence.

Step 2: Estimates. In the following, spheres are denoted by $S_r := \partial B_r(0)$. We start by testing the equation with $\bar{u}_r$. Because of the Sommerfeld-like boundary condition (5.33), we get

$$\begin{aligned}\int_{\Omega_r} f\,\bar{u}_r &= \int_{\Omega_r} \left(-\Delta u_r - \omega^2 u_r\right)\,\bar{u}_r = \int_{\Omega_r} \left\{|\nabla u_r|^2 - \omega^2 |u_r|^2\right\} - \int_{S_r} \partial_r u_r\,\bar{u}_r \\ &= \int_{\Omega_r} \left\{|\nabla u_r|^2 - \omega^2 |u_r|^2\right\} - i\omega \int_{S_r} |u_r|^2\,.\end{aligned}$$

The imaginary part of the volume integral on the right-hand side vanishes, so we get the information

$$-\omega^{-1}\mathrm{Im} \int_{\Omega_r} f\,\bar{u}_r = \int_{S_r} |u_r|^2\,. \tag{5.36}$$

Since f is supported in $B_R(0)$, this identity shows that the values of u_r on S_r can be estimated by the norm of u_r in $B_R(0)$.

We fix $\rho > R$ for the remainder of the proof. For a sequence of solutions $(u_r)_r$, we consider the sequence of norms

$$N_r := \|u_r\|_{L^2(B_\rho(0))} = \left(\int_{B_\rho(0)} |u_r|^2\right)^{1/2}\,. \tag{5.37}$$

We introduce a second assumption.

> Assumption B: Boundedness. For every sequence $r \to \infty$ for which solutions u_r exist, the sequence N_r is bounded.

Step 3: Convergences in Ω_ρ. We assume assumptions E and B, so we are given a sequence of approximate solutions u_r, for which $u_r|_{\Omega_\rho}$ is bounded in $L^2(\Omega_\rho)$. We choose a subsequence $r \to \infty$ and a limit function u such that $u_r|_{\Omega_\rho} \rightharpoonup u$ weakly in $L^2(\Omega_\rho)$ as $r \to \infty$.

Since these are solutions to Poisson problems, the convergence is automatically better; the keyword is "Caccioppoli inequality". We outline the argument: We choose a smooth cut-off function $\theta : \Omega \to [0,1]$ with support in Ω_ρ and with $\theta \equiv 1$ in a neighborhood of $\bar{\Omega}_R$. We test equation (5.31), i.e. $-\Delta u_r - \omega^2 u_r = f$, with $\bar{u}_r\,\theta^2$ and get

$$\int_{\Omega_\rho} |\nabla u_r|^2\,\theta^2 + 2\int_{\Omega_\rho} \bar{u}_r \theta\,\nabla u_r \cdot \nabla\theta = \omega^2 \int_{\Omega_\rho} |u_r|^2\,\theta^2 + \int_{\Omega_\rho} f\,\bar{u}_r\,\theta^2\,. \tag{5.38}$$

The terms on the right-hand side are bounded. The second term on the left-hand side is estimated using the Cauchy–Schwarz inequality and Young's inequality:

$$2\left|\int_{\Omega_\rho} \bar{u}_r\theta\,\nabla u_r \cdot \nabla\theta\right| \le \frac{1}{2}\int_{\Omega_\rho} |\nabla u_r|^2\,\theta^2 + C\int_{\Omega_\rho} |u_r|^2\,|\nabla\theta|^2\,.$$

The first term of the right-hand side is absorbed into the left-hand side of (5.38). The second term of the right-hand side is bounded. We therefore obtain an H^1 estimate for the solution sequence on every reduced set (e.g. on Ω_R). We get the weak H^1 and the strong L^2 convergence of the sequence $(u_r|_{\Omega_R})_r$ on Ω_R.

We can apply a similar argument again in a neighborhood of $\partial B_R(0)$: For a smooth cut-off function $\vartheta : \Omega \to [0,1]$ supported in a neighborhood of $\partial B_R(0)$, the new function $U_r = u_r\,\vartheta$ satisfies the elliptic equation

$$\Delta U_r = (-\omega^2 u_r - f)\,\vartheta + 2\nabla u_r \cdot \nabla\vartheta + u_r \Delta\vartheta$$

with zero boundary values. Since the right-hand side converges weakly in L^2, U_r (and thus also u_r) converges weakly in H^2 in a neighborhood of S_R. This means that the traces of the first derivatives on S_R also converge strongly in $L^2(S_R)$.

We summarize: On Ω_ρ, the functions u_r converge weakly in L^2 to the limit function u, which is therefore a distributional solution of the Helmholtz equation. On Ω_R, the functions u_r even converge weakly in H^1 to the limit function u, which is therefore also a weak solution of the Helmholtz equation in Ω_R and satisfies the boundary condition $u = 0$ on $\partial\Omega$. On S_R, the traces of u_r and of their first derivatives converge.

Step 4: Representation formula. We use the representation formula (5.27) for u_r in the bounded domain $D_r = B_r(0) \setminus \bar{B}_R(0)$. We first calculate the outer boundary integral, i.e. the integral over S_r. This integral is simplified because of the chosen boundary condition $\partial_\nu u_r = i\omega u_r$:

$$\begin{aligned}
&\int_{S_r} \{\partial_\nu u_r(y)\Phi(x,y) - u_r(y)\partial_\nu\Phi(x,y)\}\;dS(y)\\
&\qquad = \int_{S_r} \{i\omega u_r(y)\Phi(x,y) - u_r(y)\partial_\nu\Phi(x,y)\}\;dS(y)\\
&\qquad = \int_{S_r} u_r(y)\;\{i\omega\Phi(x,y) - \partial_\nu\Phi(x,y)\}\;dS(y)\,.
\end{aligned}$$

The magnitude of the boundary integral on the left-hand side can therefore be estimated by

$$|\;l.S.\;| \le \|u_r\|_{L^2(S_r)} \left(\int_{S_r} |i\omega\Phi(x,y) - \partial_\nu\Phi(x,y)|^2\;dS(y)\right)^{1/2}\,.$$

The factor $\|u_r\|_{L^2(S_r)}$ is bounded because of (5.36) and the boundedness of the norm N_r (independent of r). The second factor tends to 0 for every fixed $x \in \Omega$ as $r \to \infty$; this follows from the explicit form of Φ and the two expressions (5.28) and (5.29). Basically, this is nothing more than the fact that $y \mapsto \Phi(x,y)$ satisfies Sommerfeld's radiation condition; we refer to Exercise 5.6.

This preliminary consideration shows that we can pass to the limit $r \to \infty$ in the representation formula for the bounded domain. Recall that in an neighborhood of ∂B_R, the functions are smooth, and derivatives of u_r also converge. For arbitrary $x \in D_r$, we write $u_r(x)$ with the fundamental solution and form the limit $r \to \infty$:

$$
\begin{aligned}
u_r(x) = &\int_{S_r} \{\partial_\nu u_r(y)\Phi(x,y) - u_r(y)\partial_\nu \Phi(x,y)\}\, dS(y)\\
&\quad + \int_{S_R} \{\partial_\nu u_r(y)\Phi(x,y) - u_r(y)\partial_\nu \Phi(x,y)\}\, dS(y)\\
\to &\int_{S_R} \{\partial_\nu u(y)\Phi(x,y) - u(y)\partial_\nu \Phi(x,y)\}\, dS(y) =: u(x)\,.
\end{aligned}
$$

This yields a limit function u on the whole of Ω. For $|x| < \rho$, u is the local limit of u_r; for $|x| > R$, u is given by the representation formula (5.30).

The representation formula for u shows that u is also a solution of the Helmholtz equation in $\mathbb{R}^3 \setminus B_R(0)$, because $x \mapsto \Phi(x,y)$ and $x \mapsto \partial_\nu \Phi(x,y)$ are solutions for each fixed y.

Furthermore, u satisfies Sommerfeld's radiation condition, because both $\Phi(x,y)$ and each derivative $\nu_0 \cdot \nabla_y \Phi(x,y)$, for bounded y and $|x| \to \infty$, satisfy the radiation condition. This is to be proved in Exercise 5.7. Thus, under assumptions E and B, we have found a solution u for the exterior domain problem. By the assumption on the exterior domain, this solution is uniquely determined.

Step 5: Verification of Assumption B. Our goal is to derive a contradiction if we assume that there is a sequence $r \to \infty$ for which u_r exists but $N_r \to \infty$.

We fix the sequence $r \to \infty$ with $N_r \to \infty$. We normalize the solution sequence and set $\tilde{u}_r := u_r/N_r$. This sequence is a sequence of solutions to $\tilde{f}_r := f/N_r \to 0$. All considerations of steps 3 and 4 remain applicable to the sequence $\tilde{u}_r$, which is bounded by construction. We obtain a limit function $\tilde{u}$ for the right-hand side $\tilde{f} = 0$. By the assumption on Ω, uniqueness holds, so $\tilde{u} = 0$.

Recall that $\tilde{u}_r \to \tilde{u}$ is a strong L^2 convergence in Ω_R. We even found something more: For every radius $s \in (R,\rho)$, strong L^2 convergence holds on Ω_s. On the spherical shell $\Omega_\rho \setminus \Omega_s = B_\rho(0) \setminus B_s(0)$, strong L^2 convergence follows to the convergence in the representation formula (pointwise convergence and boundedness).

This is the contradiction: On the one hand, the $\tilde{u}_r$ are normalized on $B_\rho(0)$; on the other hand, they converge strongly to $\tilde{u} = 0$ on $B_\rho(0)$.

Step 6: Assumption E. We assume that there is a sequence $r \to \infty$, so that along this sequence the problems in the bounded open sets are *not* solvable. The Fredholm property of the problems implies that we then find a sequence of solutions u_r of homogeneous problems ($f = 0$). We normalize the sequence so that $\|u_r\|_{L^2(B_\rho(0))} = 1$ for all r. Steps 3 and 4 above are applicable to this sequence. They provide a limit function u that solves the homogeneous problem and radiation condition on Ω. Because of the uniqueness, u must vanish. On the other hand, u_r was normalized on $B_\rho(0)$ and converges strongly to u in $L^2(B_\rho(0))$, which follows as in step 5. This yields the desired contradiction. □

With Theorems 5.7 and 5.9, we have gained a satisfactory understanding of the Helmholtz equation in the exterior domain: The Sommerfeld condition implies uniqueness of solutions in any space dimension. There is a solution to the Sommerfeld condition in three dimensions. The proof of the existence theorem works in any dimension as soon as you have a fundamental solution.

The far field. Solutions can be characterized even more precisely: At a great distance, the solution is essentially given by its so-called far field. The corresponding statement is of particular interest in the scattering problem.

Theorem 5.10 (Far field). *Let $\Omega \subset \mathbb{R}^3$ be an exterior domain with $\mathbb{R}^3 \setminus B_R(0) \subset \Omega$. Let $u : \Omega \to \mathbb{C}$ be a solution of $-\Delta u - \omega^2 u = f$, where $f \in L^2(\Omega)$ has compact support. Then u has the asymptotic behavior for $x = r\xi$ and $r \to \infty$:*

$$u(x) = \frac{e^{i\omega r}}{r} \, [u_\infty(\xi) + O(1/r)] \, , \tag{5.39}$$

where the far-field pattern $u_\infty : S^2 \to \mathbb{C}$ is given by

$$u_\infty(\xi) = \frac{1}{4\pi} \int_{\partial B_R} \left\{ \partial_\nu u(y) \, e^{-i\omega y \cdot \xi} + u(y) \, i\omega \nu(y) \cdot \xi \, e^{-i\omega y \cdot \xi} \right\} \, dS(y) \, . \tag{5.40}$$

Proof. We start from the representation formula (5.30) for the solution,

$$u(x) = \int_{\partial B_R} \{ \partial_\nu u(y) \Phi(x, y) - u(y) \partial_\nu \Phi(x, y) \} \, dS(y) \, . \tag{5.41}$$

Insertion of the asymptotics (5.28) and (5.29) yields the result. □

A large area of mathematical research deals with a central question, the **inverse problem.** It is relevant in many applications, such as computed tomography.

> Suppose that for a "obstacle" (a bounded set $\Sigma \subset \mathbb{R}^n$), one knows, for each direction $k \in \partial B_\omega(0) \subset \mathbb{R}^n$ of an incident wave, the far field $u_\infty = u_\infty^{(k)} : S^{n-1} \to \mathbb{C}$ of the corresponding radiating wave u^{rad} from (5.24)–(5.26). Can we reconstruct the obstacle Σ from this information?

It turns out that this is a very difficult task. However, we will not deal further with the corresponding methods here. You can modify the question a bit and consider the equation $-\Delta u - \omega^2 b u = 0$ on the whole of $\mathbb{R}^n$ for an incident wave, where we use a positive coefficient $b : \mathbb{R}^n \to \mathbb{R}$, which is supposed to be identical to 1 outside a compact set. In this setting, a positive result is given in Theorem 10.5 in [2]: The knowledge of $u_\infty^{(k)}$ for all k uniquely determines the coefficient b.

Exercises

Exercise 5.1 (Radially symmetric solutions in three dimensions)**.** *We consider* $x \in \mathbb{R}^3$ *and use* $r = |x|$*. Specify the ordinary differential equation satisfied by* U *if* $u(x) = U(r)$ *is a solution of the Helmholtz equation* $-\Delta u = \lambda u$*. Verify that*

$$u(x) := \frac{e^{i\omega r}}{r}$$

is a solution of the Helmholtz equation for $\lambda = \omega^2$*.*

Exercise 5.2 (Variation of constants)**.** *Show that for a matrix* $A \in \mathbb{R}^{m\times m}$ *and the fundamental matrix* $S(t) = \exp(-At)$ *for solutions* X *of the ordinary differential equation* $\partial_t X + A \cdot X = F$ *the variation-of-constants formula*

$$X(t) = S(t - t_0)X(t_0) + \int_{t_0}^{t} S(t - s)F(s)\, ds\,.$$

Exercise 5.3 (Fundamental solution)**.** *Show that*

$$\Phi(x) := \frac{e^{i\omega|x|}}{4\pi|x|}$$

is a fundamental solution of the Helmholtz equation in the sense that

$$(-\Delta - \omega^2)\Phi = \delta_0\,.$$

Exercise 5.4 (Representation in the bounded domain)**.** *We consider a bounded Lipschitz domain* $\Omega \subset \mathbb{R}^3$ *and a solution* u *of the Helmholtz equation* $-\Delta u = \omega^2 u$ *in* Ω *with the regularity* $u \in C^2(\Omega) \cap C^1(\bar{\Omega})$*. Show that with the fundamental solution*

$$\Phi(x, y) := \frac{e^{i\omega|x-y|}}{4\pi|x - y|}$$

the following representation formula holds for arbitrary $x \in \Omega$*:*

$$u(x) = \int_{\partial\Omega} \{\partial_\nu u(y)\Phi(x, y) - u(y)\partial_\nu \Phi(x, y)\}\; dS(y)\,,$$

where we understand the last normal derivative as a derivative in the y *variable.*

Exercise 5.5 (Fredholm property)**.** *Establish the Fredholm alternative for problem* (5.35) *on the bounded domain* Ω_r*.*

Exercise 5.6 (Smallness of an integral)**.** *Show that in* $\mathbb{R}^3$*, for fixed* $x \in \mathbb{R}^3$ *and for the fundamental solution* $\Phi(x, y) := e^{i\omega|x-y|}/(4\pi|x - y|)$*, show that the Sommerfeld integral converges to zero as* $r \to \infty$*:*

$$\int_{S_r} |i\omega\Phi(x, y) - \partial_\nu \Phi(x, y)|^2\; dS(y) \to 0\,.$$

Exercise 5.7 (Sommerfeld condition for an integral)**.** *Consider* $\mathbb{R}^3$ *and the fundamental solution* $\Phi(x, y) := e^{i\omega|x-y|}/(4\pi|x - y|)$*. A solution of the Helmholtz equation is given by a representation formula: For two functions* $f, g : \partial B_R(0) \to \mathbb{C}$*, suppose that, for each* $x \in \mathbb{R}^3 \setminus \bar{B}_R(0)$*,*

$$u(x) = \int_{\partial B_R} \{f(y)\Phi(x, y) - g(y)\partial_\nu \Phi(x, y)\}\; dS(y)\,.$$

Show that u *satisfies the Sommerfeld radiation condition.*

Part III.

Helmholtz equation with periodic coefficients

6. Floquet–Bloch transform

6.1. Transformation and its inverse

We are interested in the analysis of waveguides. The typical basic domain is $\Omega = \mathbb{R} \times \Sigma$, where $\Sigma \subset \mathbb{R}^{d-1}$ is the cross-section of the waveguide; it is always assumed to be a bounded Lipschitz domain, and $d \geq 2$ is the dimension. Everything discussed here also applies in one dimension: for $d = 1$, we consider $\Omega = \mathbb{R}$ instead of $\Omega = \mathbb{R} \times \Sigma$. The target space of all functions here is always $\mathbb{C}$, so we use $L^2(\Omega) := L^2(\Omega, \mathbb{C})$ for example. The target space $\mathbb{C}^m$ with $m \in \mathbb{N}$ can also be considered; the theory can then be applied to each component individually.

We only perform the Floquet–Bloch transform in the x_1 variable. We use the following notation: For the Floquet parameter (the quasimomentum) we write $\alpha \in I := [-\pi, \pi]$. The periodicity cell is $W := (0, 1) \times \Sigma$.

Definition of Floquet–Bloch transform

Definition 6.1 (Floquet–Bloch transform)**.** *The Floquet–Bloch transform is a bounded linear mapping*

$$\mathcal{F}_{\mathrm{FB}} : L^2(\Omega) \to L^2\big(W \times I\big) , \qquad u \mapsto \hat{u} \,. \tag{6.1}$$

For smooth functions u with compact support, the transform is defined by

$$\hat{u}(x, \alpha) \ := \ \frac{1}{\sqrt{2\pi}} \sum_{\ell \in \mathbb{Z}} u(x + \ell\, e_1)\, e^{-i\ell\alpha} \,, \tag{6.2}$$

for each $x \in W$ and $\alpha \in I$. The mapping $\mathcal{F}_{\mathrm{FB}}$ from (6.1) *is defined as the continuous extension of this mapping.*

Lemma 6.2 (Well-definedness)**.** *The mapping* (6.2) *has a unique continuous extension to a bounded linear operator $\mathcal{F}_{\mathrm{FB}}$ as in* (6.1)*. The transformation $\mathcal{F}_{\mathrm{FB}}$ preserves the inner product, i.e. $\langle \mathcal{F}_{\mathrm{FB}}(u), \mathcal{F}_{\mathrm{FB}}(v) \rangle_{L^2(W \times I)} = \langle u, v \rangle_{L^2(\Omega)}$ for all $u, v \in X = L^2(\Omega)$; it is therefore an isometry.*

Proof. The central observation is the following identity for arbitrary elements $u, v \in C_c^\infty(\Omega) \subset L^2(\Omega)$:

$$\begin{aligned} \langle \hat{u}, \hat{v} \rangle_{L^2(W \times I)} &= \int_I \int_W \hat{u}(x, \alpha)\, \overline{\hat{v}(x, \alpha)}\, dx\, d\alpha \\ &\overset{(6.2)}{=} \frac{1}{2\pi} \int_I \int_W \sum_{\ell, k \in \mathbb{Z}} u(x + \ell\, e_1)\, \overline{v(x + k\, e_1)}\, e^{-i(\ell - k)\alpha}\, dx\, d\alpha \end{aligned}$$

$$= \int_W \sum_{\ell\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \delta_{k\ell}\, u(x+\ell\, e_1)\, \overline{v(x+k\, e_1)}\, dx$$

$$= \int_W \sum_{\ell\in\mathbb{Z}} u(x+\ell\, e_1)\, \overline{v(x+\ell\, e_1)}\, dx = \int_{\mathbb{R}\times\Sigma} u\, \bar{v} = \langle u, v\rangle_{L^2(\Omega)}\,. \tag{6.3}$$

The transformation $\mathcal{F}_{\mathrm{FB}}$ thus preserves the inner product and is therefore also norm-preserving.

This property also implies that the extension $\mathcal{F}_{\mathrm{FB}}$ is well defined on $L^2(\Omega)$: An arbitrary function $u \in L^2(\Omega)$ can be approximated by smooth functions with compact support $u_j \to u$. The sequence $(u_j)_j$ is a Cauchy sequence in $L^2(\Omega)$, so the images $(\hat{u}_j)_j$ are a Cauchy sequence in $L^2(W \times I)$. Therefore, the images have a limit $\hat{u}$ in this space (with the usual argument, this limit is independent of the choice of sequence). This is how the mapping $\mathcal{F}_{\mathrm{FB}} : u \mapsto \hat{u}$ is defined. The extension also preserves the inner product. □

The inverse transform

Not only will we show that $\mathcal{F}_{\mathrm{FB}}$ is invertible, we will even be able to describe the inverse operator $\mathcal{F}_{\mathrm{FB}}^{-1} : \hat{u} \mapsto u$ with a simple formula.

Definition 6.3 (The inverse transformation)**.** *The inverse transformation is the continuous linear operator*

$$\mathcal{G} : L^2(W \times I) \to L^2(\Omega)\,, \tag{6.4}$$

which maps a function $\hat{u} \in L^2(W \times I)$ to the function $u = \mathcal{G}(\hat{u})$, which is given for almost all $x \in W$ and all $k \in \mathbb{Z}$ by

$$u(x + k\, e_1) := \frac{1}{\sqrt{2\pi}} \int_I \hat{u}(x, \beta)\, e^{ik\beta}\, d\beta\,. \tag{6.5}$$

Similar to Definition 6.1, it is not immediately clear whether the $\mathcal{G}$ operator is well-defined. However, the situation is a little different: For the $\mathcal{G}$ operator, we can actually specify the target function $u : \Omega \to \mathbb{C}$ for all $\hat{u} \in L^2(W \times I)$ – however, it is not clear a priori that $u \in L^2(\Omega)$.

Lemma 6.4 ($\mathcal{G} = \mathcal{F}_{\mathrm{FB}}^{-1}$)**.** *The operator $\mathcal{G}$ from Definition 6.3 is well-defined, is a norm-preserving isomorphism, and is an inverse operator of $\mathcal{F}_{\mathrm{FB}}$.*

Proof. Step 1: The central calculation. Consider a smooth function with compact support $u \in C_c^\infty(\Omega) \subset L^2(\Omega)$ and the transformed $\hat{u} = \mathcal{F}_{\mathrm{FB}}(u)$ from (6.2). We calculate, for any $x \in W$ and any $k \in \mathbb{Z}$:

$$\mathcal{G}(\hat{u})(x + k\, e_1) = \frac{1}{\sqrt{2\pi}} \int_I \hat{u}(x, \beta)\, e^{ik\beta}\, d\beta = \frac{1}{2\pi} \int_I \sum_{\ell\in\mathbb{Z}} u(x + \ell\, e_1)\, e^{-i\ell\beta}\, e^{ik\beta}\, d\beta$$

$$= \sum_{\ell\in\mathbb{Z}} \delta_{k\ell}\, u(x + \ell\, e_1) = u(x + k\, e_1)\,.$$

So the (6.5) transformation actually restores the original function. This proves $\mathcal{G} \circ \mathcal{F}_{\mathrm{FB}}(u) = u$, at least for smooth functions u.

Step 2: The image of $\mathcal{F}_{\mathrm{FB}}$. The proof is not yet complete: From what we have shown so far, the image $\mathcal{F}_{\mathrm{FB}}(L^2(\Omega))$ could still be a proper subspace of $L^2(W \times I)$. Then $\mathcal{G}$ could be a left inverse, but not a right inverse. In addition, $\mathcal{G}$ could then not be well defined on the whole of $L^2(W \times I)$ (because the image elements are not in $L^2(\Omega)$). This illustrates that it would be helpful to know that $\mathcal{F}_{\mathrm{FB}}$ is surjective,

$$\mathcal{F}_{\mathrm{FB}}(L^2(\Omega)) = L^2(W \times I)\,. \tag{6.6}$$

To show (6.6), we first show that the image $\mathcal{F}_{\mathrm{FB}}(L^2(\Omega))$ is at least close to $L^2(W \times I)$. According to Lebesgue's theory, functions of the form $\hat{u}(x,\alpha) = \psi(x)\theta(\alpha)$ span a dense subspace for $\psi \in C_c^\infty(W)$ and $\theta : I \to \mathbb{C}$. We now also use the theory of Fourier series, which states that every smooth function $\theta : I \to \mathbb{C}$ can be approximated with linear combinations of functions $e^{-i\ell\alpha}$ in $L^2(I)$. Therefore, the functions $\hat{u}(x,\alpha) = \psi(x)e^{-i\ell\alpha}$ for $\psi : W \to \mathbb{C}$ smooth and $\ell \in \mathbb{Z}$ span a dense subspace.

It is therefore sufficient to find a preimage for a function $\hat{u}(x,\alpha) = \frac{1}{\sqrt{2\pi}}\psi(x)e^{-i\ell\alpha}$. We find it with the formula (6.5) for $\mathcal{G}$,

$$u(x + k\,e_1) = \frac{1}{2\pi}\int_I \psi(x)e^{-i\ell\beta}\,e^{ik\beta}\,d\beta = \psi(x)\,\delta_{k,\ell}\,.$$

In fact, this function u, which is supported in the segment $[\ell, \ell+1] \times \Sigma$, satisfies $\mathcal{F}_{\mathrm{FB}}(u)(x,\beta) = \hat{u}(x,\beta)$.

This shows that the image $\mathcal{F}_{\mathrm{FB}}(L^2(\Omega))$ is close to $L^2(W \times I)$. Now let $\hat{u}_j \to \hat{u}$ be a convergent sequence of image elements, i.e. $\hat{u}_j = \mathcal{F}_{\mathrm{FB}}(u_j)$. Because of the isometric property of $\mathcal{F}_{\mathrm{FB}}$, $(u_j)_j$ is also a Cauchy sequence. Thus, $u_j \to u$ for some $u \in L^2(\Omega)$ and, because of the continuity of $\mathcal{F}_{\mathrm{FB}}$, also $\mathcal{F}_{\mathrm{FB}}(u) = \hat{u}$. This is shown by (6.6).

Step 3: The properties of $\mathcal{G}$. We now find all the desired properties. The mapping $\mathcal{F}_{\mathrm{FB}}$ is injective because it is an isometry, and because of (6.6) it is also surjective. As a bijective linear continuous mapping between Banach spaces, it has an inverse $\mathcal{F}_{\mathrm{FB}}^{-1} : L^2(W \times I) \to L^2(\Omega)$.

In step 2, we saw that the $\mathcal{G}$ operator is well-defined on a dense subset of $L^2(W \times I)$. Image elements $\hat{u} = \mathcal{F}_{\mathrm{FB}}(u)$ in this dense subset are mapped to u; since $\mathcal{F}_{\mathrm{FB}}$ is norm-preserving, $\mathcal{G} : \hat{u} \mapsto u$ is also norm-preserving (on the dense subset). So $\mathcal{G}$ can be extended to a continuous operator on the whole of $L^2(W \times I)$.

The relation $\mathcal{G} \circ \mathcal{F}_{\mathrm{FB}}(u) = u$ for smooth functions u is transferred to all functions $u \in L^2(\Omega)$ because of the continuity of both operators. Therefore, $\mathcal{G}$ is left inverse to $\mathcal{F}_{\mathrm{FB}}$. However, the left inverse of an invertible operator is always its inverse, so $\mathcal{F}_{\mathrm{FB}}^{-1} = \mathcal{G}$. □

We summarize the results in a theorem.

Theorem 6.5 (Inverse Floquet–Bloch transform)**.** *The mapping $\mathcal{F}_{\mathrm{FB}} : L^2(\Omega) \to L^2(W \times I)$ from* (6.2) *is a unitary isomorphism; its inverse is given by* (6.5)*.*

6.2. Quasiperiodic functions and extensions

We'll use the same notations as above, specifically $I = [-\pi, \pi]$, $\Omega = \mathbb{R} \times \Sigma$, and $W = (0,1) \times \Sigma$ for the unit cell.

Definition 6.6 (Quasiperiodic functions)**.** *Let $\alpha \in I$. We call α the quasimomentum. We call a mapping $u : \Omega \to \mathbb{C}$ α-quasiperiodic (in the direction of e_1 with the unit cell $W = (0,1) \times \Sigma$) if, for almost every $x \in \Omega$,*

$$u(x + e_1) = e^{i\alpha}\, u(x)\,. \tag{6.7}$$

Conversely, for any function $u \in L^2(W)$, we can define the α quasiperiodic continuation of u. We use the following for (almost) every $x \in W$ and every $\ell \in \mathbb{Z}$:

$$\tilde{u}(x + \ell\, e_1) := e^{i\ell\alpha}\, u(x)\,. \tag{6.8}$$

The function $\tilde{u}$ is quasiperiodic on Ω. However, this function generally has jumps on all surfaces $\Gamma_k = \{x \in \Omega \,|\, x_1 = k\}$ with $k \in \mathbb{Z}$. This is not what we want to see later, but since the function here only has the regularity $u \in L^2_{\mathrm{loc}}(\Omega)$, we have no traces and therefore cannot speak of jumps at all.

Remark 6.7 (Images of the Floquet–Bloch transform)**.** *We observe the following: If we evaluate formula (6.2) for arbitrary $x \in \Omega$, we get an α-quasiperiodic function $\hat{u}(\cdot, \alpha) : \Omega \to \mathbb{C}$ for each α:*

$$\begin{aligned} \hat{u}(x + e_1, \alpha) &= \frac{1}{\sqrt{2\pi}} \sum_{\ell \in \mathbb{Z}} u(x + e_1 + \ell\, e_1)\, e^{-i\ell\alpha} \\ &= \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} u(x + k\, e_1)\, e^{-i(k-1)\alpha} = e^{i\alpha}\, \hat{u}(x, \alpha)\,. \end{aligned}$$

In particular, whenever these trace values can be spoken of at all, we expect that $u(1, \cdot) = e^{i\alpha}\, u(0, \cdot)$ will be used as functions on Σ.

Remark 6.8 (Simplified inverse transform)**.** *Let $u \in L^2(\Omega)$ and $\hat{u} = \mathcal{F}_{\mathrm{FB}}(u)$. For fixed $\alpha \in I$, consider the function $\hat{u}(\cdot, \alpha)$ as an α-quasiperiodic function on the whole of Ω; we achieve this by continuing $\hat{u}$ with the rule $\hat{u}(x + k\, e_1, \alpha) = \hat{u}(x, \alpha) e^{ik\alpha}$ for each $k \in \mathbb{Z}$ and $x \in W$.*

For this extension, formula (6.5) for the inverse becomes, for any $y = x + k\, e_1 \in \Omega$ with $x \in W$,

$$u(y) := \frac{1}{\sqrt{2\pi}} \int_I \hat{u}(y, \alpha)\, d\alpha\,. \tag{6.9}$$

Derivatives of u. Functions in L^2 do not have traces, so α-quasiperiodicity needs to be defined as above. However, if we look at $u \in H^1(W)$, we can evaluate the boundary values in terms of traces; with $x = (x_1, \tilde{x})$ we write $u(1, \cdot) : \Sigma \to \mathbb{C}$ for the trace on the right boundary and $u(0, \cdot) : \Sigma \to \mathbb{C}$ for the trace on the left boundary. This allows us to define α-quasiperiodic functions as such $v \in H^1(W)$ with

$$v(1, \cdot) = e^{i\alpha}\, v(0, \cdot)\,. \tag{6.10}$$

We define the space $H^1_\alpha(W)$ as the subspace of $H^1(W)$, which consists of α-quasiperiodic functions.

Let us now investigate whether we can also consider the Floquet–Bloch transform on other function spaces. A direct consequence of the definition of (6.2) is that the transformation respects derivatives: For each direction $j \le n$ and each $u \in H^1(\Omega)$, we have

$$\mathcal{F}_{\mathrm{FB}}(\partial_j u) = \partial_j(\mathcal{F}_{\mathrm{FB}} u)\,.$$

This implies that we can also think of $\mathcal{F}_{\mathrm{FB}}$ as

$$\mathcal{F}_{\mathrm{FB}} : H^1(\Omega) \to L^2\big(I, H^1(W)\big) = \big\{h \in L^2(W \times I) \big| \nabla_x h \in L^2(W \times I)\big\}\,. \tag{6.11}$$

As we have already seen, the definition of $\hat{u}$ in (6.2) is actually such that for each $\alpha \in I$, the formula of the Floquet–Bloch transform yields an α-quasiperiodic function. For continuous functions u with compact support, we can write this in the form of boundary conditions as in (6.10). Because of the continuity of $\mathcal{F}_{\mathrm{FB}}$ in the sense of (6.11), this formula extends to $u \in H^1(\Omega)$; the image $\hat{u}(\cdot,\alpha)$ is an α-quasiperiodic function for almost every $\alpha \in I$.

Therefore, $\mathcal{F}_{\mathrm{FB}}$ is a mapping of $H^1(\Omega)$ into the space

$$L^2\big(I, H^1_\alpha(W)\big) := \left\{ u \in L^2\big(I, H^1(W)\big) \;\middle|\; \begin{array}{l} u(\cdot,\alpha) \text{ is } \alpha\text{-quasiperiodic} \\ \text{for almost all } \alpha \end{array} \right\}.$$

About the notation: The target space $H^1_\alpha(W)$ depends on the parameter $\alpha \in I$, so $L^2\big(I, H^1_\alpha(W)\big)$ is not a classical Bochner space. However, the space is a closed subspace of the Bochner space $L^2\big(I, H^1(W)\big)$; we equip it with the standard norm and inner product of this ambient space.

The $\mathcal{F}_{\mathrm{FB}} : H^1(\Omega) \to L^2\big(I, H^1_\alpha(W)\big)$ operator can be inverted, just like its L^2 counterpart. The inverse is the operator $\mathcal{F}_{\mathrm{FB}}^{-1} : L^2\big(I, H^1_\alpha(W)\big) \to H^1(\Omega)$, which is again given by the formula (6.5). We note: For $\hat{u} \in L^2\big(I, H^1_\alpha(W)\big)$, for almost every $\alpha \in I$, the function $\hat{u}(\cdot,\alpha)$ is α-quasiperiodic on W. We identify this function with an α-quasiperiodic function on Ω that belongs locally to H^1 (with no jumps on the surfaces $\{x_1 \in \mathbb{Z}\}$). The simplified inverse transform from (6.9) shows that u then also belongs locally to H^1.

Exercises

Exercise 6.1 (Floquet–Bloch formula for L^2 functions)**.** *First show, using an approximation argument, that formula* (6.2) *for the Floquet–Bloch transform is valid for all $L^2(\Omega)$ functions with bounded support.*

Show that formula (6.2) *is valid for all $u \in L^2(\Omega)$, if on the right-hand side the sum is interpreted as the convergence of the partial sums in $L^2(W)$.*

Exercise 6.2 (Floquet–Bloch transform for concrete functions)**.** *We always look at the Floquet–Bloch transform $\mathcal{F}_{\mathrm{FB}}$ of the script, especially the transformation only in the variable x_1. Calculate $\mathcal{F}_{\mathrm{FB}}(u)$ for:*

(a) *$u : \mathbb{R} \times (0,h) \to \mathbb{C}$, $u(x) = \psi(x)\mathbf{1}_{(0,1)}(x_1)$, where $\mathbf{1}_A$ is the characteristic function of the set A and $\psi \in L^2(W)$ is a function on $W = (0,1) \times (0,h)$.*

(b) *$u : \mathbb{R} \times (0,h) \to \mathbb{C}$, $u(x + k\, e_1) := \psi(x)e^{ik\omega}\mathbf{1}_{(-L,L)}(k)$ for $x \in W$ and $k \in \mathbb{Z}$, where $\omega \in \mathbb{R}$ is a frequency, $L \in \mathbb{N}$ and $\psi \in L^2(W)$ for $W = (0,1) \times (0,h)$. What is the behavior for large L?*

(c) $u : \mathbb{R} \to \mathbb{R}$, $u(x) := \sin(\omega x) \boldsymbol{1}_{(-L,L)}(x)$ *for a* $\omega \in \mathbb{R}$ *and an* $L \in \mathbb{N}$.

Exercise 6.3 (Floquet–Bloch transform of a product)**.** *In the usual setting, let* $u \in L^2(\Omega)$ *and let* $b : \Omega \to \mathbb{R}$ *be a bounded function that is* 1*-periodic in* x_1*. Show that*

$$\mathcal{F}_{\mathrm{FB}}(bu) = b\ \mathcal{F}_{\mathrm{FB}}(u)\,,\ \textit{also} \quad \mathcal{F}_{\mathrm{FB}}(bu)(x,\alpha) = b(x)\,\hat{u}(x,\alpha)$$

for almost all $(x, \alpha) \in W \times I$.

7. Periodic waveguides

Our goal is to investigate elliptic operators in a waveguide $\Omega = \mathbb{R}\times\Sigma$, where $\Sigma \subset \mathbb{R}^{d-1}$ is a bounded Lipschitz domain. Let the coefficients $a, b : \Omega \to \mathbb{R}$, both of which are periodic in the direction of $x_1 \in \mathbb{R}$, i.e. $a(x + e_1) = a(x)$ and $b(x + e_1) = b(x)$ for almost all $x \in \Omega$. Furthermore, assume that both coefficients are bounded and positive: There exists $\gamma > 0$ such that $\gamma \leq a(x), b(x) \leq 1/\gamma$ for almost every $x \in \Omega$.

We are interested in the Helmholtz problem

$$-\nabla \cdot (a\, \nabla u) - \lambda\, b\, u = f\,. \tag{7.1}$$

From now on we will always impose the boundary condition $u = 0$ on $\mathbb{R} \times \partial\Sigma$; but it is no problem to prescribe other conditions here. For $f : \Omega \to \mathbb{C}$, solutions are functions $u : \Omega \to \mathbb{C}$. The parameter $\lambda \in \mathbb{R}$ is the spectral parameter.

We always understand (7.1) as a local equation in weak form. That is, we seek a function $u : \Omega \to \mathbb{C}$ that belongs to H^1 on every bounded set, has zero boundary values on $\mathbb{R} \times \partial\Sigma$ in the sense of traces, and satisfies, for each function $\varphi \in H^1_0(\Omega)$ with bounded support,

$$\int_\Omega \{a\, \nabla u \cdot \nabla\varphi - \lambda\, b\, u\, \varphi\} = \int_\Omega f\, \varphi\,. \tag{7.2}$$

7.1. Quasiperiodic problems

Instead of the problem (7.1) on the unbounded domain, consider problems on the bounded domain W. The right-hand side $\tilde{f}_\alpha$ will have the form $\tilde{f}_\alpha = \hat{f}(\cdot, \alpha)$ in our application.

Definition 7.1 (The quasiperiodic problem). *For fixed $\alpha \in I$ and a given right-hand side $\tilde{f}_\alpha \in L^2(W)$, we seek an α-quasiperiodic solution $\tilde{u}_\alpha$ of the elliptic equation in* (7.1) *on the cell W, where we impose Dirichlet conditions on* $(0, 1) \times \partial\Sigma$. *More precisely, the weak formulation is to find*

$$\tilde{u}_\alpha \in X := \left\{\tilde{u}_\alpha \in H^1_\alpha(W) \,|\, \tilde{u}_\alpha = 0 \text{ on } (0, 1) \times \partial\Sigma\right\}\,,$$

such that, for all $\varphi \in X$,

$$\int_W \{a(x)\, \nabla\tilde{u}_\alpha(x) \cdot \nabla\bar{\varphi}(x) - \lambda\, b(x)\, \tilde{u}_\alpha(x)\, \bar{\varphi}(x)\}\; dx = \int_W \tilde{f}_\alpha(x)\, \bar{\varphi}(x)\, dx\,. \tag{7.3}$$

Remark 7.2 (Extension of solutions). *Let $\tilde{u}_\alpha$ be a solution of the α-quasiperiodic problem with right-hand side $\tilde{f}_\alpha$. We identify the two functions with their α-quasiperiodic extensions. Then, in the distributional sense, the extensions solve equation* (7.1) *on the whole of Ω, especially across the interfaces with integer x_1.*

Proof. We consider a smooth test function $\psi : (1/2, 3/2) \times \Sigma \to \mathbb{C}$; in particular, we allow the test function not to vanish at an interface. We want to prove that

$$\int_\Omega \{a \nabla \tilde{u}_\alpha \cdot \nabla \psi - \lambda \, b \, \tilde{u}_\alpha \, \psi\} = \int_\Omega \tilde{f}_\alpha \, \psi \,. \tag{7.4}$$

We are introducing a new function. We set $\tilde{\psi} : W \to \mathbb{C}$, $\tilde{\psi}(x) := \psi(x)$ for $x_1 > 1/2$ and $\tilde{\psi}(x) := \psi(x+e_1)e^{-i\alpha}$ for $x_1 < 1/2$. Then $\tilde{\psi}$ on W is an α-quasiperiodic function. We now look at the first term in the equation:

$$\begin{aligned}
\int_\Omega a \nabla \tilde{u}_\alpha \cdot \nabla \psi &= \int_W a \nabla \tilde{u}_\alpha \cdot \nabla \psi + \int_{e_1+W} a \nabla \tilde{u}_\alpha \cdot \nabla \psi \\
&= \int_W a \nabla \tilde{u}_\alpha \cdot \nabla \psi + \int_W e^{i\alpha} a \nabla \tilde{u}_\alpha \cdot \nabla \psi(\cdot + e_1) \\
&= \int_{W \cap \{x_1 > 1/2\}} a \nabla \tilde{u}_\alpha \cdot \nabla \tilde{\psi} + \int_{W \cap \{x_1 < 1/2\}} a \nabla \tilde{u}_\alpha \cdot \nabla \tilde{\psi} \\
&= \int_W a \nabla \tilde{u}_\alpha \cdot \nabla \tilde{\psi} \,.
\end{aligned}$$

The other integrals are translated in the same way. Equation (7.4) then follows from equation (7.3). □

Lemma 7.3 (Equivalent equation with Floquet–Bloch I)**.** *Let $\hat{u} \in L^2(I, H^1_\alpha(W))$ be such that $\tilde{u}_\alpha := \hat{u}(\cdot, \alpha)$ is a weak solution to the quasiperiodic problem for almost all $\alpha \in I$. Then the function $u := \mathcal{F}^{-1}_{\mathrm{FB}}(\hat{u}) = (2\pi)^{-1/2} \int_I \hat{u}(\cdot, \alpha) \, d\alpha$ is an element of $H^1_0(\Omega)$ and a weak solution of* (7.1)*.*

Proof. We have already seen in the last section that the inverse-transformed function $u = \mathcal{F}^{-1}_{\mathrm{FB}}(\hat{u})$ belongs to $H^1_0(\Omega)$. We now consider any test function $\varphi \in H^1_0(\Omega)$, for which we want to prove (7.2). To do this, we use the $\hat{\varphi}(\cdot, \alpha)$ test function in (7.3) and integrate it over $\alpha \in I$. We obtain

$$\begin{aligned}
&\int_I \int_W \left\{ a(x) \nabla \hat{u}(x, \alpha) \cdot \nabla \overline{\hat{\varphi}(x, \alpha)} - \lambda \, b(x) \, \hat{u}(x, \alpha) \, \overline{\hat{\varphi}(x, \alpha)} \right\} \, dx \, d\alpha \\
&\qquad = \int_I \int_W \hat{f}(x, \alpha) \, \overline{\hat{\varphi}(x, \alpha)} \, dx \, d\alpha \,.
\end{aligned} \tag{7.5}$$

We have $\mathcal{F}^{-1}_{\mathrm{FB}}(\hat{u}) = u$, $\mathcal{F}^{-1}_{\mathrm{FB}}(a \, \hat{u}) = a \, u$ for the product, and also $\mathcal{F}^{-1}_{\mathrm{FB}}(a \nabla \hat{u}) = a \nabla u$ for the product with a gradient; see Exercise 6.3. We now read the terms in (7.5) as inner products of functions in $L^2(W \times I)$. The unitarity of the Floquet–Bloch transform allows it to be written as $L^2(\Omega)$-inner products of the inverse-transformed functions. Thus, (7.5) actually yields (7.2), the weak form of the initial equation (7.1) for u. □

Lemma 7.4 (Equivalent equation with Floquet–Bloch II)**.** *Let $u \in H^1_0(\Omega)$ be a weak solution of* (7.1)*. Then the Floquet–Bloch transform $\hat{u} := \mathcal{F}_{\mathrm{FB}}(u)$ is an element of $L^2(I, H^1_\alpha(W))$; in particular, $\hat{u}(\cdot, \alpha) \in H^1_\alpha(W)$ for almost all $\alpha \in I$. The functions $\hat{u}(\cdot, \alpha)$ are weak solutions of* (7.3) *for almost all $\alpha \in I$.*

Proof. Let $u \in H^1_0(\Omega)$ be a weak solution of (7.1) and $\hat{u} := \mathcal{F}_{\mathrm{FB}}(u)$. We have already shown that $\hat{u} \in L^2(I, H^1_\alpha(W))$. We choose any function $\psi \in H^1(W)$ that is periodic in x_1. Then the $\psi(x)e^{i\alpha x_1}$ function is an element of $H^1_\alpha(W)$. Furthermore, we choose a number $m \in \mathbb{Z}$.

Now we construct a suitable test function for u. We define $\varphi \in H^1_0(\Omega)$ as the inverse Floquet–Bloch transform of the function $\hat{\varphi}(x,\alpha) := \psi(x)e^{i\alpha(m+x_1)}$. The weak form of (7.1) yields

$$\int_\Omega \left\{ a(x)\,\nabla u(x) \cdot \nabla\overline{\varphi(x)} - \lambda\, b(x)\, u(x)\, \overline{\varphi(x)} \right\} dx = \int_\Omega f(x)\, \overline{\varphi(x)}\, dx\,. \tag{7.6}$$

Due to the unitarity of the Floquet–Bloch transform, we can replace integrals over Ω with integrals over $W \times I$ if we replace functions with their Floquet–Bloch transforms. The periodic functions a and b as factors remain unchanged during the transformation. We get exactly the relation (7.5). According to our construction, we set $\hat{\varphi}(x,\alpha) = \psi(x)e^{i\alpha(m+x_1)}$ and can write the relation in the following form:

$$\begin{aligned}\int_I \left[\int_W \left\{ a(x) \nabla\hat{u}(x,\alpha) \cdot \nabla\overline{(\psi(x)e^{i\alpha x_1})} - \lambda\, b(x)\, \hat{u}(x,\alpha)\, \overline{\psi(x)e^{i\alpha x_1}} \right\} dx \right] e^{-im\alpha}\, d\alpha \\ = \int_I \left[\int_W \hat{f}(x,\alpha)\, \overline{\psi(x)e^{i\alpha x_1}}\, dx \right] e^{-im\alpha}\, d\alpha\,.\end{aligned}$$

In this formula, m is arbitrary. Thus, all Fourier coefficients of the two terms in square brackets agree. It follows that the expressions in the square brackets match for almost all $\alpha \in I$. This is just (7.3), because $\psi(x)e^{i\alpha x_1}$ is an arbitrary α-quasiperiodic test function. □

Loosely speaking, we have obtained:

> Solving the problem on Ω with $u \in H^1(\Omega)$ is equivalent to solving a quasiperiodic problem on W for (almost) all $\alpha \in I$.

The spectrum of the quasiperiodic problem

Lemma 7.5 (Properties of the quasiperiodic problem)**.** *In the following, not only the domain W, the coefficients a and b, but also α are fixed. The quasiperiodic problem has a discrete spectrum of eigenvalues in the following sense: There is a nondecreasing sequence $(\lambda_n)_{n\in\mathbb{N}}$ of non-negative real numbers with $\lambda_n \to \infty$ as $n \to \infty$. With $\Lambda := \{\lambda_n \,|\, n \in \mathbb{N}\}$, the following statements hold:*

1. *For each $\lambda \notin \Lambda$, The quasiperiodic problem has a unique solution $\tilde{u}_\alpha$ for each right-hand side $\tilde{f}_\alpha \in L^2(W)$, and the estimate $\|\tilde{u}_\alpha\|_{L^2(W)} \le C\|\tilde{f}_\alpha\|_{L^2(W)}$ with $C = C(\lambda,\alpha)$ independent of the right-hand side.*

2. *For every $\lambda \in \Lambda$, the quasiperiodic homogeneous problem (right-hand side $\tilde{f}_\alpha = 0$) has a nontrivial solution. The solution space of the homogeneous problem is finite-dimensional. There is a complete orthonormal system of eigenfunctions.*

Proof. This follows from the spectral theorem for compact self-adjoint operators. The reformulation in terms of operators was discussed in Section 4.1. The compact operator is defined with Lax-Milgram on the Hilbert space $X \subset H^1_\alpha(W)$. For this, look at the description (7.3). □

Lemma 7.6 (Resolvent estimate)**.** *An improved statement can be made about the constant $C = C(\lambda, \alpha)$ from Lemma 7.5. With the eigenvalues $\lambda_n = \lambda_n^\alpha$, we define the distance to the spectrum for $\lambda \in \mathbb{C}$ as*

$$\rho := \min\{|\lambda - \lambda_n| \mid n \in \mathbb{N}\}\,.$$

Then there is a $C_0 = C_0(a, b)$, independent of α and λ, so that

$$C(\lambda, \alpha) \leq \frac{C_0}{\rho}\,. \tag{7.7}$$

Proof. The estimate follows from the spectral theorem: Let $(\psi_n)_n$ be the (orthonormal) eigenfunctions corresponding to the eigenvalues $(\lambda_n)_{n\in\mathbb{N}}$ of the operator $-\nabla \cdot (a\nabla)$ with the mass operator b. In doing so, we repeat the eigenvalue λ_n according according to its multiplicity. We can expand the solution and the right-hand side (with a bounded prefactor) in this basis:

$$u(x) = \sum_n c_n \psi_n(x) \quad \text{and} \quad \frac{1}{b(x)} f(x) = \sum_n d_n \psi_n(x)\,.$$

The equation is satisfied if and only if $(\lambda_n - \lambda)c_n = d_n$ for all $n \in \mathbb{N}$. Since eigenfunctions are orthonormal,

$$\|u\|^2_{L^2(W)} = \sum_n |c_n|^2 = \sum_n \left|\frac{d_n}{\lambda_n - \lambda}\right|^2 \leq \frac{1}{\rho^2} \sum_n |d_n|^2 = \frac{C_0^2}{\rho^2} \|f\|^2_{L^2(W)}\,.$$

This yields (7.7), the resolvent estimate for solutions. □

The next two lemmata use the characterization of eigenvalues with Courant-Fischer's minimax principle. We consider two Hilbert spaces $X \subset X_0$ (we think of $X = H^1$ and $X_0 = L^2$) and the operator given by a positively definite sesquilinear form $\Phi : X \times X \to \mathbb{C}$. Then the k-th eigenvalue (with respect to the X_0 mass operator) is given by

$$\lambda_k = \sup_{\substack{F \subset X \\ \dim F = k}} \inf_{0 \neq u \perp F} \frac{\Phi(u, u)}{\|u\|^2_{X_0}}\,. \tag{7.8}$$

We do not use the proof here, because the characterization (7.8) basically expresses exactly what we did in the proof of the spectral theorem 2.8.

Lemma 7.7 (Lipschitz-continuous dependence of eigenvalues)**.** *Let $k \in \mathbb{N}$. For the k-th eigenvalue of the quasiperiodic problems, the mapping*

$$I \ni \alpha \mapsto \lambda_k^\alpha \in \mathbb{R} \tag{7.9}$$

is Lipschitz continuous.

Proof. We formulate the α-quasiperiodic problems with the α-independent space $X := H^1_{\text{per}}(W) \cap \{u | u = 0 \text{ on } (0, 1) \times \partial\Sigma\}$ and the family of sesquilinear forms $\Phi_\alpha : X \times X \to \mathbb{C}$,

$$\Phi_\alpha(u, \varphi) := \int_W a(x)(\nabla + i\alpha e_1)u \cdot \overline{(\nabla + i\alpha e_1)\varphi}\,.$$

The sesquilinear forms depend Lipschitz-continuously on α: For a constant C that is independent of α and β, we have

$$|\Phi_\alpha(u,\varphi) - \Phi_\beta(u,\varphi)| \le C\,|\alpha-\beta|\,\|u\|_X\,\|\varphi\|_X\,.$$

In this proof, we exploit the Dirichlet boundary conditions; these imply that the forms Φ_α are uniformly coercive on X. For other boundary conditions, one must make the forms coercive by adding the term $\int_W bu\bar{\varphi}$.

Application of the min–max principle. The min–max principle (7.8) characterizes eigenvalues by

$$\lambda_k^\alpha = \sup_{\substack{F\subset X\\ \dim F = k}} \inf_{0\neq u\perp F} \frac{\Phi_\alpha(u,u)}{\|u\|^2_{L^2(W)}}\,.$$

In doing so, we again equip $L^2(W)$ with the inner product with the factor b.

We now look at two points $\alpha,\beta\in I$. We choose the space F in such a way that λ_k^α is realized in it. From now on, the finite-dimensional subspace F is fixed. We now use, in turn, (i) the formula for λ_k^β with F as the comparison space. (ii) The solvability of the said infimum problem with a solution $v\in X$ (existence as in the spectral theorem). We note that the solution satisfies $\|v\|_X^2 \le C_1\|v\|^2_{L^2(W)}$ with $C_1 = C_1(F)$. To show this, we choose $0\neq v_0\perp F$ arbitrarily with $\|v_0\|_{L^2(W)} = 1$ and estimate the H^1 norm for any v with $\|v\|_{L^2(W)} = 1$: $\|v\|_X^2 \le C(a,\Sigma)\Phi_\alpha(v,v) \le C(a,\Sigma)\Phi_\alpha(v_0,v_0) =: C_1(F)$. (iii) The estimate for the difference of forms. (iv) The X-estimate for the minimizer v. (v) The characterization of λ_k^α with the space F:

$$\begin{aligned}
\lambda_k^\beta &\ge \inf_{0\neq u\perp F}\frac{\Phi_\beta(u,u)}{\|u\|^2_{L^2(W)}} = \frac{\Phi_\beta(v,v)}{\|v\|^2_{L^2(W)}}\\
&\ge \frac{\Phi_\alpha(v,v)}{\|v\|^2_{L^2(W)}} - C\,|\alpha-\beta|\,\frac{\|v\|_X^2}{\|v\|^2_{L^2(W)}}\\
&\ge \inf_{0\neq u\perp F}\left(\frac{\Phi_\alpha(u,u)}{\|u\|^2_{L^2(W)}}\right) - C\,C_1\,|\alpha-\beta|\\
&= \lambda_k^\alpha - C\,C_1\,|\alpha-\beta|\,.
\end{aligned}$$

Since α and β were arbitrary and the constants are independent of their choice, this yields the claimed Lipschitz dependence. □

For any fixed $\alpha\in I$, $\lambda_n^\alpha\to\infty$ as $n\to\infty$; this follows from the spectral theorem. However, we need a quantitative version of this statement later, where the divergence is independent of α. The following lemma yields this.

Lemma 7.8 (Number of eigenvalues in a fixed interval)**.** *We consider the quasiperiodic problem from Definition 7.1. Then the following holds:*

$$\forall \lambda_* > 0\ \exists N\in\mathbb{N}\ \forall\alpha\in I: \qquad \lambda_N^\alpha \ge \lambda_*\,. \tag{7.10}$$

So for each α there are at most N eigenvalues that are smaller than λ_*.

Proof. Step 1: Compactness. We use the compactness of the $H^1_\alpha(W) \subset L^2(W)$ embedding. First, we choose finite-dimensional subspaces $F_j \subset H^1_\alpha(W)$, whose union is dense. We demand $\dim(F_j) = j$ and $F_j \subset F_{j+1}$. Compactness implies the following:

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}: \qquad u \perp F_N \Rightarrow \|u\|_{L^2(W)} \leq \varepsilon \|u\|_{H^1(W)}\ \forall u \in H^1_\alpha(W)\,. \tag{7.11}$$

In this proof, we consider the space $L^2(W)$ with the norm induced by the b coefficient.

Claim (7.11) follows from a simple contradiction argument. We assume that, for some fixed $\varepsilon > 0$, there are functions $u_N \perp F_N$ such that $1 = \|u_N\|_{L^2(W)} \geq \varepsilon \|u_N\|_{H^1(W)}$. Compactness allows us to find a subsequence and a limit u with $u_N \to u$ in $L^2(W)$. For all $\varphi \in F_j$ we have $\langle u, \varphi \rangle = \lim_N \langle u_N, \varphi \rangle = 0$ and therefore, by density, $u = 0$. This contradicts the strong convergence and normalization.

Step 2: Coercivity. We now have to exploit a coercivity property of the sesquilinear form. The form

$$\Psi_\alpha(u, \varphi) := \int_W a(x)\, \nabla u(x) \cdot \nabla \varphi(x)\, dx \tag{7.12}$$

from (7.3) satisfies with constants $c, C > 0$ (independent of α) the estimate[1]

$$\Psi_\alpha(u, u) \geq c \|u\|^2_{H^1(W)} - C \|u\|^2_{L^2(W)}\,. \tag{7.13}$$

We now choose $\varepsilon > 0$ so small that both $c/(2\varepsilon^2) \geq \lambda_*$ and $\varepsilon^2 C \leq c/2$ are satisfied. Then we can calculate for each $u \perp F_N$:

$$\Psi_\alpha(u, u) \geq c \|u\|^2_{H^1(W)} - C \|u\|^2_{L^2(W)} \geq (c - \varepsilon^2 C) \|u\|^2_{H^1(W)} \geq \frac{c}{2} \|u\|^2_{H^1(W)}\,.$$

Step 3: Minimax principle. The Courant-Fischer characterization (7.8) of eigenvalues now completes the proof. For the subspace F_N, we have

$$\inf_{u \perp F_N} \frac{\Psi_\alpha(u, u)}{\|u\|^2_{L^2(W)}} \geq \inf_{u \perp F_N} \frac{\Psi_\alpha(u, u)}{\varepsilon^2 \|u\|^2_{H^1(W)}} \geq \frac{c}{2\varepsilon^2} \geq \lambda_*\,.$$

Since the eigenvalue λ^α_N can be characterized as the supremum over all N-dimensional subspaces F, λ^α_N is greater than the above expression, i.e. greater than (or equal to) λ_*. This was the claim. □

7.2. Spectral statements for the waveguide

We will refrain from describing the problem with operators here. Instead, we use the following definition for the problem on the unbounded domain Ω.

Definition 7.9 (Spectrum of the problem)**.** *We say that $\lambda \in \mathbb{C}$ lies in the* resolvent set *if, for every $f \in L^2(\Omega)$, equation* (7.1) *has a solution $u \in L^2(\Omega)$ that is locally of class H^1; in addition, for a constant $C = C(\lambda)$, the estimate $\|u\|_{L^2(\Omega)} \leq C \|f\|_{L^2(\Omega)}$ holds.*

We say that $\lambda \in \mathbb{C}$ is in the spectrum *if λ is not in the resolvent set.*

[1] For homogeneous Dirichlet boundary conditions, this estimate is obtained with Poincaré even with $C = 0$. But we want to show here that the proof only needs the weaker estimate.

We note the following fact from the theory of partial differential equations: As a weak solution of an elliptic equation, u is always locally also of the class H^1 (Caccioppoli inequality). An estimate $\|u\|_{L^2(\Omega)} \leq C\|f\|_{L^2(\Omega)}$ therefore always implies the apparently stronger estimate $\|u\|_{H^1(\Omega)} \leq C'\|f\|_{L^2(\Omega)}$.

Remark 7.10 (Spectrum is real and positive)**.** *The spectrum is a subset of* $[0,\infty) \subset \mathbb{C}$*. In fact, for any* $\lambda \in \mathbb{C} \setminus [0,\infty)$*, the problem* (7.1) *can be solved with the Lax-Milgram theorem.*

Lemma 7.11 (Constant coefficients)**.** *For constant coefficient functions* $a(\cdot) = a_0$ *and* $b(\cdot) = b_0$*, the spectrum is the entire interval* $[0,\infty)$*.*

Proof. Consider the function $u(x) = e^{i\omega x_1}$. We apply the elliptic operator and find

$$-\nabla \cdot (a \nabla u) - \lambda\, b\, u = -a_0 \nabla \cdot (i\omega u\, e_1) - \lambda\, b_0\, u = a_0 \omega^2 u - \lambda\, b_0\, u = 0$$

for $\omega = \sqrt{\lambda\, b_0 / a_0}$. This shows that for every $\lambda > 0$ there is a nontrivial solution of the equation for $f = 0$.

The proof would thus be complete if the solution u found belonged in the function space $L^2(\Omega)$. This is not the case. One therefore also speaks of a generalized eigenfunction.

Nevertheless, the proof can easily be completed using the above function u. For each number $R \in \mathbb{N}$ we consider the cut-off function $\vartheta_R : \mathbb{R} \to [0,1]$ with $\vartheta_R(x) = 1$ for $|x| < R$ and $\vartheta_R(x) = 0$ for $|x| > R+1$, in the two intervals $(-R-1,-R)$ and $(R,R+1)$ we choose an interpolation so that $\vartheta_R \in C^2(\mathbb{R})$. We choose the same interpolation in the transition intervals for each R.

For the generalized eigenfunction u, we now look at the new functions $v_R := u\,\vartheta_R : \Omega \to \mathbb{C}$, i.e. $v_R(x) = u(x)\vartheta_R(x_1)$, from above. The squared norms $\|v_R\|^2_{L^2(\Omega)}$ grow like R as $R \to \infty$. But if we apply the operator to v_R, we get

$$f_R := -\nabla \cdot (a \nabla v_R) - \lambda\, b\, v_R = -2a_0 \nabla\vartheta_R \cdot \nabla u - a_0 (\Delta \vartheta_R)\, u\,.$$

Since the functions $\nabla\vartheta_R$ and $\Delta\vartheta_R$ are supported in only two unit intervals, the $L^2(\Omega)$ norm of the right-hand sides f_R is bounded. So there can be no estimate for solutions. □

In the proof of Lemma 7.11, we have already used a central tool: The existence of generalized eigenfunctions usually implies that λ is a spectral value.

Theorem 7.12 (Characterization of the spectrum)**.** *Consider problem* (7.1) *on the domain* Ω*, with coefficients* a *and* b *and the boundary conditions described there. Then the following characterization of the spectrum holds: A value* $\lambda \in [0,\infty)$ *is in the spectrum of the problem if and only if there is an* $\alpha \in I$ *such that*

$$\lambda \text{ is an eigenvalue of the } \alpha\text{-quasiperiodic problem on } W\,. \tag{7.14}$$

Proof. Step 1: Eigenvalues of a quasiperiodic problem are in the spectrum. This is the simpler part of the statement. If λ is an eigenvalue of an α-quasiperiodic problem on W, the corresponding eigenfunction extends to an α-quasiperiodic solution on Ω. This function is not in $L^2(\Omega)$, so it is a generalized eigenfunction.

By considering truncated versions of this function as in the proof of Lemma 7.11, we get that there can be no estimates for solutions. Therefore, λ must be a spectral value of the problem.

Step 2: All spectral values are eigenvalues of a quasiperiodic problem. We prove the contrapositive, i.e. assume that λ is not an eigenvalue of the α-quasiperiodic problem for any α. We want to show that λ is also not a spectral value of the problem. To do this, we have to look at any $f \in L^2(\Omega)$. If we can solve the equation with $u \in H^1(\Omega)$, including estimates, we have proven the statement.

For this part of the proof, we need the Floquet–Bloch transform. To the function $f \in L^2(\Omega)$, we apply the Floquet–Bloch transform and get the function $\hat{f} = \hat{f}(x, \alpha)$ from the class $\hat{f} \in L^2(W \times I)$. For each α, we now consider the α-quasiperiodic problem with the right-hand side $\hat{f}(\cdot, \alpha)$ and the solution $\hat{u}(\cdot, \alpha)$. The Floquet–Bloch inverse transform yields $u \in L^2(\Omega)$. This means that the proof is complete if we can show:

$$\begin{aligned}&\text{The solution operators are } \hat{f}(\cdot,\alpha) \mapsto \hat{u}(\cdot,\alpha)\\ &\text{bounded, independently of the parameter } \alpha \in I.\end{aligned} \tag{7.15}$$

In fact, both $\hat{u} \in L^2(W \times I)$ and the corresponding estimate $\|\hat{u}\|_{L^2(W\times I)} \leq C\|\hat{f}\|_{L^2(W\times I)} = C\|f\|_{L^2(\Omega)}$ follow from (7.15). By the isometric property of $\mathcal{F}_{\mathrm{FB}}$, we also have $\|u\|_{L^2(\Omega)} \leq C\|f\|_{L^2(\Omega)}$.

The function u is the solution of the equation (7.1) according to Lemma 7.3. This shows the solvability and the estimate; λ is not a spectral value.

Step 3: Verification of (7.15). We use the resolvent estimate (7.7) from Lemma 7.6. The eigenvalues $\lambda_n(\alpha)$ depend continuously on α according to Lemma 7.7. The interval $I = [-\pi, \pi]$ is compact, so for each n there exists a distance $\delta = \delta(n) > 0$, so

$$|\lambda - \lambda_n(\alpha)| \geq \delta \qquad \text{for all } \alpha \in I\,.$$

Due to Lemma 7.8, we only need to consider a finite number of functions $\alpha \mapsto \lambda_n(\alpha)$. We therefore receive (7.15). □

Exercises

Exercise 7.1 (An elliptic problem with periodic boundary condition)**.** *We want to solve the following problem: $-\Delta u = f$ in $W = (0,1) \times \Sigma$ with Dirichlet boundary conditions on $(0,1) \times \partial\Sigma$ and periodic boundary conditions for $x_1 = 0$ and $x_1 = 1$. Why is it not enough to impose $u(1,\cdot) = u(0,\cdot)$?*

A good weak formulation is as follows: For the space

$$X = \{u \in H^1(W)\,|\,u(1,\cdot) = u(0,\cdot)\,,\ u = 0 \text{ on } (0,1)\times\partial\Sigma\}$$

we are looking for $u \in X$ with the property

$$\int_W \nabla u \cdot \nabla\varphi = \int_W f\,\varphi \qquad \forall \varphi \in X\,.$$

What boundary conditions at $x_1 = 0$ and $x_1 = 1$ have we encoded with it?

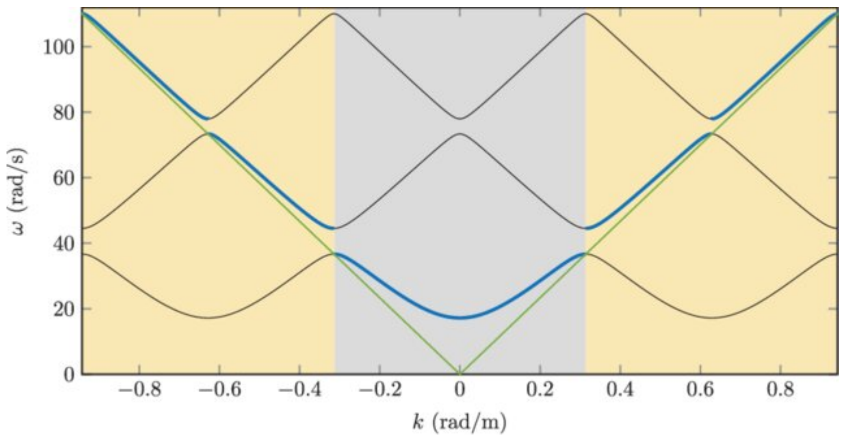

Figure 7.1.: A typical plot of dispersion curves, taken from [4]. The parameter k corresponds to our α, up to a factor of approximately 10, because the periodicity in this application is not 1. Our theory considers only the central region; in the plot, this region is continued periodically to the left and right. The vertical axis shows the frequency, i.e. the functions $\alpha \mapsto \omega_\alpha = \sqrt{\lambda_n^\alpha}$ for the first three eigenvalues, $n \in \{1, 2, 3\}$. For the periodic problem, we see that the range from 0 to about 20 lies outside the spectrum, followed by a spectral gap around 40 and another spectral gap below 80. The diagonal behavior can be interpreted as follows: For large k, the eigenfunctions behave like e^{ickx}; hence, the eigenvalues satisfy $\lambda \sim c^2 k^2$ and the frequencies satisfy $\omega \sim c|k|$.

Exercise 7.2 (An elliptic problem with quasiperiodic boundary condition)**.** *We want to solve the problem $-\Delta u = f$ again on $W = (0,1) \times \Sigma$ with Dirichlet boundary conditions on $(0,1) \times \partial\Sigma$. This time, the solution $u : W \to \mathbb{C}$ should have the property that if we continue u and f for a fixed $\alpha \in I$ to α-quasiperiodic functions on Ω, the equation should hold: $-\Delta u = f$ on all of Ω. What conditions do we want to satisfy at $x_1 = 0$ and $x_1 = 1$?*

We are looking for the solution in the form $u(x) = v(x) e^{i\alpha x_1}$ for a new function $v : W \to \mathbb{C}$. Specify a weak equation for v.

Exercise 7.3 (Spectrum of the quasiperiodic problem)**.** *Outline the proof of Lemma 7.5 in a little more detail than given there.*

Exercise 7.4 (Spectrum real and positive)**.** *Prove Remark 7.10. Use the complex Lax–Milgram theorem.*

Exercise 7.5 (Dispersion curves for constant coefficients)**.** *In the one-dimensional case, calculate all dispersion curves for constant coefficients, $a \equiv a_0 > 0$ and $b \equiv b_0 > 0$. Take advantage of the fact that all eigenfunctions are of the form $u(x) = e^{i\omega x}$. Argue why a periodic plot like the one in Figure 7.1 actually emerges.*

Exercise 7.6 (Opening of a spectral gap)**.** *We consider the one-dimensional case with $a \equiv 1$. The spectral problem is somewhat different here than in the lecture; we consider*

$$\left[-\partial_x^2 + 2\varepsilon \cos(2\pi x)\right] u = \lambda\, u\,.$$

The operator is 1-periodic, so using Floquet–Bloch theory, we can translate the problem into a family of problems on $W = (0,1)$ with parameter $\alpha \in I = [-\pi, \pi]$. If one *translates the* α*-quasiperiodic problems into the space of* 1*-periodic functions, then one must consider the following family of operators:*

$$L_\varepsilon = -(\partial_x + i\alpha)^2 + 2\varepsilon\cos(2\pi x)\,.$$

The operators L_ε *act on* 1*-periodic functions* $u \in H^1_{per}(W)$*.*

Find all eigenfunctions of L_ε *for* $\varepsilon = 0$*. For* $\alpha = \pi$*, show that the eigenvalue* $\lambda = \pi^2$ *has multiplicity two and specify two eigenfunctions.*

We are now looking for eigenfunctions of L_ε *for* $\varepsilon > 0$*. We write the (periodic) eigenfunction* u *as a Fourier series,*

$$u(x) = \sum_{n\in\mathbb{Z}} c_n e^{i2\pi nx}\,.$$

Show that u *is an eigenfunction corresponding to* $\lambda > 0$ *if and only if, for all* $n \in \mathbb{Z}$*,*

$$\Big((2\pi n + \alpha)^2 - \lambda\Big)c_n + \varepsilon(c_{n-1} + c_{n+1}) = 0\,.$$

Near the intersection point $\alpha = \pi$ *and* $\lambda = \pi^2$*, the two resonant modes* $e_0(x) = 1$ *and* $e_{-1}(x) = e^{-i2\pi x}$ *dominate. Neglecting all other Fourier coefficients, the reduced system (*$n = 0$ *and* $n = -1$ *only) is obtained as an approximation:*

$$\begin{pmatrix} \pi^2 - \lambda & \varepsilon \\ \varepsilon & \pi^2 - \lambda \end{pmatrix} \begin{pmatrix} c_0 \\ c_{-1} \end{pmatrix} = 0\,.$$

Calculate eigenvalues for $\varepsilon > 0$*. What spectral gap do we expect based on this approximation?*

8. Radiation problem in waveguides

We now want to go beyond the spectral statements, so to speak: We consider a λ from the spectrum and still want to solve equation (7.1). This seems impossible, because, roughly speaking, there are solutions with right-hand side $f = 0$. If, however, we impose an additional condition on solutions, we may be able to restore unique solvability. We will succeed in this: We have to supplement the (7.1) problem with suitable radiation conditions.

We present here an approach that was developed by Andreas Kirsch with various co-authors, see [5, 6], among others; our presentation is quite close to [7]. There is an earlier proof for most of the results by different methods, see [3]. We note that a proof that uses neither complex plane methods nor Floquet-Bloch transform is given in [13].

Notation changes compared to [7] are: k^2 becomes λ, n becomes b, the Laplace operator is generalized to $\nabla \cdot a\nabla$. The width of the unit cell was 2π, it is 1 for us. Accordingly, the boundaries of I are no longer $\pm 1/2$, but $\pm\pi$.

8.1. Families of operators

Definition 8.1 (C^1 families of operators)**.** *Let X be a Banach space and $I = [-\pi, \pi]$. We say that $(L_\alpha)_\alpha$ is a C^1* family of operators *if a $\varepsilon > 0$ exists and a C^1 mapping*

$$I_\varepsilon := (-\pi - \varepsilon, \pi + \varepsilon) \ni \alpha \mapsto L_\alpha \in \mathcal{L}(X, X)\,, \tag{8.1}$$

such that for any $\alpha \in I$, the operator L_α is a Fredholm operator with index 0.

We say that $(L_\alpha)_\alpha$ is a regular C^1 family of operators *if the following conditions are also satisfied for each $\alpha \in I$ with $\mathcal{N} := \ker(L_\alpha) \neq \{0\}$:*

(i) The operator L_α has Riesz number 1, i.e. $\mathcal{N} := \ker(L_\alpha) = \ker(L_\alpha^2)$. We then use the range $\mathcal{R} := L_\alpha(X) \subset X$ and the decomposition $X = \mathcal{N} \oplus \mathcal{R}$ with the corresponding projection P on $\mathcal{N}$.

(ii) The operator

$$M := \partial_\alpha P L_\alpha|_{\mathcal{N}} : \mathcal{N} \to \mathcal{N} \tag{8.2}$$

is invertible.

Remarks. 1. We require that every operator L_α is a Fredholm operator with index 0. This implies that for each α and $L = L_\alpha$, the subspace $\mathcal{N} := \ker(L)$ is finite-dimensional and the subspace $\mathcal{R} := L(X)$ is closed and has finite codimension; the latter agrees with the dimension of $\mathcal{N}$, since the index is equal to 0.

2. The condition $\ker(L) = \ker(L^2)$ implies that $\mathcal{N} \cap \mathcal{R} = \{0\}$. Therefore, the decomposition into a direct sum, $X = \mathcal{N} \oplus \mathcal{R}$, exists. In fact, for $u \in \mathcal{N} \cap \mathcal{R}$, we

have $u = Lx$ and $Lu = 0$, i.e. $L^2 x = 0$ and thus, according to assumption, also $Lx = 0$, i.e. $u = 0$.

3. If X is a Hilbert space with inner product $\langle \cdot, \cdot \rangle_X$ and L self-adjoint, then L always has the Riesz number 1. In fact, $L^2 x = 0$ implies that $\langle Lx, Lx \rangle_X = \langle L^2 x, x \rangle_X = 0$, so also $Lx = 0$.

4. The essential condition in the above definition is (8.2). This is a kind of transversality condition. Very lax wording (for one-dimensional $\mathcal{N}$): In α, a certain function is zero, but the derivative of the function is not equal to zero.

A central result is recorded in the following abstract theorem. We allow $J = 0$ in the following theorem, in this case $\mathcal{A}$ is the empty set. The parameter $\varepsilon > 0$ is chosen as in the definition above.

Theorem 8.2 (Families of operator equations)**.** *Let $(L_\alpha)_\alpha$ be a regular C^1 family of operators. Then the following two statements hold.*

(1) Critical values. *The set of critical values is finite: There exist a number $J \in \mathbb{N} = \{0, 1, ...\}$ and values $\{\alpha_j \,|\, j = 1, \dots, J\} \subset I$ such that*

$$\mathcal{A} \;:=\; \{\alpha \in I \,|\, \ker(L_\alpha) \neq \{0\}\} \;=\; \{\alpha_j \,|\, j = 1, \dots, J\}\,. \tag{8.3}$$

(2) Solution family. *Let $I_\varepsilon \ni \alpha \mapsto y_\alpha$ be a C^1 family of right-hand sides such that $y_{\alpha_j} \in L_{\alpha_j}(X)$ for each $j = 1, \dots, J$. Then the solution family can be continued*

$$I_\varepsilon \setminus \mathcal{A} \ni \alpha \mapsto u_\alpha := (L_\alpha)^{-1}(y_\alpha)$$

to a C^0 family on I_ε. There is a constant C, independent of the family $(y_\alpha)_\alpha$, such that

$$\sup_{\alpha \in I} \|u_\alpha\|_X \;\leq\; C \sup_{\alpha \in I} \left[\|y_\alpha\|_X + \|\partial_\alpha y_\alpha\|_X \right]. \tag{8.4}$$

Proof. Step 1: An equivalent form of the system. We start the proof by examining a point $\alpha_0 \in I$ with $\ker(L_{\alpha_0}) \neq \{0\}$. We assume $\alpha_0 = 0$ without loss of generality. The critical (non-invertible) operator is $L := L_0$. We use the $X = \mathcal{N} \times \mathcal{R}$ decomposition with the P and Q projections. We emphasize that these subspaces and projections are chosen for L and are independent of α in what follows. For α close to α_0 we write the operator L_α as

$$L_\alpha \;=\; \begin{bmatrix} PL_\alpha|_{\mathcal{N}} & PL_\alpha|_{\mathcal{R}} \\ QL_\alpha|_{\mathcal{N}} & QL_\alpha|_{\mathcal{R}} \end{bmatrix} \;:\; \mathcal{N} \times \mathcal{R} \to \mathcal{N} \times \mathcal{R}\,. \tag{8.5}$$

We then write the equation $L_\alpha u_\alpha = y_\alpha$ with $u_\alpha = (u_\alpha^N, u_\alpha^R) \in \mathcal{N} \times \mathcal{R}$ as

$$\begin{bmatrix} PL_\alpha|_{\mathcal{N}} & PL_\alpha|_{\mathcal{R}} \\ QL_\alpha|_{\mathcal{N}} & QL_\alpha|_{\mathcal{R}} \end{bmatrix} \begin{pmatrix} u_\alpha^N \\ u_\alpha^R \end{pmatrix} \;=\; \begin{pmatrix} Py_\alpha \\ Qy_\alpha \end{pmatrix}. \tag{8.6}$$

For $\alpha \neq \alpha_0 = 0$, the equation $L_\alpha u_\alpha = y_\alpha$ is therefore equivalent to the following system of equations (we divide the first line by α):

$$\tilde{L}_\alpha u_\alpha \;:=\; \begin{bmatrix} \frac{1}{\alpha} PL_\alpha|_{\mathcal{N}} & \frac{1}{\alpha} PL_\alpha|_{\mathcal{R}} \\ QL_\alpha|_{\mathcal{N}} & QL_\alpha|_{\mathcal{R}} \end{bmatrix} \begin{pmatrix} u_\alpha^N \\ u_\alpha^R \end{pmatrix} \;=\; \begin{pmatrix} \frac{1}{\alpha} Py_\alpha \\ Qy_\alpha \end{pmatrix}. \tag{8.7}$$

Relation (8.7) defines linear operators $\tilde{L}_\alpha : X \to X$ for $\alpha \neq 0$.

We would like to extend this family of operators to the point $\alpha = 0$. With $L' := (\partial_\alpha L_\alpha)|_{\alpha=0}$ and $M = PL'|_{\mathcal{N}}$ as in (8.2), we use for any $u = (u^N, u^R) \in \mathcal{N} \times \mathcal{R} = X$:

$$\tilde{L}_0 u := \begin{bmatrix} M & PL'|_{\mathcal{R}} \\ 0 & QL|_{\mathcal{R}} \end{bmatrix} \begin{pmatrix} u^N \\ u^R \end{pmatrix} . \tag{8.8}$$

We claim that the new $(-\varepsilon, \varepsilon) \ni \alpha \mapsto \tilde{L}_\alpha \in \mathcal{L}(X, X)$ family of operators is continuous. This is by definition the case for all points $\alpha \in (-\varepsilon, \varepsilon) \setminus \{0\}$. Regarding $\alpha = 0$, we note that the operators from (8.7) can be written as a difference quotient: Because $L|_{\mathcal{N}} \equiv 0$, we have

$$\frac{1}{\alpha} PL_\alpha|_{\mathcal{N}} = \frac{1}{\alpha} P(L_\alpha - L)|_{\mathcal{N}} .$$

Since we continued with $PL'|_{\mathcal{N}} = M$ for $\alpha = 0$, the resulting family in α is continuous. The same argument can be made for the second entry of the matrix: Because of $PL|_{\mathcal{R}} \equiv 0$, we can write

$$\frac{1}{\alpha} PL_\alpha|_{\mathcal{R}} = \frac{1}{\alpha} P(L_\alpha - L)|_{\mathcal{R}} .$$

The limit operator is given by the derivative in (8.8). Finally, for the third entry, we note that $QL|_{\mathcal{N}} = 0$ by the definition of $\mathcal{N}$. We conclude that the $\tilde{L}_\alpha$ family is continuous in α.

We find that the $\tilde{L}_0$ operator is invertible: The $M : \mathcal{N} \to \mathcal{N}$ operator is invertible due to the definition of a regular family. The $QL|_{\mathcal{R}} : \mathcal{R} \to \mathcal{R}$ operator is invertible due to the definition of $\mathcal{R}$ and Q. As a triangular matrix, $\tilde{L}_0$ is thus invertible.

The continuity of the $\tilde{L}_\alpha$ family together with the invertibility of $\tilde{L}_0$ yields the invertibility of $\tilde{L}_\alpha \in \mathcal{L}(X, X)$ for $\alpha \in (-\varepsilon, \varepsilon)$, possibly after reducing the size of $\varepsilon > 0$.

Step 2: Statement (1). Since $I = [-\pi, \pi] \subset \mathbb{R}$ is compact, it is sufficient to show the following claim: For each $\alpha \in I$ there is a $\varepsilon > 0$, so that $\mathcal{A} \cap (\alpha - \varepsilon, \alpha + \varepsilon)$ contains at most one point.

For $\alpha \notin \mathcal{A}$, the claim holds because small perturbations of invertible operators are in turn invertible.

We now look at $\alpha_0 \in \mathcal{A}$ and examine α in a neighborhood of α_0. For the sake of simplification of notation and without loss of generality, we assume $\alpha_0 = 0$ again. In Step 1, we showed that the equation $L_\alpha u_\alpha = y_\alpha$ has the equivalent form (8.7) and that $\tilde{L}_\alpha$ is invertible for any $\alpha \in (-\varepsilon, \varepsilon)$. It follows that L_α is invertible for any $\alpha \in (-\varepsilon, \varepsilon) \setminus \{0\}$.

Step 3: Statement (2). We look again at the situation from Step 2, where u_α solves the equation $L_\alpha u_\alpha = y_\alpha$ for $\alpha \in (-\varepsilon, \varepsilon) \setminus \{0\}$. For the right-hand side y_α, we have assumed the $y_{\alpha_j} \in L_{\alpha_j}(X)$ property for all j. In the local situation and under our assumption that the critical point is $\alpha_0 = 0$, we have $y_0 \in L(X) = \mathcal{R}$. This implies $Py_0 = 0$, and we can write the first entry of the right-hand side of (8.7) as

$$\frac{1}{\alpha} Py_\alpha = \frac{1}{\alpha} P(y_\alpha - y_0) ,$$

which can be continued for $\alpha = 0$ continuously by $P\partial_\alpha y|_0$.

The fact that the $\tilde{L}_\alpha$ family is a continuous family of invertible operators for $\alpha \in (-\varepsilon, \varepsilon)$, together with the fact that the right-hand sides of (8.7) can be continued continuously on $(-\varepsilon, \varepsilon)$, shows that the u_α family can be continued continuously. The proof yields at the same time the estimate (8.4). $\square$

Regularity of the operator family of quasiperiodic problems

We now want to examine the family of quasiperiodic problems from Definition 7.1. For fixed $\alpha \in I$, we consider the right-hand side $\hat{g}(\cdot, \alpha)$ and seek an α-quasiperiodic solution $u(\cdot, \alpha)$ of the problem

$$\int_W \left\{a(x)\,\nabla u(x,\alpha) \cdot \nabla\bar{\psi}(x) - \lambda\, b(x)\, u(x,\alpha)\,\bar{\psi}(x)\right\}\,dx = \int_W \hat{g}(\cdot,\alpha)\,\bar{\psi}(x)\,dx \quad \forall\psi\,, \tag{8.9}$$

where ψ is an arbitrary α-quasiperiodic function on W.

If we want to apply the above theory, we have to use a target space that is independent of the parameter. We have already done this in Lemma 7.7; we use the same space as there, which is

$$X := \left\{u \in H^1(W) \middle| u = 0 \text{ on } (0,1) \times \partial\Sigma \text{ and } u|_{x_1=0} = u|_{x_1=1}\right\}\,. \tag{8.10}$$

On X, we use the canonical inner product, $\langle\cdot,\cdot\rangle_X = \langle\cdot,\cdot\rangle_{H^1(W)}$.

We use the following equivalence for $U \in H^1(W)$:

$$[x \mapsto U(x)] \;\; \alpha\text{-quasiperiodic in } x_1 \iff [x \mapsto U(x)e^{-i\alpha x_1}] \;\; \text{periodic in } x_1\,. \tag{8.11}$$

It allows us to convert $H^1_\alpha(W)$ functions to $H^1_{\mathrm{per}}(W)$ functions and vice versa. The family of quasiperiodic problems (8.9) with right-hand sides $\hat{g}$ can be written by the substitutions $u(x,\alpha) = v(x,\alpha)e^{i\alpha x_1}$ and $\psi(x) = \varphi(x)e^{i\alpha x_1}$ as follows: We seek $v \in L^2(I, X)$ such that, for almost every $\alpha \in I$,

$$\begin{aligned}\int_W &\left\{a(x)\,\nabla\big(v(x,\alpha)e^{i\alpha x_1}\big) \cdot \nabla\overline{\big(\varphi(x)e^{i\alpha x_1}\big)} - \lambda\, b(x)\, v(x,\alpha)\,\overline{\varphi(x)}\right\}\,dx \\ &= \int_W \hat{g}(x,\alpha)\,\overline{\varphi(x)e^{i\alpha x_1}}\,dx\end{aligned} \tag{8.12}$$

for every $\varphi \in X$.

Now consider a fixed $\alpha \in I$. The expressions in (8.12) describe (anti-) linear functionals in φ, mappings $X \to \mathbb{C}$. By writing this functional with Riesz's representation theorem as inner products, we find elements $y_\alpha \in X$ and $L_\alpha v_0 \in X$ with

$$\langle L_\alpha v_0, \varphi\rangle_X = \int_W \left\{a(x)\,\nabla\left(v_0(x)e^{i\alpha x_1}\right) \cdot \nabla\overline{\left(\varphi(x)e^{i\alpha x_1}\right)} - \lambda\, b(x)\, v_0(x)\,\overline{\varphi(x)}\right\}\,dx\,, \tag{8.13}$$

$$\langle y_\alpha, \varphi\rangle_X = \int_W \hat{g}(x,\alpha)\,\overline{\varphi(x)e^{i\alpha x_1}}\,dx\,, \tag{8.14}$$

for each $\varphi \in X$ for $\hat{g}(\cdot,\alpha)$ and $v_0 = v(\cdot,\alpha) \in X$. With these representations, the original problem (8.23) is solved, if we find a solution $v(\cdot,\alpha) \in X \subset H^1_{\mathrm{per}}(W)$ of

$$L_\alpha v(\cdot,\alpha) = y_\alpha \tag{8.15}$$

for almost every $\alpha \in I$ and this solution family has the property $v \in L^2(I, X)$.

The operators L_α are all Fredholm operators with index 0 (we have already made this proof on bounded domains several times). The definition also immediately shows that all L_α are self-adjoint.

It is not clear whether (8.15) is solvable for every α. In fact, for spectral values λ, we know that we cannot invert the operator L_α for certain $\alpha \in I$. To solve equation (8.15), structural assumptions on g will be necessary. The background for this limitation is ultimately that we are looking for solutions to the original problem in the space $H^1(\Omega)$ with our approach, i.e. solutions with decay properties. We cannot expect all right-hand sides g to yield decaying solutions.

Now consider the single-parameter family L_α from (8.13). This family is a C^1 family due to the smooth dependence on $\alpha \mapsto L_\alpha$. Using the equivalence (8.11), the kernel $\mathcal{N}_\alpha := \ker(L_\alpha) \subset X$ is given by $\mathcal{N}_\alpha = \{e^{-i\alpha x_1} u \,|\, u \in Y^\alpha\}$ with

$$Y^\alpha \;:=\; \left\{ u \in H^1_\alpha(W) \,|\, -\nabla \cdot (a \nabla u) - \lambda b\, u = 0 \text{ in } \Omega \text{ and } u = 0 \text{ on } \mathbb{R} \times \partial\Sigma \right\}, \tag{8.16}$$

in the solution property, we have identified u with its α-quasiperiodic continuation. Since every operator L_α is a Fredholm operator, the kernel $\mathcal{N}_\alpha$ is finite-dimensional and thus Y^α is finite-dimensional. We are interested in the set of critical points

$$\mathcal{A} \;:=\; \{\alpha \in [-\pi,\pi] \,|\, \ker(L_\alpha) \neq \{0\}\}\,. \tag{8.17}$$

Without further assumptions, the set $\mathcal{A}$ can be finite or infinite. Theorem 8.2 yields that $\mathcal{A}$ is finite as soon as we can show that L_α is a regular C^1 family. This is exactly what we will get under a certain assumption.

The energy flux. We are looking at a new sesquilinear form $E : H^1(W) \times H^1(W) \to \mathbb{C}$. For $u, v \in H^1(W)$, we define

$$E(u,v) \;:=\; i \int_W a(x)\, \{u\, \partial_1 \bar{v} - \bar{v}\, \partial_1 u\}\,. \tag{8.18}$$

The arguments of E will typically be α-quasiperiodic functions, not elements of $X \subset H^1_{\mathrm{per}}(W)$. We observe that E is Hermitian, i.e. $E(u,v) = \overline{E(v,u)}$ for all u and v. In particular, $E(u,u) \in \mathbb{R}$ for all u.

The form E is related to energy fluxes through cross-sections of the form $\Gamma_t := \{t\} \times \Sigma \subset \Omega$ for $t \in \mathbb{R}$. If u and v are two solutions, $\nabla \cdot (a\nabla u) + \lambda b\, u = 0 = \nabla \cdot (a \nabla v) + \lambda b\, v$, then an application of Green's theorem in $W_{s,t} := (s,t) \times \Sigma$ for arbitrary $s < t$ yields

$$\int_{\Gamma_t} a(x)\ \{u\, \partial_1 \bar{v} - \bar{v}\, \partial_1 u\} - \int_{\Gamma_s} a(x)\ \{u\, \partial_1 \bar{v} - \bar{v}\, \partial_1 u\} = \int_{\partial W_{s,t}} a(x)\ \{u\, \partial_\nu \bar{v} - \bar{v}\, \partial_\nu u\}$$
$$= \int_{W_{s,t}} \{u\, [\nabla \cdot (a \nabla \bar{v}) + \lambda b\, \bar{v}] - \bar{v}\, [\nabla \cdot (a \nabla u) + \lambda b\, u]\} = \ 0\,,$$

where the outward normal derivative on $W_{s,t}$ is denoted by ∂_ν. The calculation shows that the flux

$$F_{u,v,t} \;:=\; i \int_{\Gamma_t} a(x)\ \{u\,\partial_1 \bar{v} - \bar{v}\,\partial_1 u\} \tag{8.19}$$

is independent of $t \in \mathbb{R}$. In particular, $E(u,v) \;=\; \int_0^1 F_{u,v,t}\,dt \;=\; F_{u,v,s}$ for each $s \in \mathbb{R}$.

We easily get that for different values of $\alpha \in (-\pi,\pi]$, the spaces Y^α are orthogonal with respect to the above sesquilinear form:

Lemma 8.3 (Orthogonality for different quasimomenta)**.** *Let $\alpha, \beta \in (-\pi,\pi]$ with $\alpha \neq \beta$ be two quasimomenta, and let $u \in Y^\alpha$ and $v \in Y^\beta$ be two solutions of the homogeneous equation. Then $E(u,v) = 0$.*

Proof. For quasiperiodic functions u and v as in the lemma, the expression from (8.19) satisfies the relationship $F_{u,v,t+1} = e^{i\alpha}e^{-i\beta}F_{u,v,t}$ by definition. On the other hand, as noted above, $F_{u,v,t}$ is independent of t. Because of $|\alpha-\beta| < 2\pi$, we conclude $F_{u,v,t} = 0$ and thus $E(u,v) = 0$. $\square$

Under an assumption on E, we obtain that L_α is a regular C^1 family.

Assumption 8.4 (Nondegeneracy)**.** *For each $\alpha \in \mathcal{A}$, let the sesquilinear form E on Y^α be nondegenerate in the following sense: For each $0 \neq \phi \in Y^\alpha$, the mapping $E(\phi,\cdot) : Y^\alpha \to \mathbb{C}$ is not identical to 0.*

Proposition 8.5 (Regularity of the operator family)**.** *Let L_α be the C^1 family of operators constructed in* (8.13)*, and suppose that Assumption 8.4 is satisfied. Then L_α is a regular C^1 family of operators in the sense of Definition 8.1.*

Proof. We fix $\alpha \in \mathcal{A}$ and consider the operator $L := L_\alpha$ with kernel $\mathcal{N} := \ker(L)$ and derivative $L' := \partial_\alpha L_\alpha$. We need to prove that $M := PL'|_{\mathcal{N}} : \mathcal{N} \to \mathcal{N}$ is invertible, where P denotes the projection onto $\mathcal{N}$. In the following calculation, the definition of L_α in (8.13) yields the first equality; we're using here that $e^{i\alpha x_1}e^{-i\alpha x_1} = 1$ is independent of α. In the second equality, we use that the terms from the first and second expressions cancel each other out if the derivatives are applied to $u(x)\,e^{i\alpha x_1}$ and $\varphi(x)\,e^{i\alpha x_1}$, respectively, but not to x_1.

$$\begin{aligned}
&\langle L'u, \varphi\rangle_X \\
&\quad = i \int_W a(x)\ \Big\{\nabla(u(x)\,x_1\,e^{i\alpha x_1}) \cdot \nabla(\overline{\varphi(x)\,e^{i\alpha x_1}}) - \nabla(u(x)\,e^{i\alpha x_1}) \cdot \nabla(\overline{\varphi(x)\,x_1\,e^{i\alpha x_1}})\Big\}\ dx \\
&\quad = i \int_W a(x)\ \Big\{u(x)\,e^{i\alpha x_1}\,\partial_1(\overline{\varphi(x)e^{i\alpha x_1}}) - \partial_1(u(x)\,e^{i\alpha x_1})\,\overline{\varphi(x)\,e^{i\alpha x_1}}\Big\}\ dx \\
&\quad = E(u\,e^{i\alpha x_1}, \varphi\,e^{i\alpha x_1})\,.
\end{aligned}$$

From this calculation, we can conclude that $PL'|_{\mathcal{N}}$ is invertible. Let $u \in \mathcal{N}$ with $PL'u = 0$. Since $L = L_\alpha$ is self-adjoint (in X), $\mathcal{N}$ and $\mathcal{R}$ are orthogonal. In this situation, $PL'u = 0$ implies that $\varphi \mapsto \langle L'u, \varphi\rangle_X$ is the zero form on $\mathcal{N}$. The above calculation together with the fact that E is nondegenerate implies that this is only possible for $u\,e^{i\alpha x_1} = 0$ and thus for $u = 0$. We get that the kernel of $PL'|_{\mathcal{N}}$ is trivial. Therefore, M is invertible from (8.2). $\square$

Corollary 8.6 (The spaces Y_j and basis functions)**.** *We consider a Helmholtz equation for which Assumption 8.4 is satisfied. In this situation, the L_α family constructed in* (8.13) *is a regular C^1 family of operators. There is a finite (possibly empty) set of values $\mathcal{A} = \{\alpha_j | j = 1, \dots, J\}$ such that*

$$Y_j \;:=\; \Big\{ u \in H^1_{\alpha_j}(W) \,|\, \nabla \cdot (a \nabla u) + \lambda b\, u = 0 \text{ in } \Omega \text{ and } u = 0 \text{ on } \mathbb{R} \times \partial\Sigma \Big\} \tag{8.20}$$

is not trivial. Each space Y_j has a finite dimension $m_j \in \mathbb{N}$, and the spaces Y_j are orthogonal with respect to E. We use the direct sum

$$Y := \bigoplus_{j=1}^{J} Y_j \subset H^1(W)\,. \tag{8.21}$$

We choose an inner product $\langle \cdot, \cdot \rangle_{Y_j}$ for each space Y_j and solve the Hermitian eigenvalue problem

$$E(\phi, \psi) \;=\; \lambda\, \langle \phi, \psi \rangle_{Y_j} \quad \text{for all } \; \psi \in Y_j \tag{8.22}$$

according to $\lambda \in \mathbb{R}$ and $\phi \in Y_j$. This yields an orthogonal basis of Y_j, consisting of eigenfunctions $\phi_{\ell,j}$, $\ell = 1, \dots, m_j$. The value $\lambda = 0$ is not an eigenvalue.

Proof. Proposition 8.5 yields that L_α is a regular family. The functional-analytic Theorem 8.2 yields that the set of critical α values is finite. Due to Lemma 8.3, the spaces Y_j are orthogonal with respect to E. Assumption 8.4 guarantees that $\lambda = 0$ is not an eigenvalue. The remaining statements repeat the definitions and follow from the Fredholm property of the operator family. The functions ϕ from (8.22) are orthogonal to each other in Y_j according to their design. □

8.2. Existence results

We look again at

$$-\nabla \cdot (a \nabla u) - \lambda b u = g \quad \text{ in } \Omega = \mathbb{R} \times \Sigma\,, \tag{8.23}$$

with the weak form

$$\int_\Omega \{ a\, \nabla u \cdot \nabla \bar{\varphi} - \lambda b\, u\, \bar{\varphi} \} = \int_\Omega g\, \bar{\varphi} \tag{8.24}$$

for all $\varphi \in H^1_0(\Omega)$. We always use Dirichlet boundary conditions: $u(\cdot) = 0$ on $\mathbb{R} \times \partial\Sigma$ for (8.23) and later $\hat{u}(\cdot, \alpha) = 0$ on $(0,1) \times \partial\Sigma$ for almost every $\alpha \in I$.

For the source terms, we introduce a weighted L^2 space. We define

$$L^2_*(\Omega) := \left\{ f : \Omega \to \mathbb{C} \,\middle|\, x \mapsto (1 + |x_1|^2) f(x) \in L^2(\Omega) \right\}\,. \tag{8.25}$$

Bounded solutions for orthogonal sources

We turn to our first existence result for the Helmholtz equation. We characterize the right-hand sides g, for which the equation (8.23) has a solution in $H^1(\Omega)$.

Theorem 8.7 (Existence of $H^1(\Omega)$ solutions with Floquet–Bloch theory)**.** *For given Σ, a, b and λ, we consider the Helmholtz equation* (8.23) *for a right-hand side $g \in L^2_*(\Omega)$. Assumption 8.4 is satisfied.*

Existence: *The Floquet–Bloch transform $\hat{g}(\cdot,\alpha)$ has the cellwise orthogonality property*

$$\langle \hat{g}(\cdot,\alpha_j), \phi\rangle_{L^2(W)} = 0 \quad \text{for all} \;\; j \in \{1,\dots,J\} \text{ and } \phi \in Y_j\,. \tag{8.26}$$

Then (8.23) *has a solution $u \in H^1(\Omega)$ with $\|u\|_{H^1} \le C\|g\|_{L^2_*(\Omega)}$ for a constant $C = C(\Sigma, a, b, \lambda)$.*

Uniqueness: *If $u \in H^1(\Omega)$ is a solution of* (8.23)*, then the orthogonality condition* (8.26) *holds. The solution u is uniquely determined.*

Proof. Existence. Using the Floquet–Bloch transform, we have shown that the equation (8.23) is equivalent to the family of equations $L_\alpha v(\cdot,\alpha) = y_\alpha$ from (8.15), $\alpha \in I = [-\pi,\pi]$. In particular, it is sufficient to find a family of solutions $v(\cdot,\alpha)$ of (8.15) and prove that $v \in L^2(I, H^1_{\mathrm{per}}(W))$. By the definition of the critical values $\mathcal{A} = \{\alpha_j | j = 1,\dots,J\}$, there is a unique solution for each $\alpha \in I \setminus \mathcal{A}$: $v(\cdot,\alpha)$. We claim that this family of solutions can be continued throughout I.

We consider one of the critical values, $\alpha = \alpha_j \in \mathcal{A}$, and a small interval $\tilde{I} = [\alpha_j - \varepsilon, \alpha_j + \varepsilon]$, which contains no other critical value. We want to apply the functional-analytic result from Theorem 8.2. We use the space X from (8.10), the operator family L_α from (8.13) and the family of right-hand sides y_α from (8.14).

We need to check the assumptions of Theorem 8.2. The operators L_α depend smoothly on α and are invertible for all $\alpha \in \tilde{I} \setminus \mathcal{A}$. We turn to the $y_{\alpha_j} \in \mathcal{R} = L_{\alpha_j}(X)$ condition. For any element $\varphi \in \mathcal{N} := \ker(L_{\alpha_j}) \subset X$, we note that $\phi(x) := \varphi(x) e^{i\alpha_j x_1} \in Y_j$; by the definition of y_α

$$\langle y_{\alpha_j}, \varphi\rangle_X \;=\; \int_W \hat{g}(x,\alpha_j)\, e^{-i\alpha_j x_1}\, \overline{\varphi(x)}\, dx \;=\; \int_W \hat{g}(\cdot,\alpha_j)\, \bar{\phi} \;=\; 0 \tag{8.27}$$

because of the orthogonality assumption (8.26). This shows that y_{α_j} is orthogonal to $\mathcal{N}$. Since L_{α_j} is self-adjoint, the $\mathcal{N}$ and $\mathcal{R}$ subspaces are orthogonal. Since L_{α_j} is also Fredholm with index 0, the space X is the orthogonal direct sum $\mathcal{N} \oplus \mathcal{R}$. Since we have shown that y_{α_j} is orthogonal to $\mathcal{N}$, we have found $y_{\alpha_j} \in \mathcal{R}$.

Proposition 8.5 yields that L_α is a regular operator family as defined in Definition 8.1; therefore, Theorem 8.2 can be applied. We get that $I \ni \alpha \mapsto v(\cdot,\alpha)$ is continuous, so $v \in L^2(I, H^1_{\mathrm{per}}(W))$ in particular. This yields an $H^1_0(\Omega)$ solution of (8.23).

We turn to the estimate for the solution. The definition of y_α (compare first equation in (8.27)) implies that $[\alpha \mapsto y_\alpha] \in C^1(I, X)$ when $[\alpha \mapsto \hat{g}(\cdot,\alpha)] \in C^1(I, L^2(W))$. We will now prove the latter.

Using the functions $g_\ell : W \to \mathbb{C}$, $g_\ell(x_1, \tilde{x}) := g(x_1 + \ell, \tilde{x})$, we can estimate the norm of g by

$$\|g\|^2_{L^2_*(\Omega)} = \sum_{\ell\in\mathbb{Z}} \int_W |g_\ell(x)|^2 [1 + |x_1 + \ell|^2]^2 dx \ge c_0 \sum_{\ell\in\mathbb{Z}} (1+\ell^2)^2 \|g_\ell\|^2_{L^2(W)}$$

for a universal constant $c_0 > 0$. This allows us to estimate the norm of a finite sum for arbitrary $0 \le m < M$, which is related to the definition of the Floquet–Bloch transform of (6.2):

$$\Bigg\| \sum_{m\le|\ell|\le M} \ell\, g(x+\ell\, e_1)\, e^{-i\ell\alpha} \Bigg\|_{L^2(W)} \le \sum_{m\le|\ell|\le M} |\ell|\, \|g_\ell\|_{L^2(W)}$$
$$\le \sum_{m\le|\ell|\le M} \frac{1}{\sqrt{1+\ell^2}} (1+\ell^2)\, \|g_\ell\|_{L^2(W)}$$
$$\le \sqrt{\sum_{m\le|\ell|\le M} \frac{1}{1+\ell^2}} \sqrt{\sum_{m\le|\ell|\le M} (1+\ell^2)^2\, \|g_\ell\|^2_{L^2(W)}} \le C_m \|g\|_{L^2_*(\Omega)}$$

with $C_m > 0$ independent of g and with arbitrarily small $C_m > 0$ if m is large. The function $\hat{g}(\cdot,\alpha)$ is defined by the Floquet–Bloch transform from (6.2). The term on the left-hand side of the above calculation is, as an infinite sum, even $\partial_\alpha \hat{g}(\cdot,\alpha)$ except for one factor. The above calculation shows that the sequence of partial sums is a Cauchy sequence, so the series is well defined. As a result of the calculation, there exists $C > 0$ such that

$$\|\hat{g}(\cdot,\alpha)\|_{L^2(W)} + \|\partial_\alpha \hat{g}(\cdot,\alpha)\|_{L^2(W)} \le C\, \|g\|_{L^2_*(\Omega)}$$

for all α. Theorem 8.2 yields the estimate (8.4) for solutions, i.e. a bound for $\hat{u} \in C^0(I, H^1(W))$, and in particular for $\hat{u} \in L^2(I, H^1(W))$. This results in the estimate for $u \in H^1(\Omega)$, namely $\|u\|_{H^1} \le C\|g\|_{L^2_*(\Omega)}$.

Uniqueness. To show the unique solvability of (8.23), it is sufficient to prove the unique solvability of (8.9) for almost every α. For each $\alpha \notin \mathcal{A}$, the equation (8.9) can be solved uniquely due to the definition of the critical α values. This already shows the uniqueness of the solution.

Let us also see that the existence of a solution implies the orthogonality relation. If $u \in H^1(\Omega)$ is a solution of (8.23), we want to show (8.26). For $\phi \in Y_j$, we calculate with the solution property of $\hat{u}(\cdot,\alpha_j)$,

$$\langle \hat{g}(\cdot,\alpha_j), \phi\rangle_{L^2(W)} = \int_W \left\{ a(x)\nabla\hat{u}(x,\alpha_j)\cdot\nabla\overline{\phi(x)} - \lambda b(x)\,\hat{u}(x,\alpha_j)\,\overline{\phi(x)} \right\} dx$$
$$= -\langle \hat{u}(\cdot,\alpha_j), (\nabla\cdot(a\nabla)+\lambda b)\phi\rangle_{L^2(W)} \;=\; 0\,,$$

where we have used the fact that integration by parts produces no boundary term for two functions in $H^1_\alpha(W)$. This proves the theorem. □

Lemma 8.8 (Orthogonality criterion)**.** *The orthogonality condition* (8.26) *is formulated using the Floquet–Bloch transforms of g. With the original function g and the space Y from* (8.21)*, an equivalent condition is*

$$\int_\Omega g(x)\,\overline{\phi(x)}\,dx \;=\; 0 \quad \textit{for all } \phi \in Y\,. \tag{8.28}$$

Proof. We fix $j \in \{1,\ldots,J\}$, set $\alpha := \alpha_j$, and select a function ϕ from Y_j. We identify ϕ with its α-quasiperiodic extension, which satisfies $\phi(x+\ell e_1) = \phi(x)e^{i\ell\alpha}$, where e_1 is the unit vector in the x_1 direction. We calculate for $\hat{g} = \mathcal{F}_{\mathrm{FB}}(g)$

$$\sqrt{2\pi}\big\langle \hat{g}(\cdot,\alpha),\phi\big\rangle_{L^2(W)} = \Big\langle \sum_{\ell\in\mathbb{Z}} g(\cdot+\ell e_1)\, e^{-i\ell\alpha}\,,\phi\Big\rangle_{L^2(W)}$$
$$= \sum_{\ell\in\mathbb{Z}} \Big\langle g(\cdot+\ell e_1),\phi(\cdot+\ell e_1)\Big\rangle_{L^2(W)} = \int_\Omega g(x)\,\overline{\phi(x)}\,dx\,.$$

We notice that the series are each well-defined because of $g \in L^2_*(\Omega)$. □

Solution for predetermined radiation conditions

In the previous subsection, we obtained a solution u of (8.23) if g satisfies the orthogonality condition (8.28). This is not the type of solution that is typically observed. In the physical problem, we have to consider the equation

$$-\nabla\cdot a\nabla u - \lambda b u = f \quad \text{ in } \Omega = \mathbb{R}\times\Sigma\,, \tag{8.29}$$

with a general right-hand side f and want to get solutions that are approximate linear combinations of outgoing waves far from the origin. Such solutions do not belong to the space $H^1(\Omega)$. We consider sources $f \in L^2_*(\Omega)$, see (8.25).

To define the radiation condition, we use two cut-off functions.

Definition 8.9 (Cut-off functions $\rho_\pm$)**.** *We say that $\rho_+,\rho_- \in C^2(\mathbb{R},[0,1])$ are admissible cut-off functions if there is a number $R>0$ such that*

$$\begin{aligned} \rho_+(x_1) &= 1\,, \quad \rho_-(x_1) = 0 \qquad \forall x_1 > R\,,\\ \rho_+(x_1) &= 0\,, \quad \rho_-(x_1) = 1 \qquad \forall x_1 < -R\,. \end{aligned}$$

Remark on cut-off functions. Formally, the radiation condition formulated below depends on the choice of $\rho_\pm$. However, we will show later that the solution u to the radiation problem does not depend on the choice of $\rho_\pm$.

The requirement $\rho_\pm \in C^2(\mathbb{R},\mathbb{R})$ can be replaced by $\rho_\pm \in C^{0,1}(\mathbb{R},\mathbb{R})$, i.e. by Lipschitz continuity of the cut-off functions.

From now on, we use the spaces Y_j and the basis functions $\phi_{\ell,j}$ as chosen in Corollary 8.6. We change the notation slightly at this point: We now combine all the basis functions $\phi_{\ell,j}$ into a new family with only one index and write $(\phi_\ell)_\ell$, where $1 \le \ell \le L := \sum_{j=1}^J m_j$. Recall that there is orthogonality with respect to the Hermitian sesquilinear form E, i.e. $E(\phi_\ell,\phi_{\ell'}) = \delta_{\ell,\ell'}\,E(\phi_\ell,\phi_\ell)$ for all ℓ,ℓ'.

Definition 8.10 (Propagating part and radiation condition)**.** *We fix admissible cut-off functions $\rho_\pm$ as in Definition 8.9. For each $\ell \le L$, the mode ϕ_ℓ is called right-going if $E(\phi_\ell,\phi_\ell) > 0$; it is called left-going when $E(\phi_\ell,\phi_\ell) < 0$. For each ℓ for which ϕ_ℓ is right-going, we set $\rho_\ell := \rho_+$ and for each ℓ for which ϕ_ℓ is left-going, we set $\rho_\ell := \rho_-$.*

(i) Propagating part. *For complex coefficients $(c_\ell)_{1\le\ell\le L}$, we say that*

$$w \;=\; \sum_{\ell=1}^{L} c_\ell\,\rho_\ell\,\phi_\ell \tag{8.30}$$

is the propagating wave function belonging to $c \in \mathbb{C}^L$*.*

(ii) Radiation condition. *We say that a solution* $u \in H^1_{loc}(\Omega)$ *of* (8.29) *satisfies the radiation condition if there exists a* $c \in \mathbb{C}^L$ *such that, with the associated propagating wave function* w*,*

$$v := u - w \in H^1(\Omega)\,. \tag{8.31}$$

Definition 8.10 allows us to show an existence result with our previously developed methods: We solve the radiation problem (8.29) by constructing $v = u - w \in H^1_0(\Omega)$ using Theorem 8.7. We can write the equation for v as

$$-\nabla \cdot a\nabla v - \lambda b v = g := f + (\nabla \cdot a\nabla w + \lambda b w)\,. \tag{8.32}$$

The expression $\nabla \cdot a\nabla w + \lambda b w$ has compact support, and we will apply Theorem 8.7 with the right-hand side g. The function g depends on the coefficient vector $c \in \mathbb{C}^L$. We will choose $c \in \mathbb{C}^L$ in such a way that g satisfies the orthogonality and (8.32) has a solution $v \in H^1_0(\Omega)$.

The formulation of the radiation condition in Definition 8.10 yields the following implications. Existence: If we find a $c \in \mathbb{C}^L$ such that (8.32) has a solution $v \in H^1_0(\Omega)$, then $u = w + v \in H^1_{\mathrm{loc}}(\bar\Omega)$ is a solution of (8.29) with the radiation condition. Uniqueness: If $u \in H^1_{\mathrm{loc}}(\bar\Omega)$ is a nontrivial solution of (8.29) with the radiation condition, then $c \in \mathbb{C}^L$ and a solution $v \in H^1_0(\Omega)$ of (8.32) exist, so either c or v are nontrivial.

Theorem 8.11 (Existence of radiating solutions)**.** *Let* Σ*,* a*,* b*,* λ *and* f *be as above; for simplicity, we assume* $a \in C^1(\bar\Omega)$ *here. Let two cut-off functions* $\tilde\rho_\pm$ *be fixed as in Definition 8.9. For an assignment* $\sigma : \{1, ..., L\} \to \{\pm 1\}$*, we set* $\rho_\ell = \tilde\rho_{\sigma(\ell)}$*. We require Assumption 8.4. Then* (8.29) *has a unique solution* $u \in H^1_{loc}(\bar\Omega)$*, which satisfies the radiation condition. For* w*,* v *and* c *from the radiation condition, the estimate*

$$\|v\|_{H^1(\Omega)} \;+\; \|w\|_{H^1(W)} \;\le\; C\,\|f\|_{L^2_*(\Omega)} \tag{8.33}$$

for a constant $C = C(\Sigma, a, b, \lambda, \rho_\pm)$*. The coefficients* c_ℓ *for* $\ell \in \{1, \dots, L\}$ *are given by*

$$c_\ell \;=\; \frac{i}{\sigma(\ell)\, E(\phi_\ell, \phi_\ell)}\,\langle f, \phi_\ell\rangle_{L^2(\Omega)}\,. \tag{8.34}$$

Proof. Existence. We want to define $c \in \mathbb{C}^L$ in the definition of w such that g satisfies the orthogonality condition (8.28) from (8.32). We use a basis function $\phi_{\ell'} \in Y_j$, which we think of as a α_j-quasiperiodic function on Ω. We want (8.28) to be satisfied by g and therefore demand the first equality. The second equation is the definition of w from (8.30).

$$\begin{aligned} -\langle f, \phi_{\ell'}\rangle_{L^2(\Omega)} &= \langle (\nabla \cdot a\nabla + \lambda b) w, \phi_{\ell'}\rangle_{L^2(\Omega)} \\ &= \sum_{\ell=1}^{L} c_\ell\, \langle (\nabla \cdot a\nabla + \lambda b)(\rho_\ell\, \phi_\ell), \phi_{\ell'}\rangle_{L^2(\Omega)}\,. \end{aligned} \tag{8.35}$$

We calculate with product rules

$$(\nabla \cdot a\nabla + \lambda b)(\rho_\ell\, \phi_\ell) = \rho_\ell(\nabla \cdot a\nabla + \lambda b)\phi_\ell + \nabla\rho_\ell \cdot a\,\nabla\phi_\ell + \nabla \cdot (\phi_\ell\, a\,\nabla\rho_\ell) \tag{8.36}$$

$$= \rho_\ell'\, a\,\partial_1\phi_\ell + \partial_1(\phi_\ell\, a\,\rho_\ell')\,. \tag{8.37}$$

The result of the calculation is to be understood in the sense of distribution, but you can also use test functions from the space $H^1_{\mathrm{loc}}(\Omega)$.

The result allows us to evaluate the inner product on the right-hand side of (8.35):

$$\begin{aligned}\big\langle(\nabla\cdot a\nabla+\lambda b)(\rho_\ell\,\phi_\ell),\phi_{\ell'}\big\rangle_{L^2(\Omega)} &= \big\langle\rho_\ell'\,a\,\partial_1\phi_\ell+\partial_1(\phi_\ell\,a\,\rho_\ell'),\phi_{\ell'}\big\rangle_{L^2(\Omega)}\\ &= \int_\Omega\bar\phi_{\ell'}\,\rho_\ell'\,a\,\partial_1\phi_\ell-\partial_1\bar\phi_{\ell'}\,a\,\rho_\ell'\,\phi_\ell \;=\; i\int_{\mathbb{R}}\rho_\ell'(t)\,F_{\phi_\ell,\phi_{\ell'},t}\,dt\end{aligned}$$

with the flux $F_{\phi_\ell,\phi_{\ell'},t}$ from (8.19). The flux is independent of t and is consistent with $E(\phi_\ell,\phi_{\ell'})=E(\phi_\ell,\phi_\ell)\,\delta_{\ell,\ell'}$. Looking at (8.35), we multiply the result by c_ℓ and sum over ℓ and get

$$\sum_{\ell=1}^{L}c_\ell\,i\int_{\mathbb{R}}\rho_\ell'(t)\,F_{\phi_\ell,\phi_{\ell'},t}\,dt \;=\; c_{\ell'}\,E(\phi_{\ell'},\phi_{\ell'})\,\sigma(\ell')\,.$$

Because of (8.35), the orthogonality condition (8.35) is equivalent to

$$-\langle f,\phi_\ell\rangle_{L^2(\Omega)} \;=\; c_\ell\,i\,E(\phi_\ell,\phi_\ell)\,\sigma(\ell)\,.$$

This is the condition we specified in (8.34).

The above calculation shows: If $(c_\ell)_\ell$ is chosen according to (8.34), the orthogonality condition (8.26) is satisfied for g. Therefore, we can determine v using Theorem 8.7. The existence statement is thus shown. With a C dependent on $\rho_\pm$, the estimate $\|v\|_{H^1(\Omega)}\le C\|g\|_{L^2_*(\Omega)}\le C(\|f\|_{L^2_*(\Omega)}+\|(\nabla\cdot a\nabla+\lambda b)w\|_{L^2(\Omega)})\le C(\|f\|_{L^2_*(\Omega)}+|c|_{\mathbb{C}^L})\le C\|f\|_{L^2_*(\Omega)}$. This implies the estimate (8.33) for solutions.

Uniqueness. Let u be a solution of the radiation problem with $f=0$. We want to show that u vanishes. The equation for $v\in H^1(\Omega)$ has the right-hand side $g=(\nabla\cdot a\nabla+\lambda b)w$. Theorem 8.7 implies that g satisfies the orthogonality condition (8.28). The above calculation shows that the orthogonality condition for g determines the coefficients $c\in\mathbb{C}^L$ by a formula; for $f=0$, the coefficients are $c=0$. Thus $w=0$. From the uniqueness statement of Theorem 8.7, it follows that $v=0$ and thus $u=0$. This proves uniqueness. □

Remark on well-definedness. We have constructed solutions $u=v+w$ for given cut-off functions $\rho_\pm$. To investigate the well-definedness of the radiation condition, we consider the consequences of choosing other admissible cut-off functions, which we denote by $\tilde\rho_\pm$.

We refer to the corresponding solutions as $u=v+w$ and $\tilde u=\tilde v+\tilde w$. Then we write

$$u-\tilde u \;=\; v-\tilde v \;+\; \sum_\ell\tilde c_\ell\,(\rho_\ell-\tilde\rho_\ell)\,\phi_\ell \;+\; \sum_\ell(c_\ell-\tilde c_\ell)\,\rho_\ell\,\phi_\ell\,.$$

We observe that $v-\tilde v+\sum_\ell\tilde c_\ell\,(\rho_\ell-\tilde\rho_\ell)\,\phi_\ell$ is in $H^1(\Omega)$. Therefore, $u-\tilde u$ satisfies not only the homogeneous Helmholtz equation, but also a radiation condition with coefficients $(c_\ell-\tilde c_\ell)_\ell$. The uniqueness result from the Theorem 8.11 implies $u=\tilde u$ and $c_\ell=\tilde c_\ell$ for all ℓ. In this sense, the choice of cut-off functions has no influence on the solution.

Remark on the regularity of a. The assumption $a \in C^1(\bar{\Omega})$ is not necessary to obtain the statement of the theorem; $a \in L^\infty(\Omega)$ is enough. To carry out the proof under this weaker assumption, one must note that the product rule (8.37) also holds in the distributional sense for $a \in L^\infty(\Omega)$.

A later part in the proof is a bit more complex: The new right-hand side is $g = f + (\nabla \cdot a\nabla + \lambda b)w$ and thus belongs only to $L^2_*(\Omega) + H^{-1}(\Omega)$, with the second part having compact support. For such right-hand sides, Theorem 8.7 can also be proved; an application of the correspondingly generalized Theorem 8.7 then yields the above existence Theorem 8.11 for $a \in L^\infty(\Omega)$.

Exercises

Exercise 8.1 (Families of operator equations)**.** *Theorem 8.2 yields for an operator family L_α and for right-hand sides y_α a solution family u_α. In addition, an estimate for solutions is shown; see* (8.4)*. Give an example showing that, in general (in the situation of Theorem 8.2), no estimate of the form*

$$\sup_{\alpha \in I} \|u_\alpha\|_X \;\leq\; C \sup_{\alpha \in I} \|y_\alpha\|_X \tag{8.38}$$

can hold.

Exercise 8.2 (One-dimensional equation with constant coefficients)**.** *We consider the Helmholtz equation on $\Omega = \mathbb{R}$ with $\lambda \in (0, \infty)$ for $a \equiv 1$ and $b \equiv 1$. Determine the set $\mathcal{A}$ of the critical values of α and the associated spaces Y^α. Check the orthogonality with respect to the sesquilinear form E, the nondegeneracy of E, and the signs of the energy fluxes.*

Exercise 8.3 (Two-dimensional kernel)**.** *We consider the two-dimensional case $d = 2$, the cross-section $\Sigma = (0, \pi)$, and the coefficient $b \equiv 1$. We choose a special $\alpha \in I = [-\pi, \pi]$ here to find an eigenspace with dimension greater than 1. Two α-quasiperiodic solutions are*

$$\phi_1(x) := e^{i\alpha x_1} \sin(2x_2) \quad \textit{and} \quad \phi_2(x) := e^{i(\alpha - 2\pi)x_1} \sin x_2 \,.$$

The two functions satisfy $-\Delta\phi_1 = (\alpha^2 + 4)\phi_1$ and $-\Delta\phi_2 = ((\alpha - 2\pi)^2 + 1)\phi_2$. Show that you can choose α in such a way that the two factors match. If you choose λ accordingly, you obtain an at least two-dimensional space Y_j.

Exercise 8.4 (Orthogonality and fluxes)**.** *We continue Exercise 8.3: For ϕ_1 and ϕ_2, prove orthogonality in $L^2(W)$ and with respect to E. Calculate the fluxes; which wave is right-going and which is left-going?*

Bibliography

[1] H. W. Alt. *Linear functional analysis.* Universitext. Springer-Verlag London, Ltd., London, german edition, 2016. An application-oriented introduction.

[2] D. Colton and R. Kress. *Inverse acoustic and electromagnetic scattering theory*, volume 93 of *Applied Mathematical Sciences.* Springer-Verlag, Berlin, second edition, 1998.

[3] S. Fliss and P. Joly. Solutions of the time-harmonic wave equation in periodic waveguides: asymptotic behaviour and radiation condition. *Arch. Ration. Mech. Anal.*, 219(1):349–386, 2016.

[4] A. Fărăgău, J. Barbosa, A. Metrikine, and K. van Dalen. Dynamic amplification in a periodic structure with a transition zone subject to a moving load: Three different phenomena. *Mathematics and Mechanics of Solids*, 05 2022.

[5] A. Kirsch and A. Lechleiter. The limiting absorption principle and a radiation condition for the scattering by a periodic layer. *SIAM J. Math. Anal.*, 50(3):2536–2565, 2018.

[6] A. Kirsch and A. Lechleiter. A radiation condition arising from the limiting absorption principle for a closed full- or half-waveguide problem. *Math. Methods Appl. Sci.*, 41(10):3955–3975, 2018.

[7] A. Kirsch and B. Schweizer. Periodic waveguides revisited: Radiation conditions, limiting absorption principles, and the space of bounded solutions. *Mathematical Methods in the Applied Sciences*, 48(2):2267–2293, 2025.

[8] A. Lamacz and B. Schweizer. Effective acoustic properties of a meta-material consisting of small Helmholtz resonators. *Discrete Contin. Dyn. Syst. Ser. S*, 10(4):815–835, 2017.

[9] F. Rellich. Über das asymptotische Verhalten der Lösungen von $\Delta u + \lambda u = 0$ in unendlichen Gebieten. *Jber. Deutsch. Math.-Verein.*, 53:57–65, 1943.

[10] B. Schweizer. *Partielle Differentialgleichungen.* Springer-Verlag, Berlin, 2013. Eine anwendungsorientierte Einführung. [An application-oriented introduction].

[11] B. Schweizer. The low-frequency spectrum of small Helmholtz resonators. *Proc. A.*, 471(2174):20140339, 18, 2015.

[12] B. Schweizer. Resonance meets homogenization: Construction of meta-materials with astonishing properties. *Jahresber. Dtsch. Math.-Ver.*, 119(1):31–51, 2017.

[13] B. Schweizer. Inhomogeneous Helmholtz equations in wave guides—existence and uniqueness results with energy methods. *European J. Appl. Math.*, 34(2):211–237, 2023.

Index