Scattering of time-harmonic acoustic waves: Helmholtz equation, boundary integral equations and BEM

Lecture notes for the "Advanced numerical methods for PDEs" class

Andrea Moiola
Department of Mathematics, University of Pavia
andrea.moiola@unipv.it
https://euler.unipv.it/moiola/T/MNAPDE2023/MNAPDE2023.html

February 27, 2023

This is a preliminary version of the notes.
Please let me know if you find any error or have any suggestion.

The main character in this course is the **Helmholtz equation** (also known as "reduced wave equation"), which is the following partial differential equation (PDE):

$$\boxed{\Delta u + k^2 u = 0.} \tag{1}$$

Here $\Delta = \sum_{j=1}^{n} \partial_{x_j}^2$ is the Laplace operator in n variables for $n \in \{1, 2, 3\}$, $k > 0$ is a real parameter called wavenumber, and u is the unknown of the equation, a scalar field defined on a subset of $\mathbb{R}^n$. We will mostly consider the two-dimensional case $n = 2$.

In the following, we first derive the Helmholtz equation from different physical phenomena, showing that it can be used to model different kinds of linear wave problems. Then we show some simple special solutions of the equation. We describe some boundary value problems (BVPs) and focus on one of them, the exterior Dirichlet problem. We show how to reformulate this as a boundary integral equation (BIE), introduce a numerical method (the boundary element method, BEM) to approximate its solution, and explain how to implement it. We also use Fredholm theory to study some analytical properties of the BIE, in particular its well-posedness, and finally sketch some alternative BIEs for the same BVP.

At the end of the course you should be able to reproduce with Matlab all the figures of these notes.

In the first few sections the approach will be more "physical" than "mathematical", so we will not make precise assumptions and will gloss over some issues such as the regularity of the objects involved or the admissibility of some operations.

All the results presented here are well-known, but scattered over several books. At the end of this document a few useful references are listed; they are the main sources used in the preparation of these notes (see also those mentioned in Remark 5.24).

1 Why is the Helmholtz equation relevant?

1.1 Acoustics

Sound waves are mechanical vibrations propagating in a fluid, which could be either a gas or a liquid. In this section we introduce the basic equations that model them. A more complete introduction can be found in [CJ77, §6] and [BK00, §3]. Being able to model, and thus to control, sound propagation is important for numerous industrial and medical applications: not only for the design of concert halls, musical instruments, microphones and loudspeakers, but also for noise and vibration mitigation (e.g. in cars and aircraft), medical ultrasound imaging, non-invasive therapy such as high-intensity focussed ultrasound surgery (HIFU), offshore oil exploration, underwater communication (sonar), bioacoustics, nondestructive testing, sensors...

We denote by $\rho(\mathbf{x}, t)$ the **density**, by $p(\mathbf{x}, t)$ the **pressure** and by $\mathbf{v}(\mathbf{x}, t)$ the **velocity** of the fluid in a point $\mathbf{x} \in \mathbb{R}^n$ at time $t \in \mathbb{R}$. We denote by ∇, div (or $\nabla\cdot$) and Δ the gradient, the divergence and the

Laplacian in the space coordinate $\mathbf{x}$ only (i.e. without derivatives in t). Conservation of mass gives the **continuity equation**

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v}) = 0$$

and conservation of momentum gives **Euler's equation**

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} + \frac{1}{\rho}\nabla p = \mathbf{0}$$

where $[(\mathbf{v}\cdot\nabla)\mathbf{v}]_j = \sum_{m=1}^n v_m \partial_{x_m} v_j$. Both PDEs are non-linear as they contain quadratic terms (products of ρ and $\mathbf{v}$ or $\mathbf{v}$ and its derivatives). We assume that all three quantities considered are small perturbations of the constant[1] static values:

$$\rho(\mathbf{x},t) = \rho_0 + \rho_\approx(\mathbf{x},t), \quad p(\mathbf{x},t) = p_0 + p_\approx(\mathbf{x},t), \quad \mathbf{v}(\mathbf{x},t) = \mathbf{v}_0 + \mathbf{v}_\approx(\mathbf{x},t) \quad \text{and that} \quad \mathbf{v}_0 = \mathbf{0}.$$

$p_\approx$ is called **acoustic pressure** or **excess pressure**. Linearising both equations around the static values (using $\frac{1}{\rho} = \frac{1}{\rho_0(1+\frac{\rho_\approx}{\rho_0})} \approx \frac{1}{\rho_0}(1-\frac{\rho_\approx}{\rho_0})$) we obtain

$$\frac{\partial \rho_\approx}{\partial t} + \rho_0 \operatorname{div}(\mathbf{v}_\approx) = 0 \quad \text{and} \quad \frac{\partial \mathbf{v}_\approx}{\partial t} + \frac{1}{\rho_0}\nabla p_\approx = \mathbf{0}. \tag{2}$$

The pressure is an increasing function of the density $p = f(\rho)$ with $p_0 = f(\rho_0)$. Linearising this relation and denoting $c^2 := \frac{\partial f}{\partial \rho}(\rho_0)$ we have $p_0 + p_\approx = f(\rho_0+\rho_\approx) \approx f(\rho_0) + c^2\rho_\approx$, thus $p_\approx = c^2\rho_\approx$. (We can call the equations (2) together with $p_\approx = c^2\rho_\approx$ the "first-order acoustic wave equation system".) Using this relation in the two linearised PDEs (2) we obtain that the pressure satisfies the **wave equation**:

$$\frac{1}{c^2}\frac{\partial^2 p_\approx}{\partial t^2} - \Delta p_\approx = 0.$$

(Here we use that the divergence of the gradient is the Laplacian, $\Delta u = \operatorname{div}\nabla u$.) Since $\rho_\approx = \frac{1}{c^2}p_\approx$, also $\rho_\approx$ satisfies the same equation.

The wave equation has been obtained from the linearisation of two "conservation laws" (for mass and momentum) and a "constitutive relation" (relating p and ρ).

In this derivation we have neglected the effect of the non-linear advection term $(\mathbf{v}\cdot\nabla)\mathbf{v}$ and, implicitly, those of fluid viscosity and gravity. A brief discussion of the validity of these assumptions is in [BK00, p. 41]. Moreover, $|p_\approx| \ll p_0$ is true for typical sounds in air: e.g. acoustic pain threshold is between 63 and $200Pa$ (130–$140dB$), while ambient pressure is $p_0 \approx 101\,325Pa$, so $|p_\approx|/p_0 \approx 6\times10^{-4}$—$2\times10^{-3}$.

Exercise 1.1: (Acoustic velocity).

- Show that the velocity $\mathbf{v}_\approx$ satisfies the vector wave equation $\frac{1}{c^2}\frac{\partial^2 \mathbf{v}_\approx}{\partial t^2} - \nabla(\nabla\cdot\mathbf{v}_\approx) = \mathbf{0}$.
- Fix $n=3$. Prove that for all vector fields $\mathbf{F} \in C^2(\mathbb{R}^3)^3$ the following vector calculus identity holds

$$\nabla(\nabla\cdot\mathbf{F}) = \boldsymbol{\Delta}\mathbf{F} + \operatorname{curl}\operatorname{curl}\mathbf{F}. \tag{3}$$

  Here $\boldsymbol{\Delta}$ is the vector Laplacian (defined componentwise) and $\operatorname{curl}\mathbf{F} = \nabla\times\mathbf{F} = (\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3}, \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1}, \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2})$ is the usual curl operator.
- Deduce that if $\mathbf{v}_\approx$ is irrotational, i.e. curl-free ($\operatorname{curl}\mathbf{v}_\approx = \mathbf{0}$), then each Cartesian component $v_{\approx,1}, v_{\approx,2}, v_{\approx,3}$ is solution of the scalar wave equation: $\frac{1}{c^2}\frac{\partial^2 v_{\approx,j}}{\partial t^2} - \Delta v_{\approx,j} = 0$, $j=1,2,3$.
- Show that if $\mathbf{v}_\approx$ is irrotational at some given time t_0, then it remains irrotational for all $t > t_0$.

Exercise 1.2: (Velocity potential). Assume that, at some initial time $t = t_0$, the velocity $\mathbf{v}_\approx(\mathbf{x},t_0)$ is gradient of a (time-independent) scalar field $\frac{1}{\rho_0}\phi_0(\mathbf{x})$. Show that the scalar field $\phi(\mathbf{x},t) := \phi_0(\mathbf{x}) - \int_{t_0}^t p_\approx(\mathbf{x},s)\,\mathrm{d}s$, called **velocity potential**, satisfies $-\frac{\partial\phi}{\partial t} = p_\approx$ and $\frac{1}{\rho_0}\nabla\phi = \mathbf{v}_\approx$. Show that the velocity potential satisfies the wave equation $\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} - \Delta\phi = 0$.

[1] In some applications, the static (time-independent) background quantities ρ_0, p_0 and $\mathbf{v}_0$ are not constant. E.g. in ocean acoustics the background density and pressure typically depend on the position; in aero-acoustics the velocity of the fluid is non-zero and variable. This leads to the presence of extra terms in the linearised differential equations.

We have seen that several quantities (the acoustic pressure $p_\approx$, the density $\rho_\approx$, and, under suitable assumptions, the velocity potential ϕ and the components of the velocity $v_{\approx,j}$) satisfy the same wave equation, so we write it for a general scalar field U:

$$\boxed{\frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} - \Delta U = 0.} \tag{4}$$

This is the prototype of second-order, linear, hyperbolic PDEs.

Exercise 1.3: (Time-domain plane waves). Show that for any smooth function $F : \mathbb{R} \to \mathbb{R}$ and any unit vector $\mathbf{d} \in \mathbb{R}^n$, $|\mathbf{d}| = 1$, the field $U(\mathbf{x}, t) = \boxed{F(\mathbf{x}\cdot\mathbf{d} - ct)}$ is a solution of the wave equation (4).

Exercise 1.3 shows that any wave profile (imagine F as a pulse, e.g. ) move across space–time with speed c, which is thus called **wave speed**.[2] Indeed, $c = \sqrt{\frac{p_\approx}{\rho_\approx}}$ and the square root of the ratio between a pressure and a mass density has the dimension of a velocity (i.e. $\sqrt{\frac{kg\,m^{-1}s^{-2}}{kg\,m^{-3}}} = \frac{m}{s}$).

Exercise 1.4: (Damped wave equation). The **damped wave equation** (or equation of telegraphy, see [CJ77, §9]) with damping parameter $\gamma > 0$ is

$$\frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} + \frac{\gamma}{c^2}\frac{\partial U}{\partial t} - \Delta U = 0. \tag{5}$$

Assume that γ is small so that γ^2 can be neglected. Show that wave profiles are damped in time with rate $\frac{\gamma}{2}$ while they propagate: for any smooth function $F : \mathbb{R} \to \mathbb{R}$ and any unit vector $\mathbf{d} \in \mathbb{R}^n$, $|\mathbf{d}| = 1$ the field $U(\mathbf{x}, t) = F(\mathbf{x}\cdot\mathbf{d} - ct)\mathrm{e}^{-\frac{\gamma}{2}t}$ is a solution of the damped wave equation (5) up to a factor $-\frac{\gamma^2}{4c^2}U \approx 0$.

Conversely, show that if U is solution of the damped wave equation (5), then $W(\mathbf{x}, t) := \mathrm{e}^{\frac{\gamma}{2}t}U(\mathbf{x}, t)$ is solution of the wave equation (4) up to a factor $\frac{\gamma^2}{4c^2}W \approx 0$.

When the acoustic waves hits an obstacle $D \subset \mathbb{R}^n$ through which it cannot propagate, on the interface between the obstacle and the fluid some boundary conditions have to be imposed. Depending on the nature of the obstacle and of the fluid, different conditions can be imposed.

- If the obstacle is "**sound-soft**" then the acoustic pressure on its boundary vanishes, i.e. $p_\approx = 0$.
- If the obstacle is "**sound-hard**" then the normal velocity on the boundary vanishes, i.e. $\mathbf{v}_\approx \cdot \mathbf{n} = 0$, where $\mathbf{n}$ is the unit normal vector on the boundary of D.

  Both sound-soft and sound-hard boundary conditions reflect all the energy carried by the wave.
- A simple way to model a more realistic absorbing boundary condition is to impose that the normal velocity is proportional to the pressure: $\mathbf{v}_\approx \cdot \mathbf{n} = \frac{\vartheta}{c\rho_0}p_\approx$ for some $\vartheta > 0$ that represents how easily the obstacle yields to the acoustic pressure. Here we have assumed that $\mathbf{n}$ points outwards of the domain where the wave propagates and into the obstacle D. We have divided by $c\rho_0$ to ensure that ϑ is dimensionless (check this fact). Deriving this relations and using the linearised Euler's equation (2), we obtain a relation involving $p_\approx$ only: $\mathbf{n}\cdot\nabla p_\approx + \frac{\vartheta}{c}\frac{\partial p_\approx}{\partial t} = 0$. This is called "**impedance**" boundary condition.
- If the obstacle is made of a different fluid, then instead of imposing boundary conditions we consider two copies of the wave equations in the two fluids, with different values of c. The two equations are coupled by suitable "transmission conditions", i.e. by imposing the continuity of the pressure and the normal displacement across the interface.

  Similarly, if the obstacle is an elastic solid, acoustic waves in the fluid generates elastic waves in the solid and vice versa. This is modelled by coupling the acoustic wave equation (4) with the (more complicated) elastodynamic wave equation, whose unknown is the point displacement, through appropriate transmission conditions, [Ihl98, §1.3]. We briefly describe the equations of elastodynamics in §1.4.

[2] Thermodynamics states that sound waves are adiabatic, i.e. there is no heat loss, and that $p = a\rho^\gamma$, where a is a proportionality constant and $\gamma = 1.4$ for air (or for a diatomic ideal gas), see [BK00, p. 39 and §7.2, (7.33)]. Indeed, taking air pressure $p_0 = 1\text{atm} \approx 101\,325 N/m^2$ and air density $\rho_0 \approx 1.225 kg/m^3$ we have $c = \sqrt{\frac{\partial p}{\partial \rho}(\rho_0)} = \sqrt{\gamma a \rho_0^{\gamma-1}} = \sqrt{\gamma p_0/\rho_0} \approx 340 m/s$, which is the speed of sound in air.

In "aeroacoustics", sound is generated by fluid turbulence, thus the source of acoustic disturbance is distributed in the bulk of the fluid. This is modelled by the **inhomogeneous wave equation**:

$$\frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} - \Delta U = F, \tag{6}$$

where $F(\mathbf{x},t)$ is the source term. Vice versa, in "vibroacoustics" the sound is generated by vibrating structures immersed in the acoustic fluid. In this case the source of disturbance is imposed as a boundary condition, justifying the interest in the homogeneous wave equation (4).

1.2 Time-harmonic waves

A **time-harmonic** function is a scalar field whose time-dependence is prescribed to be sinusoidal, in the form[3]

$$\boxed{U(\mathbf{x},t) = \Re\{u(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t}\}} = \Re\{u(\mathbf{x})\}\cos\omega t + \Im\{u(\mathbf{x})\}\sin\omega t \tag{7}$$

for an **angular frequency** $\omega > 0$ and a complex-valued field u which depends on the position in space $\mathbf{x}$ but not on the time variable t. (Here $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote real and imaginary parts, and i is the imaginary unit.) A sound wave in the form (7) is a "pure tone".

By taking the Laplacian and the second time-derivative of (7) we obtain that

$$\frac{1}{c^2}\frac{\partial^2 U}{\partial t^2}(\mathbf{x},t) - \Delta U(\mathbf{x},t) = \Re\Big\{\Big(-\frac{\omega^2}{c^2}u - \Delta u\Big)\mathrm{e}^{-\mathrm{i}\omega t}\Big\}$$

and the following crucial fact.

> If $U(\mathbf{x},t)$ is a time-harmonic (7) solution of the wave equation (4),
> then $u(\mathbf{x})$ is solution of the Helmholtz equation (1) with **wavenumber** $k := \omega/c > 0$.

This is **the main reason of the interest in the Helmholtz equation: it describes all time-harmonic solutions of the wave equation**. Any solution of the Helmholtz equation has to be interpreted via (7): multiplying by $\mathrm{e}^{-\mathrm{i}\omega t}$ and taking the real part we obtain a "physical" field depending on space and time.

The angular frequency ω is measured in radians per second and is proportional to the time frequency f, which is measured in Hertz ($1\mathrm{Hz} = 1s^{-1}$), with the relation $\omega = 2\pi f$. For brevity, in the following we simply use the word "frequency" for ω. The higher the frequency ω and the wavenumber k, the more oscillatory are the solutions of the Helmholtz equation. Sounds that are audible by humans correspond to values of f ranging approximately between 20Hz and 20 000Hz.

Another important acoustic quantity is the **wavelength** $\boxed{\lambda := \frac{2\pi}{k}}$. From $k = \frac{\omega}{c}$ and $\omega = 2\pi f$, we have $\lambda = \frac{c}{f}$, thus the wavelengths of audible sounds in the atmosphere range between $17mm$ and $17m$. The wavelength is a characteristic length associated to the Helmholtz operator $\Delta + k^2$: as we will see in §2.2, λ is the spatial period of some important periodic Helmholtz solutions.

Sound-soft conditions translate to Dirichlet boundary conditions $u = 0$, sound-hard to Neumann $\mathbf{n}\cdot\nabla u = 0$, and impedance to Robin $\mathbf{n}\cdot\nabla u - \mathrm{i}k\vartheta u = 0$.

Exercise 1.5: (Amplitude and phase). Show that a time-harmonic function $U(\mathbf{x},t)$ as in (7) (not necessarily solution of any PDE) can be written in terms of an **amplitude** function $A(\mathbf{x}) \geq 0$ and a (real) **phase** function $\phi(\mathbf{x})$ as

$$U(\mathbf{x},t) = A(\mathbf{x})\cos\big(\omega(t-\phi(\mathbf{x}))\big).$$

Express A and ϕ in terms of u and vice versa.

At the point $\mathbf{x}$, the wave U oscillates between the values $-A(\mathbf{x})$ and $A(\mathbf{x})$. Different points $\mathbf{x}_1$ and $\mathbf{x}_2$ reach the maximum periodically at different times dictated by ϕ: show that they are synchronised if and only of $\phi(\mathbf{x}_1) - \phi(\mathbf{x}_2)$ is an integer multiple of $\frac{2\pi}{\omega}$.

Exercise 1.6: (Inhomogeneous wave and Helmholtz equations). Let U be time-harmonic as in (7) and be a solution of the inhomogeneous wave equation (6) with $F(\mathbf{x},t) = \Re\{f(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t}\}$. Show that u satisfies the inhomogeneous Helmholtz equation $-\Delta u - k^2 u = f$.

Exercise 1.7: (Time reversal). Show that complex conjugation of Helmholtz solutions "reverses time": if U is the space–time wave solution associated to the Helmholtz solution u, then the space–time wave solution W associated to $\overline{u} = \Re u - \mathrm{i}\Im u$ satisfies $W(\mathbf{x},t) = U(\mathbf{x},-t)$.

[3] Unfortunately, several references use the opposite convention $U(\mathbf{x},t) = \Re\{u(\mathbf{x})\mathrm{e}^{\mathrm{i}\omega t}\}$, with a different sign at the exponent. This causes changes in the signs and conjugation in all formulas in the following.

Exercise 1.8: (Helmholtz solutions oscillate around 0). Show that if u is a Helmholtz solution defined on an open set and $\mathbf{x}_*$ is an interior local maximum (minimum, respectively) of its real part, then $\Re u(\mathbf{x}_*) \geq 0$ ($\Re u(\mathbf{x}_*) \leq 0$, respectively). This means that $\Re u$ can look like  but not like .

Exercise 1.9: (Time-harmonic loop). Verify that for every time-harmonic U as in (7)

$$U(\mathbf{x},0) = \Re\{u(\mathbf{x})\}, \quad U\Big(\mathbf{x},\frac{\pi}{2\omega}\Big) = \Im\{u(\mathbf{x})\}, \quad U\Big(\mathbf{x},\frac{\pi}{\omega}\Big) = -\Re\{u(\mathbf{x})\}, \quad U\Big(\mathbf{x},\frac{3\pi}{2\omega}\Big) = -\Im\{u(\mathbf{x})\}.$$

This means that when the wave U has value equal to the real part of u, it will shift towards a value equal to the imaginary part of u. We will use this to determine direction a time-harmonic wave u is moving, from the comparison of his real and imaginary part.

Exercise 1.10: (Helmholtz equation with complex wavenumber). Show that if U is a time-harmonic solution of the damped wave equation (5) then it is solution of the Helmholtz equation with complex wavenumber k, such that $k^2 = \omega(\omega + \mathrm{i}\gamma)/c^2$. (We always choose the root k with $\Re k > 0$ and $\Im k \geq 0$).

This shows that the solutions of the Helmholtz equation with complex wavenumber k can be understood as waves that are attenuated while they propagate, i.e. they are absorbed by the medium through which they propagate. The larger the imaginary part of the wavenumber, the stronger the damping. A negative imaginary part of k corresponds to $\gamma < 0$ in (5) and to waves increasing in time, which is an unphysical situation.

Remark 1.11: (Waves in heterogeneous media). We have assumed that the medium through which the wave propagates is uniform. In the more general case of an acoustic wave propagating through heterogeneous materials, both the sound speed c and the static density ρ_0 depend on the position $\mathbf{x}$. Repeating the reasoning done above, one obtains the wave equation $\frac{1}{\rho_0(\mathbf{x})c^2(\mathbf{x})}\frac{\partial^2 p_\approx}{\partial t^2} - \operatorname{div}(\frac{1}{\rho_0(\mathbf{x})}\nabla p_\approx) = 0$. Assuming time-harmonic behaviour (7) for $U = p_\approx$ we have the Helmholtz equation with variable coefficients $\operatorname{div}(\frac{1}{\rho_0(\mathbf{x})}\nabla u) + \frac{\omega^2}{\rho_0(\mathbf{x})c^2(\mathbf{x})}u = 0$, which is often written as $\operatorname{div}(\frac{1}{\rho_0(\mathbf{x})}\nabla u) + k^2 n(\mathbf{x})u = 0$ and n is called refractive index. In the following we do not consider this more general problem and we stick to the constant-coefficients case; see e.g. [CK2, §8] for more details on this problem.

Remark 1.12: (Is the Helmholtz equation elliptic?). According to the standard classification of second-order linear PDEs, Helmholtz equation is clearly elliptic: its principal part (the second order term) is simply the Laplacian. Indeed it shares many properties with Laplace equation, e.g. "elliptic regularity" (all solutions are C^∞ in their domain). But often the word "elliptic" is used to denote problems that satisfy the assumptions of Lax–Milgram theorem. We will see that the typical variational forms of Helmholtz BVPs do not satisfy this requirement, so under this respect the equation is not elliptic. Indeed, it does not satisfy some other typical properties of elliptic PDEs, such as the maximum principle; moreover it is closely related to the wave equation, which is the prototypical hyperbolic equation.

Remark 1.13: (Helmholtz equation = wave equation + Fourier transform). Fourier analysis tells us that any "reasonable" (e.g. square-integrable) time-dependent field U can be written as a continuous linear combination of time-harmonic fields $\mathrm{e}^{\mathrm{i}\omega t}\widehat{U}(\mathbf{x},\omega)$ with different frequencies $\omega \in \mathbb{R}$, where $\widehat{U}$ is its Fourier transform (in time):

$$U(\mathbf{x},t) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\omega t}\widehat{U}(\mathbf{x},\omega)\,\mathrm{d}\omega \qquad \text{with} \qquad \widehat{U}(\mathbf{x},\omega) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\omega t}U(\mathbf{x},t)\,\mathrm{d}t.$$

Reasoning as above, we can verify that, if U is solution of the wave equation with wave speed c, then its Fourier transform $\widehat{U}$ evaluated at a given frequency ω, i.e. $u(\mathbf{x}) = \widehat{U}(\mathbf{x},\omega)$, is solution of the Helmholtz equation with wavenumber $k = \omega/c$. Thus any solution of the wave equation is a linear combination of infinitely many solutions of the Helmholtz equation at different wavenumbers. (Roughly speaking, this is how human ears process sound: different parts of the ear receive and transmit to the brain different frequencies.) Numerically, often one approximates a wave equation solution by solving several Helmholtz problems. This is an important reason for studying the Helmholtz equation, even if we were not interested in problems at a fixed frequency.

When we study U and the wave equation we say that we work "in time domain"; when we study $\widehat{U}$ or u and the Helmholtz equation we say that we work "in frequency domain".

1.3 Electromagnetism

Although the Helmholtz equation is usually associated to acoustic waves, it is important also in the modelling of other kinds of linear waves, e.g. electromagnetic ones. In this section and in the next one we fix $n = 3$, i.e. we consider 3D problems. Electromagnetic waves in a homogeneous material, in the absence of charges, are described by the **Maxwell's equations**:

$$\operatorname{curl} \mathcal{E}(\mathbf{x},t) + \mu \frac{\partial \mathcal{H}}{\partial t}(\mathbf{x},t) = \mathbf{0}, \qquad \operatorname{curl} \mathcal{H}(\mathbf{x},t) - \epsilon \frac{\partial \mathcal{E}}{\partial t}(\mathbf{x},t) - \sigma \mathcal{E}(\mathbf{x},t) = \mathbf{0}, \tag{8}$$

where $\mathcal{E}$ is the electric field, $\mathcal{H}$ the magnetic field, ϵ the electric permittivity, μ the magnetic permeability, and σ the conductivity. The parameters ϵ, μ and σ represent the properties of the material through which the wave propagates. As we consider a homogeneous, isotropic medium, ϵ and μ are positive constants, σ is a non-negative constant (0 in a dielectric, i.e. a perfect insulator, and positive in a conducting medium). The first equation is called Faraday law, the second one Ampère law.

If both the electric and the magnetic field are time-harmonic, i.e. $\mathcal{E}(\mathbf{x},t) = \Re\{\mathbf{E}(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t}\}$ and $\mathcal{H}(\mathbf{x},t) = \Re\{\mathbf{H}(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t}\}$ for some $\omega > 0$, then the time-independent fields $\mathbf{E}$ and $\mathbf{H}$ satisfy the **time-harmonic Maxwell's equations**:

$$\operatorname{curl} \mathbf{E}(\mathbf{x}) - \mathrm{i}\omega\mu \mathbf{H}(\mathbf{x}) = \mathbf{0}, \qquad \operatorname{curl} \mathbf{H}(\mathbf{x}) + \mathrm{i}\omega\epsilon \mathbf{E}(\mathbf{x}) - \sigma \mathbf{E}(\mathbf{x}) = \mathbf{0}. \tag{9}$$

These are two vector-valued PDEs with two vector fields as unknowns. Eliminating $\mathbf{H}$, we obtain the **second-order time-harmonic Maxwell's equations** for the electric field:

$$\operatorname{curl}\operatorname{curl} \mathbf{E} - k^2 \mathbf{E} = \mathbf{0} \qquad \text{with} \quad k^2 = \omega^2 \epsilon\mu + \mathrm{i}\omega\sigma\mu. \tag{10}$$

Since $\operatorname{div}\operatorname{curl} \mathbf{v} = 0$ for any vector field $\mathbf{v}$, any solution of (10) is divergence-free (solenoidal). Then the expansion (3) of the $\operatorname{curl}\operatorname{curl}$ operator implies that each component of the solution of the second-order Maxwell's equations (10) is solution of the Helmholtz equation with (possibly complex, if $\sigma > 0$) wavenumber k:

$$\Delta E_j + k^2 E_j = 0 \qquad \text{for} \quad j = 1,2,3.$$

The speed of propagation of electromagnetic waves (e.g. of light) is $c = \frac{1}{\sqrt{\epsilon\mu}} > 0$ and the damping factor (as in Exercise 1.10) is $\gamma = \Im\{k^2\}\frac{c^2}{\omega} = \frac{\sigma}{\epsilon} \geq 0$.

As any other PDE, time-harmonic Maxwell's equations are complemented by boundary conditions. When the domain under consideration is surrounded by a metal, through which the electric field does not penetrate, then typically one imposes the "**perfect electric conductor**" (PEC) boundary conditions, which impose that the *tangential* component of the electric field vanishes. In formulas this is $\mathbf{E} \times \mathbf{n} = \mathbf{0}$, where $\times$ denote the vector product and $\mathbf{n}$ is the unit normal vector on the boundary. In terms of the magnetic field, the PEC boundary conditions correspond to the vanishing of the *normal* component: $\mathbf{H} \cdot \mathbf{n} = 0$. This is easy to verify for a plane boundary, e.g. $\Pi = \{x_1 = 0\}$: in this case $\mathbf{E} \times \mathbf{n} = \mathbf{E} \times (1,0,0) = (0, E_3, -E_2)$ so $E_2 = E_3 = 0$ on the whole plane Π, and $\mathbf{H} \cdot \mathbf{n} = \frac{1}{\mathrm{i}\omega\mu} \operatorname{curl} \mathbf{E} \cdot (1,0,0) = \frac{1}{\mathrm{i}\omega\mu}(\frac{\partial E_3}{\partial x_2} - \frac{\partial E_2}{\partial x_3}) = 0$.

We also often encounter **impedance boundary conditions**: $\mathbf{H} \times \mathbf{n} - \vartheta(\mathbf{n} \times \mathbf{E}) \times \mathbf{n} = \frac{1}{\mathrm{i}k}\mathbf{g}$, or equivalently $\mu^{-1} \operatorname{curl} \mathbf{E} \times \mathbf{n} - \mathrm{i}k\vartheta(\mathbf{n} \times \mathbf{E}) \times \mathbf{n} = \mathbf{g}$, for a positive parameter ϑ and a boundary source term $\mathbf{g}$. Here $(\mathbf{n} \times \mathbf{E}) \times \mathbf{n} = \mathbf{E} - (\mathbf{E} \cdot \mathbf{n})\mathbf{n}$ is the tangential component of $\mathbf{E}$.

Exercise 1.14: (Maxwell-vs-Helmholtz). Complete the proof of the following statement. For $k \in \mathbb{C}$, $k \neq 0$, a vector field $\mathbf{v}$ is solution of $\operatorname{curl}\operatorname{curl} \mathbf{v} - k^2 \mathbf{v} = \mathbf{0}$ if and only if it is divergence-free and each of its three components is solution of the Helmholtz equation $\Delta v_j + k^2 v_j = 0$, $j = 1,2,3$.

Exercise 1.15: (Alternative derivation). We have shown that the components of the time-harmonic solutions of the Maxwell's equations (8) are Helmholtz solutions. Show again the same fact performing the same operations in different order. First eliminate $\mathcal{H}$ from (8) obtaining second-order Maxwell's equations in time-domain. Then verify that each component of $\mathcal{E}$ satisfies the wave equation (4). Finally assume that $\mathcal{E}$ is time-harmonic.

Remark 1.16: (Current density). Often the conductivity term $\sigma\mathcal{E}(\mathbf{x},t)$ in the time-domain Maxwell's equation is modelled as a *given* current density $\mathcal{J}(\mathbf{x},t)$ and treated as a datum. If this is assumed to be time-harmonic $\mathcal{J}(\mathbf{x},t) = \Re\{\mathbf{J}(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t}\}$ we obtain the inhomogeneous time-harmonic Ampère law $\operatorname{curl} \mathbf{H} + \mathrm{i}\omega\epsilon\mathbf{E} = \mathbf{J}$ and the second-order equation $\operatorname{curl} \frac{1}{\mu} \operatorname{curl} \mathbf{E} - \omega^2\epsilon\mathbf{E} = \mathrm{i}\omega\mathbf{J}$. In absence of charges, the current density is divergence free: $\operatorname{div} \mathbf{J} = 0$ (more generally we would have the continuity equation $\operatorname{div} \mathbf{J} = -\frac{\partial \rho}{\partial t}$, where ρ is the charge density), so the component of the electric field satisfy the inhomogeneous Helmholtz equation $\Delta E_j + k^2 E_j = -\mathrm{i}\omega\mu J_j$.

Remark 1.17: (1 Maxwell PDE $\Rightarrow$ 3 Helmholtz PDEs, 1 Maxwell BVP $\not\Rightarrow$ 3 Helmholtz BVPs). We have seen that time-harmonic Maxwell solutions are componentwise Helmholtz solutions. However, in general one cannot reduce the solution of a boundary value problem for the Maxwell's equations (10) to three independent Helmholtz problems for E_1, E_2, E_3, because the boundary conditions required are different.

For instance, when we impose PEC boundary conditions, only the tangential component of the electric field vanishes. This is equivalent to the imposition of two scalar boundary conditions (e.g. on two Cartesian components if the domain is a cube) for three unknown scalar fields and three scalar PDEs (Helmholtz); the boundary value problem is closed by the condition $\operatorname{div}\mathbf{E}=\mathbf{0}$ which intertwines the three components. Being able to solve/approximate Helmholtz BVPs is not enough to solve/approximate Maxwell BVPs.

We see in the next remark that the decoupling of the scalar components and the reduction of Maxwell's problems to Helmholtz ones can be performed when symmetries are present.

Remark 1.18: (TE and TM modes). ([Néd01, p. 5], [CJ77, §86], [BK00, §6.8]) The Helmholtz equation is important in dimensional reductions of the Maxwell's equations. Maxwell's equations simplify when we assume that the dependence on one of the Cartesian variables of all components of the fields is a given complex exponential, i.e.

$$\mathbf{E}(x_1,x_2,x_3)=\widetilde{\mathbf{E}}(x_1,x_2)\mathrm{e}^{\mathrm{i}\eta x_3},\qquad \mathbf{H}(x_1,x_2,x_3)=\widetilde{\mathbf{H}}(x_1,x_2)\mathrm{e}^{\mathrm{i}\eta x_3}. \tag{11}$$

This is relevant when we consider the propagation of waves through very long objects such as optical fibres. In this case the curl becomes

$$\operatorname{curl}\mathbf{H}=\mathrm{e}^{\mathrm{i}\eta x_3}\Big(\frac{\partial\widetilde{H}_3}{\partial x_2}-\mathrm{i}\eta\widetilde{H}_2,\ \ \mathrm{i}\eta\widetilde{H}_1-\frac{\partial\widetilde{H}_3}{\partial x_1},\ \ \frac{\partial\widetilde{H}_2}{\partial x_1}-\frac{\partial\widetilde{H}_1}{\partial x_2}\Big).$$

Maxwell's equations (9) (with $\sigma=0$) become

$$\begin{aligned}
&\frac{\partial\widetilde{E}_3}{\partial x_2}-\mathrm{i}\eta\widetilde{E}_2-\mathrm{i}\omega\mu\widetilde{H}_1=0, && \frac{\partial\widetilde{H}_3}{\partial x_2}-\mathrm{i}\eta\widetilde{H}_2+\mathrm{i}\omega\epsilon\widetilde{E}_1=0,\\
&\mathrm{i}\eta\widetilde{E}_1-\frac{\partial\widetilde{E}_3}{\partial x_1}-\mathrm{i}\omega\mu\widetilde{H}_2=0, && \mathrm{i}\eta\widetilde{H}_1-\frac{\partial\widetilde{H}_3}{\partial x_1}+\mathrm{i}\omega\epsilon\widetilde{E}_2=0,\\
&\frac{\partial\widetilde{E}_2}{\partial x_1}-\frac{\partial\widetilde{E}_1}{\partial x_2}-\mathrm{i}\omega\mu\widetilde{H}_3=0, && \frac{\partial\widetilde{H}_2}{\partial x_1}-\frac{\partial\widetilde{H}_1}{\partial x_2}+\mathrm{i}\omega\epsilon\widetilde{E}_3=0.
\end{aligned}$$

If $\widetilde{E}_3=0$, with some manipulation one can see that all other field components can be computed from $\widetilde{H}_3$ ($\widetilde{E}_1=-(\mathrm{i}\omega\epsilon-\frac{\mathrm{i}\eta^2}{\omega\mu})^{-1}\frac{\partial\widetilde{H}_3}{\partial x_2}$, $\widetilde{E}_2=(\mathrm{i}\omega\epsilon-\frac{\mathrm{i}\eta^2}{\omega\mu})^{-1}\frac{\partial\widetilde{H}_3}{\partial x_1}$, $\widetilde{H}_1=-\frac{\eta}{\omega\mu}\widetilde{E}_2$, $\widetilde{H}_2=\frac{\eta}{\omega\mu}\widetilde{E}_1$, and therefore $(\widetilde{H}_1,\widetilde{H}_2)=\frac{\mathrm{i}\eta}{\omega^2\epsilon\mu-\eta^2}\nabla\widetilde{H}_3$) and that $\widetilde{H}_3$ itself is solution of the 2D Helmholtz equation $\Delta\widetilde{H}_3+(\omega^2\epsilon\mu-\eta^2)H_3=0$. These solutions are called "**transverse-electric (TE) modes**", since the electric field is perpendicular to the x_3 axis, along which the wave propagates. Similarly, the "**transverse-magnetic (TM) modes**" are solutions with $\widetilde{H}_3=0$, which can be computed by solving the same Helmholtz equation for $\widetilde{E}_3$. All Maxwell solutions in form (11) are sum of a TE and a TM mode.

Now assume that we want to compute the solutions of Maxwell's equation in an infinite cylinder $\Omega\times\mathbb{R}=\{\mathbf{x}=(x_1,x_2,x_3)\in\mathbb{R}^3:(x_1,x_2)\in\Omega,x_3\in\mathbb{R}\}$, where $\Omega\subset\mathbb{R}^2$, and PEC boundary conditions are imposed on $\partial\Omega\times\mathbb{R}$. A given TE mode satisfies the PEC conditions $\widetilde{\mathbf{H}}\cdot\mathbf{n}=0$ if the Neumann condition $\mathbf{n}\cdot\nabla\widetilde{H}_3=0$ holds, while a TM mode has to satisfy the Dirichlet one $\widetilde{E}_3=0$. Thus there exists an electromagnetic wave propagating through the "waveguide" $\Omega\times\mathbb{R}$ with frequency η in the x_3 direction only if the 2D Helmholtz problem admits a non-trivial solution. This is the same as saying that $\omega^2\epsilon\mu-\eta^2$ is either a Dirichlet (TM) or a Neumann (TE) eigenvalue for the 2D Laplacian in Ω. Thanks to symmetry, 3D Maxwell's problems have been reduced to 2D Helmholtz ones.

Exercise 1.19: (TEM modes). Assume that the (non-trivial) pair $(\mathbf{E},\mathbf{H})$ is a "TEM mode", which means that it is simultaneously a TE and a TM mode: it is in the form (11) with $\widetilde{E}_3=\widetilde{H}_3=0$. Show the following facts.

- $\eta^2=\omega^2\epsilon\mu$, i.e. $\eta=k$: the wavenumber in the x_3 direction coincide with the free-space wavenumber.
- $\mathbf{H}=\sqrt{\epsilon/\mu}\,\mathbf{e}_3\times\mathbf{E}$, where $\mathbf{e}_3=(0,0,1)$. This means that $\mathbf{E}$ and $\mathbf{H}$ are orthogonal vectors in the x_1x_2-plane.
- If ϕ is a 2D harmonic function ($\Delta\phi=0$), then $\mathbf{E}=\nabla\phi(x_1,x_2)\mathrm{e}^{\mathrm{i}kx_3}$ and $\mathbf{H}=\sqrt{\epsilon/\mu}\,\mathbf{e}_3\times\mathbf{E}$ constitute a TEM mode.
- If the domain $\Omega\subset\mathbb{R}^2$ is simply connected then there is no non-trivial TEM mode with PEC conditions propagating through $\Omega\times\mathbb{R}$. (This is a main motivation for the use of coaxial cables.)

The Helmholtz equation is used in place of the Maxwell's equations when the effects of the wave polarisation (the direction in which the field points) are neglected. This is often done, for example, in the Fresnel, Fraunhofer and Kirchhoff descriptions of light diffraction by apertures.

Remark 1.20: (Reality is more complicated than this!). The setting considered in this section is a special case of much more general ones, which are needed in many applications. If different materials are present in the region considered, or the properties of the material vary in space, then ϵ, μ, σ are function of position. In this case, for instance, to obtain (10) we cannot simply move μ to the second term and find componentwise solution of the Helmholtz equation, but we obtain some more general elliptic equations. If the material is anisotropic, then the coefficients are modelled by symmetric positive definite matrices (semi-definite in case of σ). Since the polarisation of a material given an impinging electromagnetic field is not immediate, the multiplications $\epsilon\mathcal{E}$ and $\mu\mathcal{H}$ in (8) are more precisely modelled as convolutions in time between $\mathcal{E}/\mathcal{H}$ and suitable kernels; however in frequency-domain these give rise to standard products $\epsilon(\omega)\mathbf{E}$ and $\mu(\omega)\mathbf{H}$ where now the coefficients depends on the frequency ω. In some materials and regimes (e.g. in lasers) the coefficients ϵ and μ need to be modelled as non-linear operators acting on $\mathbf{E}$ and $\mathbf{H}$: this is the field of non-linear optics.

1.4 Elastodynamics

Mechanical vibrations propagating in solids have more complicated behaviour than those in fluids, as two different types of waves can be present. The relevant PDEs are similar to the wave and the Helmholtz equations, with some complications due to the fact that the unknown is a vector field and the differential operator (in the space variable $\mathbf{x}$) is not as simple as the Laplacian. The **Navier's equations** are the system of PDEs that describes small-amplitude vibrations in (homogeneous, isotropic) solid objects:

$$\rho\frac{\partial^2\mathbf{U}}{\partial t^2} = (\lambda+2\mu)\nabla\operatorname{div}\mathbf{U} - \mu\operatorname{curl}\operatorname{curl}\mathbf{U} + \mathbf{F}.$$

Here $\mathbf{U}(\mathbf{x},t)$ is the **displacement** vector field, describing the position of a material point of the object with respect to the rest position; $\mathbf{F}$ are the forces per unit of volume (e.g. gravity forces); the positive parameters λ and μ are the **Lamé constants**[4], describing the elastic properties of the material; and $\rho > 0$ is the mass density of the medium at rest. These are the main equation of "linear elastodynamics", while "elastostatics" studies the equilibrium case where $\mathbf{U}$ is independent of time (e.g. it studies small deformations of an object under a static load). In absence of external forces acting in the volume of the body ($\mathbf{F} = \mathbf{0}$), and if the waves are time-harmonic with angular frequency ω (i.e. (7) holds for each component of $\mathbf{U}$), the Navier's equations become

$$(\lambda+2\mu)\nabla\operatorname{div}\mathbf{u} - \mu\operatorname{curl}\operatorname{curl}\mathbf{u} + \omega^2\rho\,\mathbf{u} = \mathbf{0}. \tag{12}$$

We define the wavenumber of pressure (longitudinal) and shear (transverse) waves, respectively, as:

$$k_P := \omega\sqrt{\frac{\rho}{\lambda+2\mu}}, \qquad k_S := \omega\sqrt{\frac{\rho}{\mu}}.$$

We define the scalar and vector potential, respectively, as

$$\chi := -\frac{\lambda+2\mu}{\omega^2\rho}\operatorname{div}\mathbf{u} = -\frac{\operatorname{div}\mathbf{u}}{k_P^2}, \qquad \boldsymbol{\psi} := \frac{\mu}{\omega^2\rho}\operatorname{curl}\mathbf{u} = \frac{\operatorname{curl}\mathbf{u}}{k_S^2}. \tag{13}$$

From (12), we can use these potentials to represent $\mathbf{u}$:

$$\mathbf{u} = -\frac{\lambda+2\mu}{\omega^2\rho}\nabla\operatorname{div}\mathbf{u} + \frac{\mu}{\omega^2\rho}\operatorname{curl}\operatorname{curl}\mathbf{u} = \nabla\chi + \operatorname{curl}\boldsymbol{\psi}, \tag{14}$$

which is a "Helmholtz decomposition" of the displacement field. With some manipulation we obtain

$$\Delta\chi + k_P^2\chi \overset{(13),\Delta=\operatorname{div}\nabla}{=} -\operatorname{div}\nabla\frac{\operatorname{div}\mathbf{u}}{k_P^2} - \operatorname{div}\mathbf{u} \overset{(12)}{=} -\frac{1}{k_P^2}\operatorname{div}\left(\frac{\mu}{\lambda+2\mu}\operatorname{curl}\operatorname{curl}\mathbf{u} - k_P^2\mathbf{u}\right) - \operatorname{div}\mathbf{u} \overset{\operatorname{div}\operatorname{curl}=0}{=} 0,$$

$$\operatorname{curl}\operatorname{curl}\boldsymbol{\psi} - k_S^2\boldsymbol{\psi} \overset{(13)}{=} \operatorname{curl}\operatorname{curl}\frac{\operatorname{curl}\mathbf{u}}{k_S^2} - \operatorname{curl}\mathbf{u} \overset{(12)}{=} \frac{1}{k_S^2}\operatorname{curl}\left(\frac{\lambda+2\mu}{\mu}\nabla\operatorname{div}\mathbf{u} + k_S^2\mathbf{u}\right) - \operatorname{curl}\mathbf{u} \overset{\operatorname{curl}\nabla=\mathbf{0}}{=} \mathbf{0}.$$

[4] Sometimes these equations are written in terms of the Poisson ratio ν and Young's modulus E, which are other relevant material parameters related to the Lamé constants by the relations $\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}$ and $\mu = \frac{E}{2(1+\nu)}$.

This means that the scalar and vector potentials satisfy Helmholtz and Maxwell's equations, respectively.

The decomposition (14) shows that any solution $\mathbf{u}$ of Navier's equations (12) is sum of two terms. The first one is a curl-free, longitudinal, time-harmonic wave propagating at speed $c_P = \frac{\omega}{k_P} = \sqrt{\frac{\lambda+2\mu}{\rho}}$; this is called **pressure wave** (P-wave). The second one is a divergence-free, transverse, time-harmonic wave propagating at (lower) speed $c_S = \frac{\omega}{k_S} = \sqrt{\frac{\mu}{\rho}}$; this is called **shear wave** (S-wave). In seismology, P-waves and S-waves are called primary and secondary waves, respectively, because after an earthquake they reach a give point the surface in this order, due to their different speeds.

In particular, all time-harmonic elastic waves can be 'assembled' from solutions of two copies of the Helmholtz equation with different wavenumbers. In some applications, such as seismic imaging for oil retrieval, Navier's equations are sometimes approximated by the scalar Helmholtz equation, neglecting shear waves.

The limit $\mu \to 0$ corresponds to a fluid material, elasticity reduces to acoustics and shear waves disappear: Navier's equations tend to $\nabla \operatorname{div} \mathbf{u} + k_P^2 \mathbf{u} = \mathbf{0}$, which is the equation satisfied by the acoustic displacement and the acoustic velocity.

The Dirichlet boundary condition for the Navier's equations consists in imposing a given displacement on the bondary: $\mathbf{u} = \mathbf{g}$. The Neumann boundary condition requires the **traction** operator $\mathbf{T}(\mathbf{u}) := 2\mu \frac{\partial \mathbf{u}}{\partial \mathbf{n}} + \lambda \mathbf{n} \operatorname{div} \mathbf{u} + \mu \mathbf{n} \times \operatorname{curl} \mathbf{u}$, with $\mathbf{n}$ the outward-pointing unit normal; setting $\mathbf{T}(\mathbf{u}) = \mathbf{g}$ on the boundary corresponds to imposing the action of a force on the surface of the body.

When an elastic solid is in contact with a fluid, elastic vibration in the solid generate acoustic waves in the fluid, and vice versa. The simulation of this interaction is important for noise mitigation in vehicles and aircraft. On the boundary between the fluid and the solid domain, one has to impose "transmission conditions", to ensure the continuity of pressure and particle displacement (in formulas: $p\mathbf{n} = -\mathbf{T}(\mathbf{u})$ and $\mathbf{n} \cdot \nabla p = \omega^2 \rho\, \mathbf{u} \cdot \mathbf{n}$, where p is the fluid pressure), see [Ihl98, §1.3]. This is sometimes called "strong coupling". If the pressure forces of the fluid on the solid are negligible, then one can impose a "weak coupling": first compute the elastic vibrations of the solid and use them as input for the computation of the acoustic field.

Porous materials (e.g. soil, rocks, biological tissues, foam sound absorbers, ...) are solid materials containing pores, which are spaces filled by a fluid (water, oil, air, ...). The theory of *poroelasticity*, first developed by M.A. Biot, describes the mechanical vibrations of porous media and arises from the interaction between elastic waves in the solid structure and acoustic waves in the fluid component.

All properties mentioned here have an analogue in time-domain, as opposed to frequency-domain.

More information can be found e.g. in [BK00, §5] and [Ihl98, §1.2].

Remark 1.21: (Navier's equations with strain and stress tensors). We give some more notation and write Navier's equations in different equivalent forms to help relate other references to this section. Using identity (3), equation (12) can also be written as

$$(\lambda + \mu) \nabla \operatorname{div} \mathbf{u} + \mu \boldsymbol{\Delta} \mathbf{u} + \omega^2 \rho \mathbf{u} = \mathbf{0}.$$

We denote by $\underline{\underline{\mathbf{D}}}\mathbf{v}$ the Jacobian of the vector field $\mathbf{v}$, by $\underline{\underline{\mathbf{D}}}^S\mathbf{v} := \frac{1}{2}(\underline{\underline{\mathbf{D}}}\mathbf{v} + \underline{\underline{\mathbf{D}}}^\top \mathbf{v})$ the symmetric gradient, by $\mathbf{div}$ the (row-wise) vector divergence of matrix fields, and by $\underline{\underline{\mathrm{Id}}}$ the 3×3 identity matrix. The symmetric gradient of the displacement $\underline{\underline{\mathbf{D}}}^S(\mathbf{u})$ is called "Cauchy strain tensor" and often denoted $\underline{\underline{\boldsymbol{\varepsilon}}}$: it is a matrix field measuring the deformation of the solid body. Using the identity $2\,\mathbf{div}\,\underline{\underline{\mathbf{D}}}^S = \nabla \operatorname{div} + \boldsymbol{\Delta} = 2\nabla \operatorname{div} - \operatorname{curl}\operatorname{curl}$, equation (12) can be written in the form

$$\mathbf{div}\, \underline{\underline{\boldsymbol{\sigma}}} + \omega^2 \rho \mathbf{u} = \mathbf{0}, \qquad \text{where } \underline{\underline{\boldsymbol{\sigma}}} := 2\mu\, \underline{\underline{\mathbf{D}}}^S \mathbf{u} + \lambda\, (\operatorname{div} \mathbf{u})\underline{\underline{\mathrm{Id}}} = 2\mu\, \underline{\underline{\boldsymbol{\varepsilon}}} + \lambda \operatorname{Tr}(\underline{\underline{\boldsymbol{\varepsilon}}})\underline{\underline{\mathrm{Id}}}$$

is called "Cauchy stress tensor". Then the traction operator on the boundary can be written as $\mathbf{T}(\mathbf{u}) = \underline{\underline{\boldsymbol{\sigma}}}\mathbf{n}$ (the matrix–vector product between the stress tensor and the unit normal to the boundary). The elastic wave equations for more general anisotropic linear materials are still written in the form $\mathbf{div}\, \underline{\underline{\boldsymbol{\sigma}}} + \omega^2 \rho \mathbf{u} = \mathbf{0}$ but the strain and stress tensors are related by the more general relation $\underline{\underline{\boldsymbol{\sigma}}} = \mathsf{C}\underline{\underline{\boldsymbol{\varepsilon}}}$, where C is the fourth-order "stiffness tensor".

We mention a few other examples of oscillatory phenomena where the Helmholtz equation plays a role.

Remark 1.22: (Membrane vibrations). The small-amplitude vibrations of a membrane are described by the scalar wave equation (4) in two variables and, in the time-harmonic case, by the Helmholtz equation. This model is accurate under the assumptions that the rest position is flat, the displacement is small and vertical (so the vertical component of the displacement is the unknown, the horizontal components are zero), the membrane is perfectly flexible and elastic, in particular it does not resist to bending, and tension forces act only tangentially to the membrane. Typical examples are drums, loudspeakers, microphones and our own eardrums.

Exercise 1.23: (Helmholtz equation and Kirchhoff–Love plates). In mechanical engineering, flat, thin, elastic structures are called "plates" and often modelled by Kirchhoff–Love theory. The flat three-dimensional object is represented by a domain in $\mathbb{R}^2$. In the "pure bending" time-harmonic case, the plate displacement in the direction perpendicular to the plate satisfies the fourth-order equation $-\Delta^2 u + k^4 u = 0$, where $\Delta^2 = \Delta\Delta$ is the bi-Laplacian operator.

- Show that solutions of the Helmholtz equation $\Delta u + k^2 u = 0$ and of the reaction–diffusion equation $\Delta u - k^2 u = 0$ solve also the fourth-order equation.
- Let u be a smooth solution of $-\Delta^2 u + k^4 u = 0$. Define two fields $w_\pm := \Delta u \pm k^2 u$. Show that they satisfy $\Delta w_\pm \mp k^2 w_\pm = 0$ and that $u = \frac{1}{2k^2}(w_+ - w_-)$.

  This means that all time-harmonic Kirchhoff–Love solutions can be written as sums of oscillatory and boundary-layer components, that are solution of Helmholtz and reaction–diffusion equations, respectively.

Remark 1.24: (Helmholtz for water waves). The "shallow water equations" are a non-linear model for the propagation of small-amplitude water waves in the sea when the horizontal length scale is much greater than the sea depth, [BK00, §4.7]. Their linearisation leads to the two-dimensional wave equation, so under time-harmonic assumptions we obtain the Helmholtz equation once again, possibly with varying coefficients. In this setting the Helmholtz equation is sometimes called "Berkhoff equation".

2 Particular solutions of the Helmholtz equation

We now focus on the construction of some simple analytical solutions of the Helmholtz equation $\Delta u + k^2 u = 0$ in 2D and we study some of their qualitative properties. This is useful to understand some typical features of all Helmholtz solutions. Moreover, in the next sections we will use some of these functions to study some boundary value problems and to construct integral operators.

Plots and time-harmonic animations are available on the course webpage
https://euler.unipv.it/moiola/T/MNAPDE2023/MNAPDE2023anim.html

2.1 The one-dimensional case

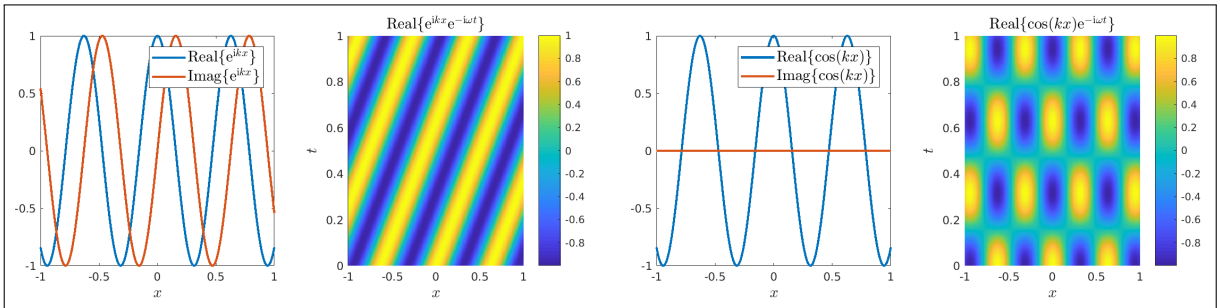


Figure 1: Panel 1: the real and the imaginary part of the one-dimensional propagative wave $u(x) = \mathrm{e}^{\mathrm{i}kx}$ over the interval $(-1, 1)$, with $k = 10$.
Panel 2: the real part of the corresponding solution of the wave equation $U(x,t) = \Re\{\mathrm{e}^{\mathrm{i}kx}\mathrm{e}^{-\mathrm{i}\omega t}\}$, plotted as function of x (horizontal axis) and t (vertical axis). Here we are taking $\omega = k$ and $c = 1$. The wave propagates towards the right endpoint of the space interval.
Panels 3 and 4: the same plots for the stationary wave $u(x) = \cos(kx) = \frac{1}{2}(\mathrm{e}^{\mathrm{i}kx} + \mathrm{e}^{-\mathrm{i}kx})$. The waves oscillates in time but does not propagate: the peaks (yellow parts) appear at the same locations in space.

We begin with the (boring) simpler case of one space dimension ($n = 1$). In this case, the Helmholtz equation reduces to the ordinary differential equation: $u'' + k^2 u = 0$. All solutions are in the form

$$u(x) = c_1 \cos(kx) + c_2 \sin(kx) \qquad \text{for some } c_1, c_2 \in \mathbb{C}.$$

Equivalently

$$u(x) = C_1 \mathrm{e}^{\mathrm{i}kx} + C_2 \mathrm{e}^{-\mathrm{i}kx} \qquad \text{for some } C_1, C_2 \in \mathbb{C}.$$

All 1D Helmholtz solutions are periodic with period $\lambda = \frac{2\pi}{k}$.

Let us fix $c = 1$, so $\omega = kc = k$. When we expand the time-dependence of the corresponding solutions (7) of the wave equation, we see that $u(x) = \mathrm{e}^{\mathrm{i}kx}$ corresponds to $U(x,t) = \Re\{\mathrm{e}^{\mathrm{i}kx-\mathrm{i}kt}\} = \cos(k(x-t))$, which is a wave propagating to the right. On the other hand, $u(x) = \cos(kx)$ corresponds to $U(x,t) = \Re\{\cos(kx)\mathrm{e}^{-\mathrm{i}kt}\} = \cos(kx)\cos(kt)$, which oscillates in time but maintains the same space profile and does not propagate. See Figure 1 and the animations.

2.2 Plane waves

We have seen in Exercise 1.3 that the space–time field $U(\mathbf{x},t) = F(\mathbf{x}\cdot\mathbf{d} - ct)$, propagating in the direction of $\mathbf{d}$ at speed c, is solution of the wave equation (here $\mathbf{d} \in \mathbb{R}^2$ is a unit vector and F a smooth real function). To have a Helmholtz solution, we want U to be time-harmonic, i.e. $U(\mathbf{x},t) = \Re\{u(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t}\}$. A simple way to reconcile these two expressions is to choose $F(z) = \Re\{\mathrm{e}^{\mathrm{i}kz}\}$ so that $U(\mathbf{x},t) = \Re\{\mathrm{e}^{\mathrm{i}(k\mathbf{x}\cdot\mathbf{d}-\omega t)}\} = \cos(k\mathbf{x}\cdot\mathbf{d} - \omega t)$ (recalling that $\omega = kc$) and

$$u(\mathbf{x}) = \boxed{\mathrm{e}^{\mathrm{i}k\mathbf{x}\cdot\mathbf{d}}} = \cos(k\mathbf{x}\cdot\mathbf{d}) + \mathrm{i}\sin(k\mathbf{x}\cdot\mathbf{d}).$$

This is a time-harmonic **propagative plane wave**, which propagates in the direction $\mathbf{d}$. Plane waves are arguably the simplest solutions of the Helmholtz equation.

Another way to obtain plane waves is to look for Helmholtz solutions that are independent of one of the Cartesian variables. If $u(x_1,x_2) = \tilde{u}(x_1)$, then $\tilde{u}$ has to satisfy $\tilde{u}'' + k^2\tilde{u} = 0$, so $\tilde{u}(x_1) = c_1\cos(kx_1) + c_2\sin(kx_2)$ for some $c_1, c_2 \in \mathbb{C}$. Propagative plane waves correspond to the choice $c_1 = 1$, $c_2 = \mathrm{i}$.

Plane waves have constant **amplitude** $|u(\mathbf{x})|$ and are constant on the lines perpendicular to $\mathbf{d}$. Their complex argument $\arg(u(\mathbf{x})) = k\mathbf{x}\cdot\mathbf{d}$ in a point $\mathbf{x}$ is called **phase**. Plane waves are periodic in the direction $\mathbf{d}$ with period equal to the wavelength $\lambda = \frac{2\pi}{k}$: the wavelength is precisely the distance between two consecutive peaks of a plane wave. A translation along a vector $\mathbf{v}$ corresponds to a multiplication by a complex factor of absolute value 1, i.e. it is a phase shift: $u(\mathbf{x}+\mathbf{v}) = \mathrm{e}^{\mathrm{i}k(\mathbf{x}+\mathbf{v})\cdot\mathbf{d}} = \mathrm{e}^{\mathrm{i}k\mathbf{v}\cdot\mathbf{d}}u(\mathbf{x})$.

The sum and the difference of two plane waves with opposite directions are called **stationary**, or **standing**, plane waves:

$$\mathrm{e}^{\mathrm{i}k\mathbf{x}\cdot\mathbf{d}} + \mathrm{e}^{-\mathrm{i}k\mathbf{x}\cdot\mathbf{d}} = 2\cos(k\mathbf{x}\cdot\mathbf{d}), \qquad \mathrm{e}^{\mathrm{i}k\mathbf{x}\cdot\mathbf{d}} - \mathrm{e}^{-\mathrm{i}k\mathbf{x}\cdot\mathbf{d}} = 2\mathrm{i}\sin(k\mathbf{x}\cdot\mathbf{d}).$$

As in the one-dimensional case of §2.1, the reason why these are called stationary while $\mathrm{e}^{\mathrm{i}k\mathbf{x}\cdot\mathbf{d}}$ is called propagative is clear if one looks at the evolution in time of the corresponding time-domain wave $U(\mathbf{x},t) = \Re\{u(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t}\}$; see Figure 2 and the animations.[5]

Exercise 2.1: (Conjugate of a plane wave). Show that the complex-conjugate of a plane wave is a plane wave propagating in the opposite direction, in accordance with Exercise 1.7.

Exercise 2.2: (Periodicity of sum of plane wave). Show that a linear combination of two propagative plane waves with directions the unit vectors $\mathbf{d}$ and $\mathbf{p}$, with $\mathbf{d} \neq \pm\mathbf{p}$, is periodic both in the direction of the bisectrix $\mathbf{b} = \mathbf{d} + \mathbf{p}$ and in the direction of its orthogonal $\mathbf{b}^\perp = \mathbf{d} - \mathbf{p}$. Show that the period of the linear combination in these two directions is $\frac{\lambda}{\cos(\alpha/2)}$ and $\frac{\lambda}{\sin(\alpha/2)}$, respectively, where α is the angle between $\mathbf{d}$ and $\mathbf{p}$.

Exercise 2.3: (Vector plane waves). Show that the vector plane wave $\mathbf{E}(\mathbf{x}) = \mathbf{A}\mathrm{e}^{\mathrm{i}k\mathbf{x}\cdot\mathbf{d}}$ is solution of Maxwell's equations $\operatorname{curl}\operatorname{curl}\mathbf{E} - k^2\mathbf{E} = \mathbf{0}$ if and only if $\mathbf{d}\cdot\mathbf{d} = 1$ and $\mathbf{d}\cdot\mathbf{A} = 0$. This means that the amplitude vector is orthogonal to the propagation direction, i.e. electromagnetic plane waves are transverse waves. (The formula $\mathbf{u}\times(\mathbf{v}\times\mathbf{w}) = \mathbf{v}(\mathbf{u}\cdot\mathbf{w}) - \mathbf{w}(\mathbf{u}\cdot\mathbf{v})$ might help.)

Show that Navier's equations (12) support both transverse plane waves $\mathbf{A}\mathrm{e}^{\mathrm{i}k_S\mathbf{x}\cdot\mathbf{d}}$, with $\mathbf{d}\cdot\mathbf{A} = 0$, and longitudinal ones $\mathbf{d}\mathrm{e}^{\mathrm{i}k_P\mathbf{x}\cdot\mathbf{d}}$. Longitudinal elastic waves are faster and have longer wavelengths than transverse ones.

Here $\mathbf{A} \in \mathbb{C}^3$. If $\Re\mathbf{A}$ and $\Im\mathbf{A}$ are parallel to one another, then the plane wave $\mathbf{u}(\mathbf{x}) = \mathbf{A}\mathrm{e}^{\mathrm{i}k\mathbf{x}\cdot\mathbf{d}}$ is said to have linear polarisation. If $\Re\mathbf{A}$ and $\Im\mathbf{A}$ are perpendicular to one another and $|\Re\mathbf{A}| = |\Im\mathbf{A}|$ (so that $\Im\mathbf{A} = \pm\Re\mathbf{A}\times\mathbf{d}$), then the polarisation is circular. In all other cases the polarisation is called elliptical. To understand the meaning of these names, draw the graph of $t \mapsto \mathbf{U}(\mathbf{x},t) = \Re\{\mathbf{u}(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t}\}$ in the plane perpendicular to $\mathbf{d}$, for a given $\mathbf{x}$. See also [BK00, §6.5].

[5] In one of the animations available online you can observe a sketch of the motion of the fluid particles subject to a time-harmonic plane wave: each particle oscillates back and forth harmonically around a fixed position, and never moves far, even if the wave propagates over the whole space. This happens for all solutions of the Helmholtz and the wave equations: pressure, energy and momentum are transported while matter oscillates but does not move away.

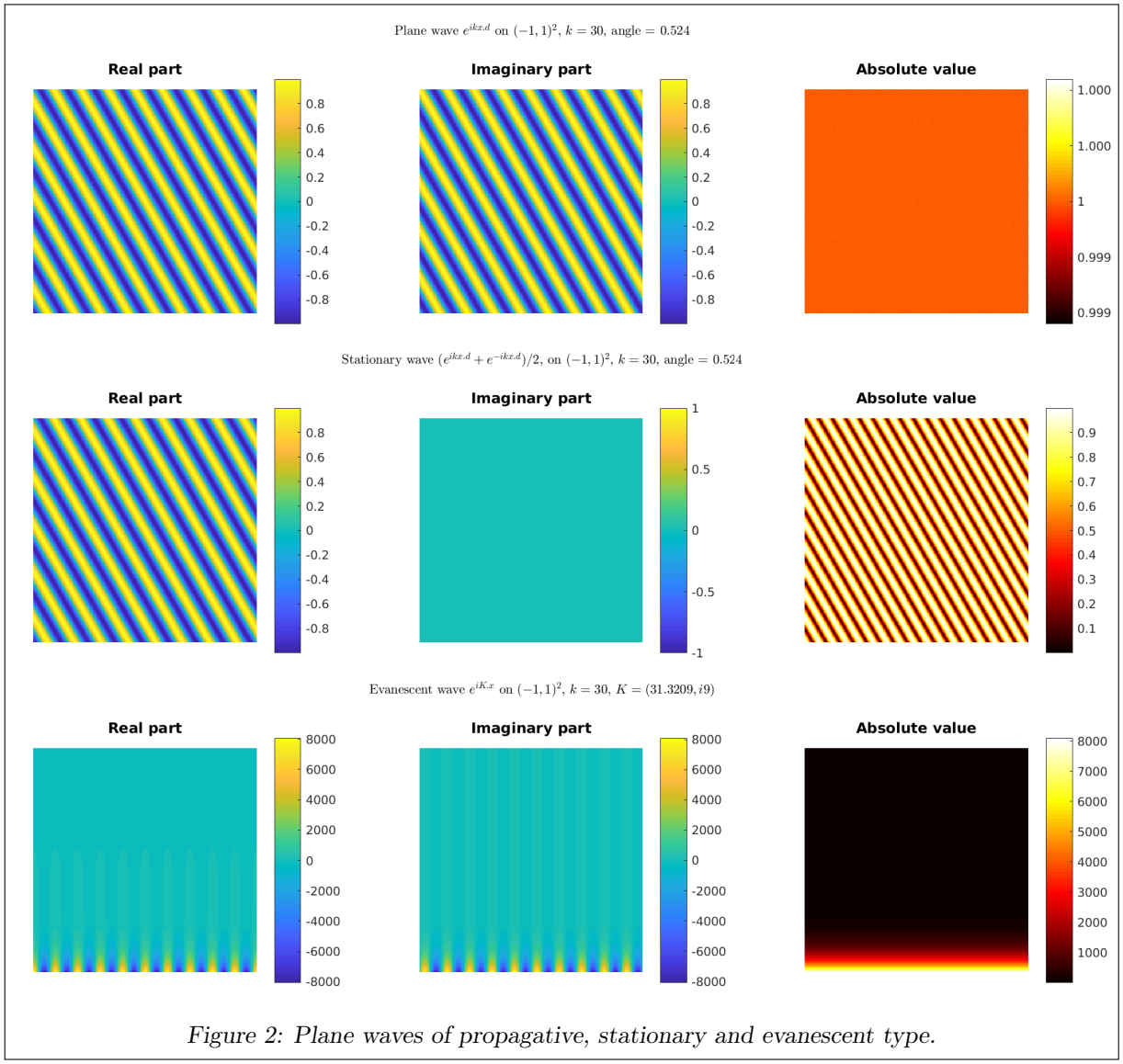


Figure 2: Plane waves of propagative, stationary and evanescent type.

2.2.1 Evanescent plane waves

Propagative and stationary waves are not the only solutions of the Helmholtz equation that are "separable" in Cartesian coordinates (i.e. that can be written as $u(\mathbf{x}) = u_1(x_1)u_2(x_2)$). If we look for functions in the form $u(\mathbf{x}) = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}} = \mathrm{e}^{\mathrm{i}(k_1x_1+k_2x_2)}$ satisfying $\Delta u + k^2u = 0$, we see that we need a "wavevector" $\mathbf{k} \in \mathbb{C}^2$ with $\mathbf{k}\cdot\mathbf{k} = k_1^2 + k_2^2 = k^2$. If both k_1 and k_2 are real then we obtain again the plane waves. If at least one of the two is not real then we have a new kind of waves, called evanescent (plane) waves. Expanding $\mathbf{k} = \mathbf{k}_R + \mathrm{i}\mathbf{k}_I$ with $\mathbf{k}_R, \mathbf{k}_I \in \mathbb{R}^2$, we have $\boxed{u(\mathbf{x}) = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}} = \mathrm{e}^{\mathrm{i}\mathbf{k}_R\cdot\mathbf{x}}\mathrm{e}^{-\mathbf{k}_I\cdot\mathbf{x}}}$: this field oscillates in the direction $\mathbf{k}_R$ with wavenumber $|\mathbf{k}_R| \geq k$ and decays exponentially in the orthogonal direction $\mathbf{k}_I$ ($|u(\mathbf{x})| = \mathrm{e}^{-\mathbf{k}_I\cdot\mathbf{x}}$). The orthogonality of $\mathbf{k}_R$ and $\mathbf{k}_I$ is a consequence of $\mathbf{k}\cdot\mathbf{k} \in \mathbb{R}$. See Figure 2 for a representation.

Evanescent (non-plane) waves typically appear at the interface between different materials or near boundaries; important examples in elasticity are Rayleigh waves, which include the surface waves generated by earthquakes.

Exercise 2.4: (Evanescent plane wave computations). Verify the statements made in the paragraph.

Exercise 2.5: (Complex parametrisation of plane waves). Show that all 2D plane waves, either propagative or evanescent, can be written in the form $\mathrm{e}^{\mathrm{i}k(x_1\cos\theta + x_2\sin\theta)} = \mathrm{e}^{\frac{k}{2}(\mathrm{i}(\nu+\frac{1}{\nu})x_1 + (\nu-\frac{1}{\nu})x_2)}$, parametrised by $0 \neq \nu \in \mathbb{C}$ or $\theta \in \mathbb{C}$, with $\nu = \mathrm{e}^{\mathrm{i}\theta}$.

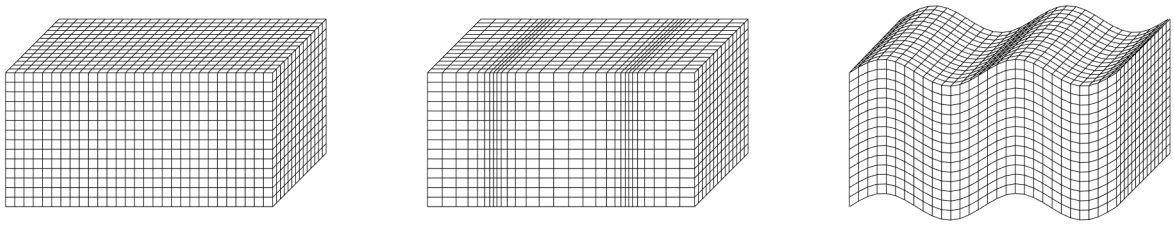


Figure 3: The deformation of an elastic medium perturbed by plane waves. Left: the unperturbed medium. Centre: a longitudinal plane wave $\mathbf{d}\mathrm{e}^{\mathrm{i}k\mathbf{x}\cdot\mathbf{d}}$. *Right: a transverse plane wave* $\mathbf{A}\mathrm{e}^{\mathrm{i}k\mathbf{x}\cdot\mathbf{d}}$ *with* $\mathbf{A}\cdot\mathbf{d}=0$. *In this example the wave propagates horizontally* ($\mathbf{d}=(1,0,0)$) *and* $\mathbf{A}=(0,0,1)$. *See §1.4 and Exercise 2.3. Acoustic waves are longitudinal waves, since fluids cannot support transverse waves.*

2.3 Circular waves and Bessel functions

We have seen Helmholtz solutions that are separable in Cartesian coordinates, we now look for those that are separable in the polar coordinates (r,θ), where $(x_1,x_2)=(r\cos\theta, r\sin\theta)$. The 2D Laplacian in polar coordinates reads

$$\Delta u = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 u}{\partial\theta^2}=\frac{\partial^2 u}{\partial r^2}+\frac{1}{r}\frac{\partial u}{\partial r}+\frac{1}{r^2}\frac{\partial^2 u}{\partial\theta^2}.$$

If we have a separable Helmholtz solution $u(\mathbf{x})=f(r)g(\theta)$, the functions f,g have to satisfy

$$f''(r)g(\theta)+\frac{1}{r}f'(r)g(\theta)+\frac{1}{r^2}f(r)g''(\theta)+k^2f(r)g(\theta)=\Delta u+k^2u=0.$$

The angular component g has to be periodic of period 2π, so we take the circular harmonic $g(\theta)=\mathrm{e}^{\mathrm{i}\ell\theta}$, for $\ell\in\mathbb{Z}$. Then $g''(\theta)=-\ell^2 g(\theta)$, so we can cancel g from the expression above, multiply by r^2, and obtain that f satisfies

$$r^2f''(r)+rf'(r)+(r^2k^2-\ell^2)f(r)=0. \tag{15}$$

For $k=1$, this is called **Bessel differential equation**: it is a linear, second-order ODE with variable coefficients, it depends on the index ℓ and degenerates at $r=0$. Two linearly independent real-valued solutions are the **Bessel functions of the first kind** and order ℓ, denoted $\boxed{J_\ell(r)}$, and the **Bessel function of the second kind** (or Neumann functions) and order ℓ, denoted $\boxed{Y_\ell(r)}$. Explicit expressions (e.g. as power series or integral representations), plenty of useful formulas and graphs can be found online on the "NIST Digital Library of Mathematical Functions" [DLMF]. See also Appendix B.

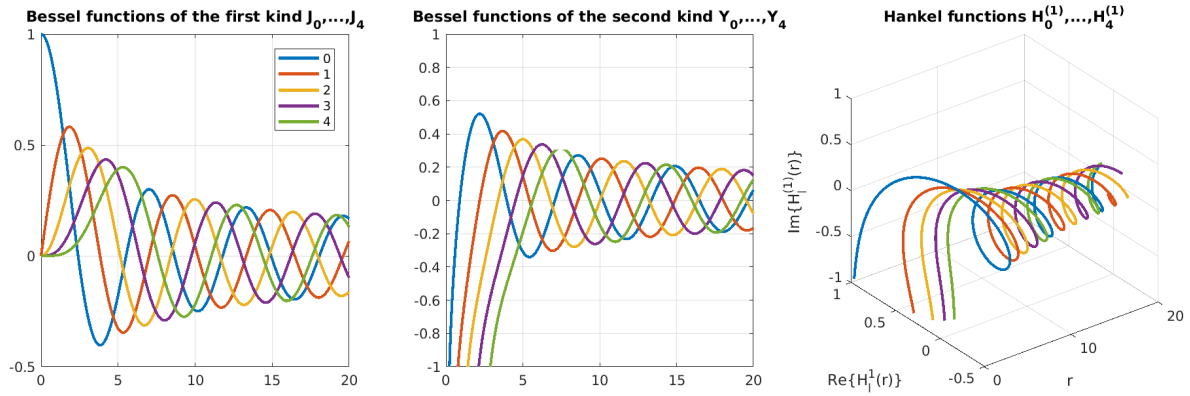


Figure 4: The Bessel functions of first and second kind, and the Hankel functions, for $\ell=0,\dots,4$.

The Bessel functions of the first and second kind for $\ell=0,\dots,4$ are plotted in Figure 4. We see that both families of functions oscillate around 0 and decay slowly for $r\to\infty$. The distance between two successive zeros of either J_ℓ or Y_ℓ is slightly shorter than π for $\ell=0$ and slightly longer than π for $\ell\neq 0$. The main difference is that the $J_\ell(r)$s are smooth over $\mathbb{R}$, while the $Y_\ell(r)$s have a singularity at $r=0$; the higher ℓ the stronger the singularity. Useful formulas are $J_{-\ell}(r)=(-1)^\ell J_\ell(r)$ and $Y_{-\ell}(r)=(-1)^\ell Y_\ell(r)$.

The **Hankel functions** (sometimes called Bessel functions of the third kind) are the complex-valued linear combinations

$$\boxed{H_\ell^{(1)}(r)} := J_\ell(r) + \mathrm{i}Y_\ell(r), \qquad \boxed{H_\ell^{(2)}(r)} := J_\ell(r) - \mathrm{i}Y_\ell(r) = \overline{H_\ell^{(1)}(r)}. \tag{16}$$

The right panel of Figure 4 shows the first few Hankel functions: the argument r is one of the axis, the real and the imaginary parts of $H_\ell^{(1)}(r)$ are on the other two axes. An important property of the Hankel functions is that the magnitude $r \mapsto |H_\ell^{(1)}(r)|$ is a monotonically decreasing function ($|H_\ell^{(1)}(r)| \approx \sqrt{2/(\pi r)}$ for large r). For increasing r the complex number $H_\ell^{(1)}(r)$ spirals clockwise towards the origin.

Bessel and Hankel functions can be used in Matlab with the commands `besselj`, `bessely` and `besselh`.

Exercise 2.6: (Bessel equation). Verify that if f_1 is solution of (15) for $k = 1$, then, for any $k > 0$, $f_k(r) := f_1(kr)$ solves (15).

Exercise 2.7: (Bessel function asymptotics). Compare numerically the plots of the Bessel functions against the asymptotics for small and large (positive) arguments (from [DLMF, §10.7]):

$$\begin{aligned} &J_\ell(z) \sim \frac{z^\ell}{\ell!\, 2^\ell} \quad \ell \in \mathbb{N}_0, \qquad Y_0(z) \sim \frac{2}{\pi}\log z, \qquad Y_\ell(z) \sim -\frac{(\ell-1)!\, 2^\ell}{\pi z^\ell} \quad \ell \in \mathbb{N}, \qquad z \to 0, \\ &J_\ell(z) \sim \sqrt{\frac{2}{\pi z}}\cos\Big(z - \frac{\ell\pi}{2} - \frac{\pi}{4}\Big), \qquad Y_\ell(z) \sim \sqrt{\frac{2}{\pi z}}\sin\Big(z - \frac{\ell\pi}{2} - \frac{\pi}{4}\Big) \quad \ell \in \mathbb{N}_0, \qquad z \to \infty. \end{aligned} \tag{17}$$

From what we have said, we deduce that for any $\ell \in \mathbb{Z}$ the two fields

$$\boxed{J_\ell(kr)\mathrm{e}^{\mathrm{i}\ell\theta}, \qquad Y_\ell(kr)\mathrm{e}^{\mathrm{i}\ell\theta}}$$

and their linear combinations are the solutions of the Helmholtz equations that are separable in polar coordinates. They are called **circular waves** or **Fourier–Bessel** functions. Of all the elements of the 2-dimensional space $\operatorname{span}\{J_\ell(kr)\mathrm{e}^{\mathrm{i}\ell\theta}, Y_\ell(kr)\mathrm{e}^{\mathrm{i}\ell\theta}\}$, only $J_\ell(kr)\mathrm{e}^{\mathrm{i}\ell\theta}$ is defined in the whole of $\mathbb{R}^2$, while all the others are defined in the punctured plane $\mathbb{R}^2 \setminus \{0\}$. They are all of class C^∞ in their domain of definition. From the angular dependence, we see that all these function are invariant under rotations of an angle multiple of $2\pi/|\ell|$; a rotation by an angle α corresponds to a multiplication by a complex factor $\mathrm{e}^{\mathrm{i}\ell\alpha}$.

Special circular waves are the **Fourier–Hankel** functions, namely

$$\boxed{H_\ell^{(1)}(kr)\mathrm{e}^{\mathrm{i}\ell\theta}} = J_\ell(kr)\mathrm{e}^{\mathrm{i}\ell\theta} + \mathrm{i}Y_\ell(kr)\mathrm{e}^{\mathrm{i}\ell\theta}, \qquad \boxed{H_\ell^{(2)}(kr)\mathrm{e}^{\mathrm{i}\ell\theta}} = J_\ell(kr)\mathrm{e}^{\mathrm{i}\ell\theta} - \mathrm{i}Y_\ell(kr)\mathrm{e}^{\mathrm{i}\ell\theta}, \qquad \ell \in \mathbb{Z}.$$

We will see soon why the the Fourier–Hankel functions $H_\ell^{(1)}(kr)\mathrm{e}^{\mathrm{i}\ell\theta}$ are important for problems posed in unbounded domains.

Plotting the time evolution (7) of these fields, one notes that the Fourier–Bessel functions $J_\ell(kr)\mathrm{e}^{\mathrm{i}\ell\theta}$ and $Y_\ell(kr)\mathrm{e}^{\mathrm{i}\ell\theta}$ rotate around the origin (anticlockwise if $\ell > 0$, clockwise if $\ell < 0$) and do not propagate in the radial direction. The Fourier–Hankel functions $H_\ell^{(1)}(kr)\mathrm{e}^{\mathrm{i}\ell\theta}$ rotate and move towards infinity, while the $H_\ell^{(2)}(kr)\mathrm{e}^{\mathrm{i}\ell\theta}$ towards the origin. This can be seen in the animations on the course webpage. Figure 5 shows some circular waves.

Exercise 2.8: (Singular circular waves are not in H^1). Show that the Fourier–Bessel functions $Y_\ell(kr)\mathrm{e}^{\mathrm{i}\ell\theta}$ and all the Fourier–Hankel functions do not belong to $H^1(\Omega)$, for any domain Ω containing the origin $\mathbf{0}$, because of the singularity at that point.

Use the small-argument asymptotics (17), the derivative formula $Y_\ell' = \frac{1}{2}(Y_{\ell-1} - Y_{\ell+1})$ [DLMF, eq. 10.6.1], and recall how to compute gradients and integrals in polar coordinates.

Remark 2.9: (Special Helmholtz solutions in 3D). Plane waves in $\mathbb{R}^3$ are defined exactly as in 2D.

The 3D analogous of circular waves are called "spherical waves". In their expression, Bessel and Hankel functions are substituted by the similar "spherical Bessel functions" and "spherical Hankel functions", denoted $j_\ell, y_\ell, h_\ell^{(1)}, h_\ell^{(2)}$. The angular component $\mathrm{e}^{\mathrm{i}\ell\theta}$ is substituted by the "spherical harmonics" Y_ℓ^m, which are smooth functions defined on the unit sphere and indexed by two indices ℓ and m. All these functions are described in details in e.g. [Néd01, §2.4, 2.6] or [CK2, §2.3, 2.4].

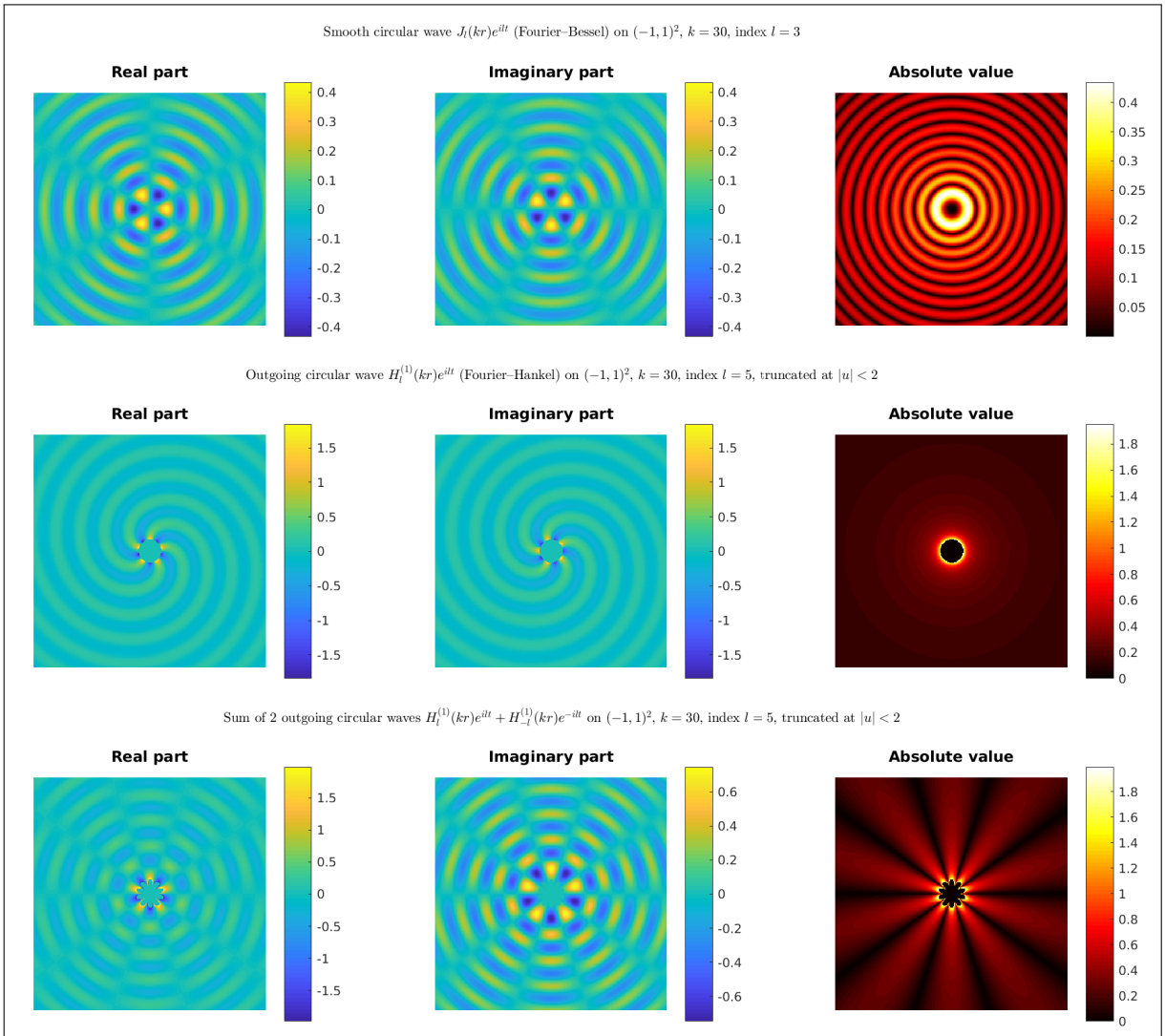


Figure 5: The Fourier–Bessel function $J_3(kr)\mathrm{e}^{\mathrm{i}3\theta}$, the Fourier–Hankel function $H_5^{(1)}(kr)\mathrm{e}^{\mathrm{i}5\theta}$, and the sum of two Fourier–Hankel functions with opposite indices $H_5^{(1)}(kr)\mathrm{e}^{\mathrm{i}5\theta}+H_{-5}^{(1)}(kr)\mathrm{e}^{-\mathrm{i}5\theta}$. Since the Hankel functions are unbounded at the origin, the field has been truncated to zero in the region where its absolute value is larger than 2.

2.4 Other remarks on the Helmholtz equation

Remark 2.10: (Herglotz functions). For $g\in L^2(0,2\pi)$, the field $u(\mathbf{x})=\int_0^{2\pi} g(\varphi)\mathrm{e}^{\mathrm{i}k(x_1\cos\varphi+x_2\sin\varphi)}\,\mathrm{d}\varphi\in C^\infty(\mathbb{R}^2)$ is called **Herglotz function** with kernel g. It can be thought as a continuous linear combination of plane waves with different directions $(\cos\varphi,\sin\varphi)$ weighted by $g(\varphi)$. Some interesting cases are the following.

- When g approximates a Dirac δ function centred at $\varphi_\star$ then u approximates the plane wave with direction $\mathbf{d}=(\cos\varphi_\star,\sin\varphi_\star)$.
- When g is constant in a small interval of $(0,2\pi)$ and 0 otherwise, then u approximates a plane wave in a strip of the plane and decays away from it. In some applications this is more realistic than a plane wave, which has an infinite propagating front. You can see the plot of such a Herglotz function in Figure 6.
- When g is a circular harmonic $g(\varphi)=\mathrm{e}^{\mathrm{i}\ell\varphi}$ we obtain a Fourier–Bessel function $u(\mathbf{x})=(2\pi\mathrm{i}^\ell)J_\ell(kr)\mathrm{e}^{\mathrm{i}\ell\theta}$.
  (Prove this fact using the Jacobi–Anger formula $\mathrm{e}^{\mathrm{i}z\cos\alpha}=\sum_{\ell\in\mathbb{Z}}\mathrm{i}^\ell J_\ell(z)\mathrm{e}^{\mathrm{i}\ell\alpha}$ and the $L^2(0,2\pi)$-orthogonality of circular harmonics.)

If u is a Herglotz function with kernel g, its translate $\tilde u(\mathbf{x})=u(\mathbf{x}+\mathbf{c})$ is also a Herglotz function with kernel $\tilde g(\varphi)=\mathrm{e}^{\mathrm{i}k(c_1\cos\varphi+c_2\sin\varphi)}g(\varphi)$.

Plot with Matlab some Herglotz functions with different kernels.

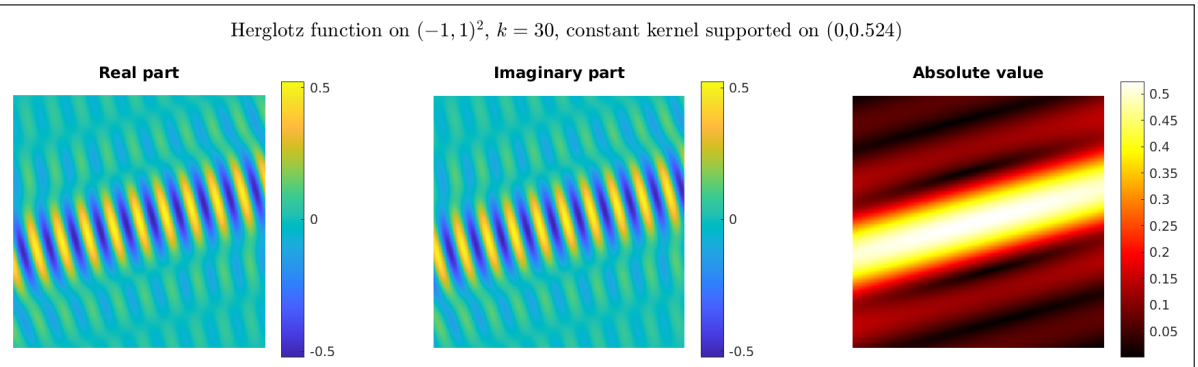


Figure 6: The Herglotz function with kernel $g(\varphi)=1$ if $0<\varphi<\frac{\pi}{6}$ and 0 otherwise, on $(-1,1)^2$. See Remark 2.10.

Exercise 2.11: (PDEs for phase and amplitude). Let c,ω be positive constants. Assume that the real-valued phase function ϕ satisfies the non-linear "eikonal equation" $|\nabla\phi|=c^{-1}$ and that the complex-valued amplitude A satisfies the complex, ϕ-dependent, diffusion–transport–reaction equation $\Delta A+2\mathrm{i}\omega\nabla\phi\cdot\nabla A+(\mathrm{i}\omega\Delta\phi)A=0$.

This can be interpreted as saying that ϕ varies at constant speed and that A is advected in the direction of variation of ϕ and it spreads. (Note that differently from Exercise 1.5 here we allow A to take complex values.)

- Prove that $u(\mathbf{x})=A(\mathbf{x})\mathrm{e}^{\mathrm{i}\omega\phi(\mathbf{x})}$ solves the Helmholtz equation (with $k=\frac{\omega}{c}$).
- Show that some affine ϕ and any constant A satisfy these conditions and allow to construct plane waves.
- Show $\phi(\mathbf{x})=|\mathbf{x}|$ and $A(\mathbf{x})=H_0^{(1)}(k|\mathbf{x}|)\mathrm{e}^{-\mathrm{i}k|\mathbf{x}|}$ satisfy the eikonal equation and the diffusion–transport–reaction equation in $\mathbb{R}^2\setminus\{\mathbf{0}\}$, respectively (with $c=1$ and $k=\omega$). (Use Bessel differential equation 15.)

  Check numerically that A does not oscillate.

Careful: not all Helmholtz solutions can be written in this form; often u is a sum of terms $u=\sum_j A_j\mathrm{e}^{\mathrm{i}\omega\phi_j}$, each of them with well-defined phase and amplitude.

Exercise 2.12: (Helmholtz and Schrödinger equations). Let u be a Helmholtz solution defined on a domain in the xy-plane. Fix a number $k_0>0$.

- Assume that $u(x,y)=\mathrm{e}^{\mathrm{i}k_0y}\psi(x,y)$, where the "envelope" ψ is a field that varies slowly in the variable y, more precisely that $|\frac{\partial^2\psi}{\partial y^2}|\ll|k\frac{\partial\psi}{\partial y}|$. Show that ψ approximately satisfies the Schrödinger equation

$$2\mathrm{i}k_0\frac{\partial\psi}{\partial y}+\frac{\partial^2\psi}{\partial x^2}+(k^2-k_0^2)\psi=0.$$

- Now assume that u can be written in polar coordinates as $u(r,\theta)=H_0^{(1)}(k_0r)\Psi(r,\theta)$ where Ψ varies slowly in the radial variable: $|\frac{\partial^2\Psi}{\partial r^2}|\ll|k\frac{\partial\Psi}{\partial r}|$. Using the approximation (17) for $k_0r\gg1$, show that Ψ approximately satisfies the Schrödinger equation with (r,θ) in place of (x,y).

This equation is used to model waves propagating in directions close to a leading one (either the y direction or the radial one). This regime is called "paraxial approximation". We factored out the leading term (either a plane or a circular wave), and showed that the if the remainder behaves smoothly in the dominant direction, then we can approximate it with a PDE that is first-order in this variable (y or r).

This PDE is also called "parabolic wave equation" and has the same mathematical form of the linear Schroedinger equation used in quantum mechanics, where the variable y is the time variable.

Exercise 2.13: (The Helmholtz Poynting vector). The Poynting vector $\mathbf{S}:=\Re\{\mathbf{E}\times\overline{\mathbf{H}}\}$ (often defined with a multiplicative constant) denotes the direction in which the energy of a time-harmonic electromagnetic field $(\mathbf{E},\mathbf{H})$ flows. Note the conjugation on $\mathbf{H}$.

- Compute the Poynting vector of the plane wave in Exercise 2.3 (you first need to compute $\mathbf{H}$).
- Show that the Poynting vector of a TE and a TM mode ((11) with $\widetilde{E}_3=0$ and $\widetilde{H}_3=0$, respectively) are

$$\mathbf{S}_{\mathrm{TE}}=\Re\Big\{\frac{\mathrm{i}\omega\mu}{\eta^2-\omega^2\epsilon\mu}\overline{\widetilde{H}_3}\nabla\widetilde{H}_3+\frac{\eta}{\omega\mu}(|\widetilde{E}_1|^2-|\widetilde{E}_2|^2)\hat{\mathbf{e}}_3\Big\},$$
$$\mathbf{S}_{\mathrm{TM}}=\Re\Big\{\frac{\mathrm{i}\omega\epsilon}{\omega^2\epsilon\mu-\eta^2}\widetilde{E}_3\nabla\overline{\widetilde{E}_3}+\frac{\omega\epsilon\eta}{(\eta^2-\omega\epsilon\mu)^2}|\nabla\widetilde{E}_3|^2\hat{\mathbf{e}}_3\Big\}.$$

Show that if $\eta = 0$ then $\mathbf{S}_{\mathrm{TE}} = \Im\{\frac{1}{\omega\epsilon}\overline{\widetilde{H}_3}\nabla\widetilde{H}_3\}$ and $\mathbf{S}_{\mathrm{TM}} = \Im\{\frac{1}{\omega\mu}\overline{\widetilde{E}_3}\nabla\widetilde{E}_3\}$ lie in the x_1x_2-plane.

Recalling that $\widetilde{H}_3$ and $\widetilde{E}_3$ are Helmholtz solutions in two dimensions, this suggests to define the Poynting vector for a 2D Helmholtz solution u as

$$\mathbf{S} = \mathbf{S}(u) := \Im\Big\{\frac{1}{k}\overline{u}\nabla u\Big\} = \Im\Big\{-\frac{1}{k}u\nabla\overline{u}\Big\} = \Re\Big\{\frac{1}{\mathrm{i}k}\overline{u}\nabla u\Big\}.$$

Verify the following facts.

- If u is Helmholtz solution, its Poynting vector is solenoidal: $\nabla\cdot\mathbf{S} = 0$.
- If u is a complex multiple of a real field, i.e. a standing wave, then $\mathbf{S} = \mathbf{0}$.
- The Poynting vector of a propagating or evanescent plane wave identifies the propagation direction: $\mathbf{S}(\mathrm{e}^{\mathrm{i}k\mathbf{x}\cdot\mathbf{d}}) = \mathbf{d}$ and $\mathbf{S}(\mathrm{e}^{\mathrm{i}\mathbf{k}_R\cdot\mathbf{x}}\mathrm{e}^{-\mathbf{k}_I\cdot\mathbf{x}}) = \frac{\mathrm{e}^{-2\mathbf{k}_I\cdot\mathbf{x}}}{k}\mathbf{k}_R$.
- For a field written in polar coordinates as $u(\mathbf{x}) = f(r)g(\theta)$, we have $\nabla u = f'g\hat{\mathbf{r}} + \frac{1}{r}fg'\hat{\boldsymbol{\theta}}$, so that $\mathbf{S} = \frac{1}{k}\Im\{f'\overline{f}|g|^2\hat{\mathbf{r}} + \frac{1}{r}|f|^2g'\overline{g}\hat{\boldsymbol{\theta}}\}$.
- $$\mathbf{S}\big(\mathrm{e}^{\mathrm{i}\ell\theta}J_\ell(kr)\big) = \frac{\ell}{kr}|J_\ell(kr)|^2\hat{\boldsymbol{\theta}}, \qquad \mathbf{S}\big(\mathrm{e}^{\mathrm{i}\ell\theta}H_\ell^{(1)}(kr)\big) = \frac{\ell}{kr}|H_\ell^{(1)}(kr)|^2\hat{\boldsymbol{\theta}} + \frac{2}{\pi kr}\hat{\mathbf{r}}.$$

  This is consistent with the fact that smooth Fourier–Bessel functions rotate around the origin (anticlockwise for $\ell > 0$ and clockwise for $\ell < 0$) and do not propagates radially, while Fourier–Hankel functions rotate and simultaneously propagate outwards.

  Hint: use the Wronskian identity $J_\ell(z)Y_\ell'(z) - Y_\ell(z)J_\ell'(z) = \frac{2}{\pi z}$ ([DLMF, eq. 10.5.E2]).

Compare the results with the animations on the webpage.

3 Analytical tools

We introduce a few mathematical tools that will be useful in the following. Much more detail on the content of this section can be found e.g. in the first part of [SBH19].

3.1 Lipschitz domains

We say that an open set $\Omega \subset \mathbb{R}^2$ with bounded boundary is Lipschitz if (1) there is a finite family of open sets $\{W_j\}_{j=1,\dots,J}$ that cover $\partial\Omega$ (i.e. $\partial\Omega \subset \bigcup_{j=1,\dots,J} W_j$), (2) there is a family of rotated Lipschitz hypographs $\{\Omega_j\}_{j=1,\dots,J}$ (i.e. $\Omega_j = \{(x_1', x_2') : x_2' < f_j(x_1')\}$, where (x_1', x_2') is a system of rotated Cartesian coordinates in $\mathbb{R}^2$ and f_j is a Lipschitz function), and (3) $W_j \cap \Omega = W_j \cap \Omega_j$.

Intuitively, for each $\mathbf{x} \in \partial\Omega$, there is a neighbourhood where the boundary can be represented as the graph of a Lipschitz function, and Ω lies only on one side of $\partial\Omega$. Smooth domains and polygons are Lipschitz. Domains with cusps (such as $\{0 < \sqrt{x_2} < x_1 < 1\}$) or cracks (such as $\{|\mathbf{x}| < 1, x_2 \neq 0\}$) are not allowed.

We also say that Ω is of class C^m, $m = 0, 1, \dots, \infty$, if the functions f_j are of class C^m.

An important property of Lipschitz domains is that the unit normal vector field $\mathbf{n}$ is defined almost everywhere on their boundary (a.e. with respect to the 1-dimensional measure). E.g. on a polygon the unit normal is defined everywhere except that at corners.

3.2 Function spaces on Lipschitz domains

To study boundary value problems we need some function spaces. Let $\Omega \subset \mathbb{R}^2$ be an open, Lipschitz set. We denote by $\boxed{\mathscr{D}(\Omega)}$ the space of the "test functions": complex-valued C^∞ functions defined on Ω whose support is compactly contained in Ω.

We denote by $\boxed{L^2(\Omega)}$ the usual Lebesgue space of complex-valued, square-integrable functions. This is a Hilbert space with inner product $(v, w)_{L^2(\Omega)} = \int_\Omega v\overline{w}\,\mathrm{d}\mathbf{x}$ and norm $\|v\|^2_{L^2(\Omega)} = \int_\Omega |v|^2\,\mathrm{d}\mathbf{x}$. Since we are dealing with complex-valued functions, the inner product is a sesquilinear form (with a complex conjugation on the second entry) and the norm requires the use of the absolute value of the argument.

Definition 3.1: ($H^1(\Omega)$ and $H^1_0(\Omega)$). The Sobolev space $\boxed{H^1(\Omega)}$ is space of complex-valued $L^2(\Omega)$ functions, whose first (distributional) partial derivatives are in $L^2(\Omega)$. It is provided with the following seminorm, norm, and inner product:

$$|v|^2_{H^1(\Omega)} := \|\nabla v\|^2_{L^2(\Omega)^2} = \left\|\frac{\partial v}{\partial x_1}\right\|^2_{L^2(\Omega)} + \left\|\frac{\partial v}{\partial x_2}\right\|^2_{L^2(\Omega)}, \qquad \|v\|^2_{H^1(\Omega)} := |v|^2_{H^1(\Omega)} + \|v\|^2_{L^2(\Omega)},$$

$$(v,w)_{H^1(\Omega)} := \left(\frac{\partial v}{\partial x_1}, \frac{\partial w}{\partial x_1}\right)_{L^2(\Omega)} + \left(\frac{\partial v}{\partial x_2}, \frac{\partial w}{\partial x_2}\right)_{L^2(\Omega)} + (v,w)_{L^2(\Omega)} = \int_\Omega (\nabla v \cdot \nabla \overline{w} + v\overline{w})\,\mathrm{d}\mathbf{x}.$$

The space $\boxed{H^1_0(\Omega)}$ is the subspace of $H^1(\Omega)$ of the elements that can be approximated in $H^1(\Omega)$ norm by a sequence of elements of $\mathscr{D}(\Omega)$.

Remark 3.2: (Distributional derivatives). When we say that a partial derivative "in the sense of distributions" $\frac{\partial v}{\partial x_1}$ of an $L^2(\Omega)$ function is in $L^2(\Omega)$, we mean that there is a function $w \in L^2(\Omega)$ such that $\int_\Omega v \frac{\partial \varphi}{\partial x_1} = -\int_\Omega w\varphi$ for all $\varphi \in \mathscr{D}(\Omega)$.

[6] We define also the subspace of $H^1(\Omega)$ of the functions with square-integrable (distributional) Laplacian: $\boxed{H^1(\Omega;\Delta)} = \{v \in H^1(\Omega) : \Delta v \in L^2(\Omega)\}$.

Finally, we say that $u \in \boxed{H^1_{\mathrm{loc}}(\overline{\Omega})}$ if the restriction of u to any bounded open subset D of Ω belongs to $H^1(D)$, [Sayas06, p. 12]. If Ω is bounded, then $H^1_{\mathrm{loc}}(\overline{\Omega}) = H^1(\Omega)$, while if Ω is unbounded then the "local space" $H^1_{\mathrm{loc}}(\overline{\Omega})$ includes functions that do not decay at infinity. For instance, all plane and circular waves belong to $H^1_{\mathrm{loc}}(\overline{\Omega}) \setminus H^1(\Omega)$ if Ω is the complement of a bounded set (which must contain the origin, if the circular waves are the singular ones). We haven't defined a norm on $H^1_{\mathrm{loc}}(\overline{\Omega})$, so this is not a Hilbert space; on the other hand the $H^1(D)$ norms are seminorms on $H^1_{\mathrm{loc}}(\overline{\Omega})$. Similarly, $u \in \boxed{H^1_{\mathrm{loc}}(\overline{\Omega};\Delta)}$ if the restriction of u to any bounded open subset D of Ω belongs to $H^1(D;\Delta)$.

The elements of $H^1(\Omega)$ are in general not continuous, so their point evaluation is not well-defined. However we will see that their values on the boundary of Ω, or any other Lipschitz curve, are well-defined.

3.3 Spaces on boundaries

3.3.1 The circle

We will need spaces of functions defined on boundary of Lipschitz sets with regularity weaker than H^1 and stronger than L^2. How to define functions with "half derivative"?

Let $\mathbb{S}^1 = \{\mathbf{x} \in \mathbb{R}^2, \|\mathbf{x}\| = 1\}$ be the unit circle. For a function v defined on $\mathbb{S}^1$ we write $v(\theta)$, with $\theta \in [0, 2\pi]$, for its value in polar coordinates. We say that $v \in L^2(\mathbb{S}^1)$ if $\|v\|^2_{L^2(\mathbb{S}^1)} := \int_0^{2\pi} |v|^2\,\mathrm{d}\theta < \infty$ and $v \in H^1(\mathbb{S}^1)$ if $\|v\|^2_{H^1(\mathbb{S}^1)} := \int_0^{2\pi}(|v|^2 + |v'|^2)\,\mathrm{d}\theta < \infty$, where v' is the derivative in the angular coordinate.

The expansion of v in circular harmonics is $v(\theta) = \sum_{\ell\in\mathbb{Z}} \hat{v}_\ell \mathrm{e}^{\mathrm{i}\ell\theta}$ for a sequence of coefficients $\hat{v}_\ell \in \mathbb{C}$, which can be computed as $\hat{v}_\ell = \frac{1}{2\pi}\int_0^{2\pi} v(\theta)\mathrm{e}^{-\mathrm{i}\ell\theta}\,\mathrm{d}\theta$. We can compute the $L^2(\mathbb{S}^1)$ scalar product and the $L^2(\mathbb{S}^1)/H^1(\mathbb{S}^1)$ norms using this expansion, exploiting the orthogonality $\int_0^{2\pi} \mathrm{e}^{\mathrm{i}\ell\theta}\mathrm{e}^{-\mathrm{i}\ell'\theta}\,\mathrm{d}\theta = 2\pi\delta_{\ell,\ell'}$ and the derivation formula $\frac{\partial}{\partial\theta}\mathrm{e}^{\mathrm{i}\ell\theta} = \mathrm{i}\ell\mathrm{e}^{\mathrm{i}\ell\theta}$:

$$(v,w)_{L^2(\mathbb{S}^1)} = \int_0^{2\pi} v(\theta)\overline{w(\theta)}\,\mathrm{d}\theta = \int_0^{2\pi} \sum_{\ell\in\mathbb{Z}} \hat{v}_\ell \mathrm{e}^{\mathrm{i}\ell\theta} \sum_{\ell'\in\mathbb{Z}} \overline{\hat{w}_{\ell'}\mathrm{e}^{\mathrm{i}\ell'\theta}}\,\mathrm{d}\theta = 2\pi\sum_{\ell\in\mathbb{Z}} \hat{v}_\ell\overline{\hat{w}_\ell}, \tag{18}$$

$$\|v\|^2_{L^2(\mathbb{S}^1)} = \int_0^{2\pi} |v|^2\,\mathrm{d}\theta = 2\pi\sum_{\ell\in\mathbb{Z}} |\hat{v}_\ell|^2, \qquad \|v\|^2_{H^1(\mathbb{S}^1)} = \int_0^{2\pi}(|v|^2 + |v'|^2)\,\mathrm{d}\theta = 2\pi\sum_{\ell\in\mathbb{Z}} |\hat{v}_\ell|^2(1+\ell^2).$$

(Write down the intermediate computations.) Thus, a function defined on the circle is in $L^2(\mathbb{S}^1)$ if the sequence of its Fourier coefficients is an element of the sequence space $l^2(\mathbb{Z}) = \{(a_\ell), \ell \in \mathbb{Z}, \|(a_\ell)\|^2_{l^2} := \sum_{\ell\in\mathbb{Z}} |a_\ell|^2 < \infty\}$, and in $H^1(\mathbb{S}^1)$ if its Fourier coefficients weighted with $(1+\ell^2)^{1/2}$ are in $l^2(\mathbb{Z})$. This means that the Fourier coefficients of functions in $H^1(\mathbb{S}^1)$ must decay faster for $\ell \to \infty$ (faster than $|\ell|^{-3/2}$) than those of the general function of $L^2(\mathbb{S}^1)$ (for which a decay slightly faster than $|\ell|^{-1/2}$ is enough).

[6] The two terms in the definition of the $H^1(\Omega)$ norm and inner product are not dimensionally homogeneous, so when dealing with Helmholtz problems we often weigh the norm with the wavenumber $k > 0$ as $\|v\|^2_{H^1_k(\Omega)} := |v|^2_{H^1(\Omega)} + k^2\|v\|^2_{L^2(\Omega)}$.

This suggests a way to define Sobolev spaces with other regularities:

$$\boxed{\|v\|^2_{H^s(\mathbb{S}^1)} := 2\pi \sum_{\ell\in\mathbb{Z}} |\hat v_\ell|^2 (1+\ell^2)^s,} \qquad \boxed{H^s(\mathbb{S}^1)} := \left\{ v(\theta) = \sum_{\ell\in\mathbb{Z}} \hat v_\ell \mathrm{e}^{\mathrm{i}\ell\theta} : \ \|v\|_{H^s(\mathbb{S}^1)} < \infty \right\} \quad \forall s\in\mathbb{R}. \tag{19}$$

For $s=0$ and $s=1$ we recover $H^0(\mathbb{S}^1) = L^2(\mathbb{S}^1)$ and $H^1(\mathbb{S}^1)$ as defined above. High-order Fourier coefficients correspond to rapidly oscillating components: imposing that they decay faster in ℓ is the same as imposing some regularity on v. So the higher s the smoother are the elements of $H^s(\mathbb{S}^1)$: for any real $s_- < s_+$, $H^{s_+}(\mathbb{S}^1) \subset H^{s_-}(\mathbb{S}^1)$ and $\|v\|_{H^{s_-}(\mathbb{S}^1)} \le \|v\|_{H^{s_+}(\mathbb{S}^1)}$. If $s > 1/2$, the elements of $H^s(\mathbb{S}^1)$ are continuous functions, if $s\ge 0$ they are $L^2(\mathbb{S}^1)$ classes of equivalence, for $s<0$ they can only be understood as distributions.

Exercise 3.3: (Dense embeddings). Show that for all $s_- < s_+$ the space $H^{s_+}(\mathbb{S}^1)$ is a *dense* subspace of $H^{s_-}(\mathbb{S}^1)$. This means that for all $v\in H^{s_-}(\mathbb{S}^1)$ and $\epsilon>0$ there is $w\in H^{s_+}(\mathbb{S}^1)$ with $\|v-w\|_{H^{s_-}(\mathbb{S}^1)} \le \epsilon$. Show also that $H^{s_+}(\mathbb{S}^1) \ne H^{s_-}(\mathbb{S}^1)$.

Hint: look for a common subspace of all $H^s(\mathbb{S}^1)$ that is dense in each of them.

Exercise 3.4: (Special elements of $H^s(\mathbb{S}^1)$).

- Compute the Fourier series of $\chi(\theta) = \begin{cases} 1 & 0<\theta<\pi, \\ 0 & \pi<\theta<2\pi. \end{cases}$ Show that $\chi\in H^s(\mathbb{S}^1)$ if and only if $s<\frac12$.
- Can you find a $v\in H^{\frac12}(\mathbb{S}^1)\setminus C^0(\mathbb{S}^1)$?
- Show that a delta function $\delta_{\theta_\star}$, $\theta_*\in[0,2\pi]$, belongs to $H^s(\mathbb{S}^1)$ if and only if $s<-\frac12$.

Exercise 3.5: (Random Sobolev functions on boundaries). We want to visualise how the decay of the Fourier coefficients $\hat v_\ell$ affects the regularity of $v(\theta) = \sum_{\ell\in\mathbb{Z}} \hat v_\ell \mathrm{e}^{\mathrm{i}\ell\theta}$. To this purpose, generate and plot a function v on $\mathbb{S}^1$ with random Fourier coefficients that decay in such a way that $v\in H^{s-\epsilon}(\mathbb{S}^1)\setminus H^s(\mathbb{S}^1)$. Of course you need to truncate the series after a finite number of terms. Observe the behaviour of v for different values of s.

3.3.2 General boundaries

Given a Lipschitz bounded domain Ω, if there is a bi-Lipschitz map $\Phi : \overline{B_1} = \{|\mathbf{x}|\le 1\} \to \overline{\Omega}$ that maps $\mathbb{S}^1$ in $\partial\Omega$, we can define the space $\boxed{H^s(\partial\Omega)}$ as the space of functions v defined on $\partial\Omega$ whose pullback $v^*(\mathbf{x}) = v(\Phi(\mathbf{x}))$ is an element of $H^s(\mathbb{S}^1)$. For $-1\le s\le 1$ this gives a well-defined space $H^s(\partial\Omega)$ independently of the choice of Φ. Different maps Φ give equivalent norms. The space $H^1(\partial\Omega)$ is precisely the space of $L^2(\partial\Omega)$ functions whose tangential first derivative $\nabla_T v$ is in $L^2(\partial\Omega)$; moreover $v\mapsto (\int_{\partial\Omega}(|v|^2+|\nabla_T v|^2)\,\mathrm{d}s)^{1/2}$ is an equivalent norm on $H^1(\partial\Omega)$.

If we do not have such a map Φ, $H^s(\partial\Omega)$ (for $-1\le s\le 1$) can be defined in a slightly more complicated way, using the cover of $\partial\Omega$ from the definition of a Lipschitz set and a so-called "partition of unity".

Several other definitions of the spaces $H^s(\partial\Omega)$ exist and can be found in the literature. The norms obtained with different definitions are equivalent (for $-1\le s\le 1$) but not necessarily equal.

In most of these notes we will use $H^s(\partial\Omega)$ only for $s=\pm 1/2$: $H^{\frac12}(\partial\Omega)\subset L^2(\partial\Omega)\subset H^{-\frac12}(\partial\Omega)$.

3.3.3 Duality product

An important property of these spaces is that $H^{-s}(\partial\Omega)$ can be identified to the anti-dual space of $H^s(\Omega)$, i.e. the space of anti-linear continuous functionals on $H^s(\partial\Omega)$. Let us look at what this means in the case of $\mathbb{S}^1$. If $v,w\in L^2(\mathbb{S}^1)$, we have seen in (18) that we can write their scalar product as the scalar product of the Fourier coefficients in the sequence space $l^2(\mathbb{Z})$ (times 2π). We want to use the same $l^2(\mathbb{Z})$ product of Fourier coefficients also when either v or w is not in $L^2(\mathbb{S}^1)$. The Fourier coefficients of this function decay slowly, thus for the series in (18) to be finite we need the other function (w or v) to have fast-converging Fourier coefficients, i.e. to be smoother. We define the **duality product**

$$\langle v, w\rangle_{\mathbb{S}^1} = \Big\langle \sum_{\ell\in\mathbb{Z}} \hat v_\ell \mathrm{e}^{\mathrm{i}\ell\theta}, \sum_{\ell\in\mathbb{Z}} \hat w_\ell \mathrm{e}^{\mathrm{i}\ell\theta} \Big\rangle_{\mathbb{S}^1} := 2\pi \sum_{\ell\in\mathbb{Z}} \hat v_\ell \overline{\hat w_\ell}$$

whenever the sum is bounded. If $v\in H^s(\mathbb{S}^1)$ and $w\in H^{-s}(\mathbb{S}^1)$ for some $s\in\mathbb{R}$ then this series is bounded: by the Cauchy–Schwarz inequality in $l^2(\mathbb{Z})$,

$$|\langle v,w\rangle_{\mathbb{S}^1}| \le 2\pi \sum_{\ell\in\mathbb{Z}} (1+\ell^2)^{s/2} |\hat v_\ell| (1+\ell^2)^{-s/2} |\hat w_\ell| \le \|v\|_{H^s(\mathbb{S}^1)} \|w\|_{H^{-s}(\mathbb{S}^1)}.$$

Moreover, if $v, w \in L^2(\mathbb{S}^1)$, then the duality product coincides with the L^2 inner product: $\langle v, w\rangle_{\mathbb{S}^1} = \int_{\mathbb{S}^1} v(\theta)\overline{w}(\theta)\,\mathrm{d}\theta$.

Similarly, it is possible to define a duality product $\langle\cdot,\cdot\rangle_{\partial\Omega}$ on $\partial\Omega$, i.e. a sesquilinear form acting on Sobolev functions defined on $\partial\Omega$ such that

$$|\langle v, w\rangle_{\partial\Omega}| \le C \|v\|_{H^{\frac12}(\partial\Omega)} \|w\|_{H^{-\frac12}(\partial\Omega)} \qquad \forall v \in H^{\frac12}(\partial\Omega),\ w \in H^{-\frac12}(\partial\Omega), \text{ and}$$
$$\langle v, w\rangle_{\partial\Omega} = \int_{\partial\Omega} v\overline{w}\,\mathrm{d}s, \qquad \text{if } w \in L^2(\partial\Omega).$$

Because of this, we sometimes abuse the notation and write the duality product simply as $\int_{\partial\Omega} v\overline{w}\,\mathrm{d}s$, even when one of the two distributions is not in L^2 and the product is not strictly speaking an integral. We write $\langle v, w\rangle_{\partial\Omega}$ also when $v \in H^{-\frac12}(\partial\Omega)$ and $w \in H^{\frac12}(\partial\Omega)$. To be more clear we might write $\langle v, w\rangle_{H^{\frac12}(\partial\Omega)\times H^{-\frac12}(\partial\Omega)}$ and $\langle v, w\rangle_{H^{-\frac12}(\partial\Omega)\times H^{\frac12}(\partial\Omega)}$, depending on the regularity of the arguments.

Exercise 3.6: (Dual norms). Let the domain Ω admit a bi-Lipschitz map $\Phi : \overline{B_1} \to \overline{\Omega}$ as above. For each v, w defined on $\partial\Omega$ let v^*, w^* be their pullbacks on $\mathbb{S}^1$ and define $\|v\|_{H^s(\Gamma)} := \|v^*\|_{H^s(\mathbb{S}^1)}$ and $\langle v, w\rangle_\Gamma := \langle v^*, w^*\rangle_{\mathbb{S}^1}$.

Show that $\|v\|_{H^s(\Gamma)} = \sup_{0\neq w\in H^{-s}(\Gamma)} \frac{|\langle v,w\rangle_\Gamma|}{\|w\|_{H^{-s}(\Gamma)}}$.

Hint: show it first for $\Gamma = \mathbb{S}^1$.

3.3.4 Trace operators

These spaces are closely related to **trace operators**. Given a smooth function $v \in C^1(\overline{\Omega})$, we define its **Dirichlet and Neumann traces**, respectively, as

$$\boxed{\gamma v := v|_{\partial\Omega}, \qquad \partial_{\mathbf{n}} v := \mathbf{n}\cdot\gamma(\nabla v)}$$

where $\mathbf{n}$ is the outward-pointing unit normal vector field on $\partial\Omega$ (which is defined almost everywhere if the domain is Lipschitz, by Rademacher theorem). Can we define these traces for more general functions, such as elements of Sobolev spaces?

Theorem 3.7: (Trace theorem). The Dirichlet trace γ can be extended to a surjective continuous operator mapping $\gamma : H^1(\Omega) \to H^{\frac12}(\partial\Omega)$. The kernel of γ is $H^1_0(\Omega)$.
The Neumann trace $\partial_{\mathbf{n}}$ can be extended to a surjective continuous operator $\partial_{\mathbf{n}} : H^1(\Omega;\Delta) \to H^{-\frac12}(\partial\Omega)$.

This theorem says that the spaces $H^{\frac12}(\partial\Omega)$ and $H^{-\frac12}(\partial\Omega)$ are precisely the spaces of the Dirichlet and the Neumann traces of $H^1(\Omega)$ and $H^1(\Omega;\Delta)$ functions, respectively. See [Spence14, p. 6] for more details.

The trace operators are local, i.e. γv and $\partial_{\mathbf{n}} v$ depend only on the value of v in an arbitrary small neighbourhood of $\partial\Omega$. Thus, if Ω is unbounded (but $\partial\Omega$ is still bounded), in the trace theorem we can substitute $H^1(\Omega)$ and $H^1(\Omega;\Delta)$ with $H^1_{\mathrm{loc}}(\overline{\Omega})$ and $H^1_{\mathrm{loc}}(\overline{\Omega};\Delta)$, respectively.

Exercise 3.8: (Equivalent norms on boundaries). Prove that the following are equivalent norms on $H^{\pm\frac12}(\partial\Omega)$:

$$\|u\|_{H^{\frac12}_{\#}(\partial\Omega)} := \inf\{\|U\|_{H^1(\Omega)} : \gamma U = u\}, \qquad \|u\|_{H^{-\frac12}_{\#}(\partial\Omega)} := \inf\{\|U\|_{H^1(\Omega;\Delta)} : \partial_{\mathbf{n}} U = u\}.$$

Hint: use that all continuous linear surjective operators between Hilbert spaces admit continuous right inverses [SBH19, Lemma 4.1].

3.4 Green's identities

The divergence theorem says that for any $\mathbf{F} \in C^1(\overline{\Omega})^2$ we have $\int_\Omega \operatorname{div}\mathbf{F}\,\mathrm{d}\mathbf{x} = \int_{\partial\Omega} \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}s$, where $\mathbf{n}$ is the outward pointing unit normal vector field on $\partial\Omega$. The product rule for the divergence is $\operatorname{div}[w\mathbf{G}] = \nabla w\cdot\mathbf{G} + w\operatorname{div}\mathbf{G}$. The combination of these two ingredients gives **Green's first and second identity** for the Helmholtz equation: for $u, w \in C^2(\overline{\Omega})$, $k \in \mathbb{R}$

$$\int_\Omega (\Delta u + k^2 u) w\,\mathrm{d}\mathbf{x} = \int_{\partial\Omega} \partial_{\mathbf{n}} u\ \gamma w\,\mathrm{d}s + \int_\Omega (k^2 uw - \nabla u\cdot\nabla w)\,\mathrm{d}\mathbf{x}, \tag{20}$$
$$\int_\Omega \Big((\Delta u + k^2 u)w - u(\Delta w + k^2 w)\Big)\,\mathrm{d}\mathbf{x} = \int_{\partial\Omega} (\partial_{\mathbf{n}} u\ \gamma w - \gamma u\ \partial_{\mathbf{n}} w)\,\mathrm{d}s. \tag{21}$$

Exercise 3.9: (Complete proof). Write in detail the proof of (20)–(21) for $u, v \in C^2(\overline{\Omega})$.

Do these identities hold for Sobolev functions?

Proposition 3.10: (Green's identities in Sobolev spaces). If Ω is a bounded Lipschitz domain, Green's first identity (20) holds for $u \in H^1(\Omega;\Delta)$ e $w \in H^1(\Omega)$. Green's second identity (21) holds for $u, v \in H^1(\Omega;\Delta)$.
The boundary integrals must be understood as the duality products $\langle \partial_{\mathbf{n}} u, \gamma\overline{w}\rangle_{\partial\Omega}$ and $\langle \gamma u, \partial_{\mathbf{n}}\overline{w}\rangle_{\partial\Omega}$.

Green's identities are the main tools in the derivation of boundary integral equations. This is the reason why the Sobolev spaces we need are $H^1(\Omega)$, $H^1(\Omega;\Delta)$ and their trace spaces $H^{\pm\frac{1}{2}}(\partial\Omega)$.

3.5 Variational problems, Fredholm theory, Gårding inequality

As we do for Laplace equation, we typically write Helmholtz problems in variational form. The abstract linear **variational problem** in the (complex) Hilbert space H is

$$\text{find} \quad u \in H \qquad \text{such that} \quad \mathcal{A}(u,w) = \mathcal{F}(w) \qquad \forall w \in H, \tag{22}$$

where $\mathcal{A}$ is a sesquilinear form in H and $\mathcal{F}$ is a anti-linear functional in H. We recall that an anti-linear functional satisfies $\mathcal{F}(\lambda v + \mu w) = \overline{\lambda}\mathcal{F}(v) + \overline{\mu}\mathcal{F}(w)$ for all $\lambda, \mu \in \mathbb{C}$ and $v, w \in H$. A sesquilinear form is linear in the first argument and anti-linear in the second one.

Given a variational problem such as (22), the (conforming) **Galerkin method** consists of choosing an N-dimensional subspace $V_N \subset H$ and a basis $\varphi_1, \dots, \varphi_N$, and of looking for a solution of the problem restricted to V_N:

$$\text{find} \quad u_N \in V_N \qquad \text{such that} \quad \mathcal{A}(u_N, w_N) = \mathcal{F}(w_N) \qquad \forall w_N \in V_N. \tag{23}$$

This is achieved computationally by solving the $N \times N$ linear algebraic system $\underline{\underline{\mathbf{A}}}\mathbf{U} = \mathbf{F}$ where $A_{j,m} := \mathcal{A}(\varphi_m, \varphi_j)$, $F_j := \mathcal{F}(\varphi_j)$ and $u_N = \sum_{j=1}^N U_j \varphi_j$.

Given a continuous sesquilinear form $\mathcal{A}$, we can associate a linear bounded operator $A : H \to H^*$ by $(Au)(w) = \langle Au, w\rangle_{H^*\times H} = \mathcal{A}(u,w)$ for all $u, w \in H$, where H^* is the anti-dual of H, [Spence14, Lemma 5.4]. The operator A admits a continuous inverse if and only if, for all $\mathcal{F} \in H^*$, the variational problem (22) is well-posed.

In the case of Laplace equation, the crucial result from functional analysis is **Lax–Milgram** theorem: if $\mathcal{A}$ is **continuous** ($|\mathcal{A}(u,w)| \le C_{\mathcal{A}} \|u\|_H \|w\|_H \ \forall u, w \in H$) and **coercive**[7] ($\Re\{\mathcal{A}(w,w)\} \ge \gamma_{\mathcal{A}} \|w\|_H^2 \ \forall w \in H$), and $\mathcal{F}$ is continuous ($|\mathcal{F}(w)| \le C_{\mathcal{F}} \|w\|_H \ \forall w \in H$), then the corresponding variational problem is well-posed. Moreover, several good properties of all Galerkin discretisations follow (well-posedness, stability bounds, quasi-optimality, bounds on the number of linear solver iterations, ...).

Unfortunately, for most variational formulations of the Helmholtz equation coercivity does not hold, so we cannot rely on Lax–Milgram. The branch of functional analysis that we need is called "Fredholm theory" and studies compact perturbations of operators. We recall the definitions of two classes of bounded linear operators between Hilbert (or Banach) spaces; see e.g. [Sayas06, p. 21].

Definition 3.11: (Compact and Fredholm operators). A linear operator $K : H_1 \to H_2$ is **compact** if the image of a bounded sequence admits a converging subsequence ($(v_j)_{j\in\mathbb{N}} \subset H_1, \|v_j\|_H \le C \ \forall j \in \mathbb{N} \Rightarrow \exists j_m \to \infty, w \in H_2$ such that $Kv_{j_m} \to w$).
A bounded linear operator is a **Fredholm operator** if it is sum of an invertible one and a compact one (more precisely, we should say it is a Fredholm operator of index 0).

[8] We can think at Fredholm operators as "small" perturbations of invertible operators. The main result is the "Fredholm alternative", which, in its simplest form, reads as follows; [Brezis11, Thm. 6.6(d)], [SBH19, §8.1, 8.6].

[7] Here terminology can be confusing. Coercivity is sometimes called "sign-definiteness", "V-ellipticity", "strong ellipticity", or "strict coercivity", see [Spence14, §5.2]. In some of these cases, the word "coercive" is used for sesquilinear forms satisfying a Gårding inequality, which is a weaker condition. Here we follow the convention of [Spence14].

[8] Well, this is not exactly the definition you find in functional analysis textbooks. E.g. [Brezis11, p. 168] defines $T : H_1 \to H_2$ (continuous operator between Hilbert spaces, for simplicity) as "Fredholm with index $\mathrm{Ind}(T)$" if $\dim(\ker T), \dim(\mathrm{Im}T)^\perp < \infty$ and $\mathrm{Ind}(T) := \dim(\ker T) - \dim(\mathrm{Im}T)^\perp$. Let us check that this definition, with index 0, is equivalent to ours.

An invertible operator is Fredholm with index 0; then by [Brezis11, p. 169, ⋆1(c)] any operator in the form invertible+compact as in Definition 3.11 is Fredholm with index 0.

Conversely, assume that $T : H_1 = (\ker T \oplus (\ker T)^\perp) \to H_2 = (\mathrm{Im}T \oplus (\mathrm{Im}T)^\perp)$ is Fredholm with index 0. Let $\{\phi_1, \dots, \phi_n\}$ be an orthonormal basis of $\ker T$ and $\{\psi_1, \dots, \psi_n\}$ be an orthonormal basis of $(\mathrm{Im}T)^\perp$ (which are both finite-dimensional and have the same dimension). Define $K : H_1 \to H_2$ as $K\phi_j = \psi_j$ and $Kv = 0 \ \forall v \in (\ker T)^\perp$; this operator has finite rank, so it is compact. Then $L = T + K$ is an invertible operator and $T = L - K$ is in the form invertible+compact, as desired.

To verify that an operator is Fredholm we will always use Definition 3.11.

Theorem 3.12: (Fredholm alternative). Let $T : H_1 \to H_2$ be a Fredholm operator.
T is injective if and only if it is surjective. In this case its inverse T^{-1} is bounded.

This theorem has this name because it states that when we have a Fredholm operator T then only two "alternatives" are possible: either T is injective and surjective, or is neither injective nor surjective. Fredholm alternative says that if we want to prove that a Fredholm operator is invertible, then it suffices to prove its injectivity. A useful idea to keep in mind is that, under this respect, Fredholm operators behave like square matrices.[9]

How do we use Fredholm alternative? The general strategy is the following. We will show that some linear operator T mapping "problem solution" to "problem data" are Fredholm. When we can show that the homogeneous problem (with data equal to 0) only admits the trivial solution (i.e. T is injective), Fredholm alternative guarantees that *all* data admit a solution (i.e. T is surjective), which is unique by injectivity, and that the solution is controlled by the data (i.e. T^{-1} is bounded). **A linear problem whose "solution-to-data" operator is Fredholm is well-posed if the same operator is injective.**

Typically, in time-harmonic problems, sesquilinear forms are not coercive but satisfy a weaker inequality, called Gårding inequality.

Definition 3.13: (Gårding inequality). Let $H \subset V$ be two Hilbert spaces provided with the norms $\|\cdot\|_H$ and $\|\cdot\|_V$, and let the embedding $H \hookrightarrow V$ be continuous. A sesquilinear form satisfies a Gårding inequality if there exists two positive constants α and C_V such that

$$\Re\{\mathcal{A}(v,v)\} \geq \alpha \|v\|_H^2 - C_V \|v\|_V^2 \qquad \forall v \in H. \tag{24}$$

[10] Here we follow the notation of [Spence14, §5.3], where the letters H and V are swapped with respect to the classical choice for Hilbert triples (as in [Brezis11, p. 136]).

The main use of Gårding inequality comes from the next result, see [Spence14, Theorem 5.20].

Proposition 3.14: (Gårding ⇒ Fredholm). Assume that H and V are as in Definition 3.13, the embedding $H \hookrightarrow V$ is compact and the continuous sesquilinear form $\mathcal{A}(\cdot,\cdot)$ satisfies the Gårding inequality (24).
Then the operator $A : H \to H^*$ associated to $\mathcal{A}(\cdot,\cdot)$ is Fredholm.

Proof. Let $i : H \to V$ be the (compact) inclusion map, and $T : V \to H^*$ defined by $\langle Tv, w\rangle_{H^*\times H} = (v, iw)_V$, for $v \in V$ and $w \in H$, where $(\cdot,\cdot)_V$ is the scalar product in V. Then, by Cauchy–Schwarz in V, T is continuous: $\|Tv\|_{H^*} = \sup_{w\in H} \frac{|\langle Tv,w\rangle_{H^*\times H}|}{\|w\|_H} \leq \sup_{w\in H} \frac{\|v\|_V \|iw\|_V}{\|w\|_H} \leq \|v\|_V \|i\|_{H\to V}$. Define $B := A + C_V T \circ i : H \to H^*$, where $C_V > 0$ is the value in (24). Since $T \circ i$ is compact, in order to prove that A is Fredholm, it is enough to see that B is invertible. Then the sesquilinear form

$$\mathcal{B}(u,w) := \langle Bu, w\rangle_{H^*\times H} = \langle Au, w\rangle_{H^*\times H} + C_V \langle T \circ iu, w\rangle_{H^*\times H} = \mathcal{A}(u,w) + C_V (iu, iw)_V, \qquad u, w \in H,$$

is continuous and coercive in H, which, by Lax–Milgram, implies that B is invertible. (Actually we have proved something stronger, that A is sum of a compact operator and a *coercive* one.) □

Proposition 3.14 and Theorem 3.12 imply that, in order to prove well-posedness of a variational problem (22) whose sesquilinear form $\mathcal{A}(\cdot,\cdot)$ satisfies a Gårding inequality, it suffices to study the homogeneous problem (with $\mathcal{F} = 0$):

Corollary 3.15: (Well-posedness from Gårding). Assume that:

- $H \subset V$ are Hilbert spaces and the embedding $H \hookrightarrow V$ is compact,
- the sesquilinear form $\mathcal{A}(\cdot,\cdot)$ is continuous in H and satisfies the Gårding inequality (24),
- the homogeneous variational problem, $\mathcal{A}(u_0, w) = 0$ for all $w \in H$, admits only the trivial solution $u_0 = 0$.

Then also the inhomogeneous problem (22) is well-posed, for any $\mathcal{F} \in H^*$.

[9] Indeed, an invertible linear operator between finite-dimensional spaces corresponds to a square matrix. All finite-range operators are compact because all bounded sequences of $\mathbb{R}^n$ and $\mathbb{C}^n$ admit converging subsequences. Thus the operators between finite-dimensional spaces that are Fredholm are precisely those associated to square matrices. We know from linear algebra that a square matrix is injective if and only if it is surjective. This proves the Fredholm alternative in the case of operators between finite-dimensional spaces.

(In this $\mathbb{C}^n \to \mathbb{C}^n$ case, Fredholm operators are compact, but in infinite dimensions a Fredholm operator is not compact.)

[10] We could take a more general definition using the inequality $\Re\{\sigma\mathcal{A}(v,v)\} \geq \alpha \|v\|_H^2 - C_V \|v\|_V^2$ for some $0 \neq \sigma \in \mathbb{C}$ (the same σ for all v). Then all consequences of the Gårding inequality would follow precisely in the same way. However all variational problems with $\mathcal{A}(\cdot,\cdot)$ that satisfy this inequality can be reduced to equivalent problems satisfying (24) simply by multiplying the sesquilinear form and the linear functional by σ.

To be able to exploit Corollary 3.15, we need to know when the embedding between two function spaces is compact. A classical result, called Rellich embedding theorem, says that $H^1(\Omega) \hookrightarrow L^2(\Omega)$ is compact, when Ω is a bounded Lipschitz domain[11]. Similarly, one can show that $H^{s_+}(\partial\Omega) \hookrightarrow H^{s_-}(\partial\Omega)$ is compact for all $s_+ > s_-$, [Néd01, Theorem 2.5.7].

Exercise 3.16: (Compactness of Sobolev embeddings). Let $s_- < s_+$ be real numbers. Show the compactness of Sobolev embedding $i : H^{s_+}(\mathbb{S}^1) \to H^{s_-}(\mathbb{S}^1)$, in the simple case of the boundary of the unit circle.

The key tool to use is [Brezis11, Cor. 6.2]: given a linear operator $T : H_1 \to H_2$ and $T_j : H_1 \to H_2$ finite-rank operators (i.e. with finite-dimensional image) for $j \in \mathbb{N}$, if $\|T - T_j\|_{H_1 \to H_2} \to 0$ then T is compact.

Construct a sequence of finite-rank "truncated embedding" operators $i_L : H^{s_+}(\mathbb{S}^1) \to H^{s_-}(\mathbb{S}^1)$ for $L \in \mathbb{N}$, such that

$$\|i - i_L\|^2_{H^{s_+}(\mathbb{S}^1) \to H^{s_-}(\mathbb{S}^1)} \leq \frac{1}{(1+L^2)^{s_+ - s_-}}.$$

Exercise 3.17: (Compact and Fredholm operators in sequence spaces). Let l^2 be the Hilbert space of squared-summable complex sequences $x = (x_j)_{j\in\mathbb{N}}$, $x_j \in \mathbb{C}$, equipped with $\|x\|^2_{l^2} = \sum_{j\in\mathbb{N}} |x_j|^2$.

- Define the right and left complex shift operators $R, L : l^2 \to l^2$ as $(Rx)_1 = 0$, $(Rx)_{j+1} = x_j$ and $(Lx)_j = x_{j+1}$ for $j \in \mathbb{N}$. Show that R and L are neither Fredholm nor compact.

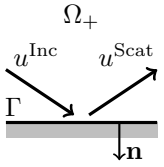
  Are the composition LR and RL either Fredholm or compact? Which of L, R, LR, RL is invertible?
- Show that for any $J \in \mathbb{N}$ the truncation operator $T_J : l^2 \to l^2$ defined by $(T_J x)_j = x_j$ if $j \leq J$ and $(T_J x)_j = 0$ if $j > J$ is compact.
- Show that the operator $T : l^2 \to l^2$ defined by $Tx = (x_3, 0, -x_3, 2x_4, 2x_5, \dots, (Tx)_j = 2x_j, \dots)$ is Fredholm.

An extensive description of the relations between variational problems satisfying Lax–Milgram assumptions, Gårding inequality or the inf-sup inequality, and the consequences for the Galerkin method, can be found in [Spence14, §5.3].

4 Boundary value problems for the Helmholtz equation

4.1 Plane waves reflected by a straight line

To understand what happens when a wave hits an impenetrable obstacle we start with a very simple case that can be solved analytically. Let $u^{\mathrm{Inc}}(\mathbf{x}) = \mathrm{e}^{\mathrm{i}k\mathbf{x}\cdot\mathbf{d}}$ be a plane wave with $|\mathbf{d}| = 1$, $d_1 \geq 0$ and $d_2 \leq 0$ (i.e. propagating rightward and downward in the plane, $\searrow$). This is called the "**incoming field**", or "**incident field**". Assume that we truncate the domain of propagation to the upper half plane $\Omega^+ = \{x_2 > 0\}$ and on the horizontal line $\Gamma = \{x_2 = 0\}$ we impose some boundary conditions that reflect the impinging wave. We call u^{Scat} the reflected wave, i.e. the "**scattered field**", and $u^{\mathrm{Tot}} = u^{\mathrm{Inc}} + u^{\mathrm{Scat}}$ the "**total field**". u^{Tot} is the physical field we would measure in a point of the half plane.



Given u^{Inc}, which is a datum, we now want to find u^{Tot} that satisfies the Helmholtz equation in the upper half plane, and satisfies some desired homogeneous boundary conditions on Γ. This is the same as saying that we want u^{Scat} that satisfies the Helmholtz equation in the same region, and satisfies boundary conditions that depend on u^{Inc} on Γ.

By the law of reflection, we expect u^{Scat} to propagate upwards and to make with the horizontal line Γ the same angle as u^{Inc}, $\nearrow$. This means that u^{Scat} is a plane wave with direction $\widetilde{\mathbf{d}} = (d_1, -d_2)$: $u^{\mathrm{Scat}}(\mathbf{x}) = A\mathrm{e}^{\mathrm{i}k\widetilde{\mathbf{d}}\cdot\mathbf{x}} = A\mathrm{e}^{\mathrm{i}k(x_1 d_1 - x_2 d_2)}$ for some reflection coefficient $A \in \mathbb{C}$ that gives the amplitude and the phase of u^{Scat} itself. The coefficient A depends on the particular type of boundary condition chosen.

- When the line Γ is sound-soft, the Dirichlet trace of the total field u^{Tot} vanishes on Γ:

$$0 = u^{\mathrm{Tot}}(x_1, 0) = u^{\mathrm{Inc}}(x_1, 0) + u^{\mathrm{Scat}}(x_1, 0) = \mathrm{e}^{\mathrm{i}kx_1 d_1} + A\mathrm{e}^{\mathrm{i}kx_1 d_1} \qquad \forall x_1 \in \mathbb{R} \qquad \Rightarrow\ A = -1.$$

- When the line Γ is sound-hard, the Neumann trace of the total field u^{Tot}, i.e. its normal derivative, vanishes on Γ:

$$0 = \frac{\partial u^{\mathrm{Tot}}}{\partial x_2}(x_1, 0) = \frac{\partial}{\partial x_2}\Big(u^{\mathrm{Inc}}(x_1, 0) + u^{\mathrm{Scat}}(x_1, 0)\Big) = \mathrm{i}kd_2\mathrm{e}^{\mathrm{i}kx_1 d_1} - A\mathrm{i}kd_2\mathrm{e}^{\mathrm{i}kx_1 d_1} \quad \forall x_1 \in \mathbb{R} \quad \Rightarrow\ A = 1.$$

[11] See e.g. [Brezis11, Thm. 9.16] and [SBH19, Prop. 7.5] for a more general version.

- To impose the impedance boundary condition $\frac{\partial u^{\mathrm{Tot}}}{\partial \mathbf{n}} - \mathrm{i}k\vartheta u^{\mathrm{Tot}} = 0$ with $\vartheta > 0$, we recall that $\mathbf{n}$ is the outward-pointing unit normal vector on Γ so $\mathbf{n} = (0, -1)$. In this case we have

$$0 = \frac{\partial u^{\mathrm{Tot}}}{\partial \mathbf{n}} - \mathrm{i}k\vartheta u^{\mathrm{Tot}} = -\frac{\partial u^{\mathrm{Tot}}}{\partial x_2} - \mathrm{i}k\vartheta u^{\mathrm{Tot}} = \Big(\mathrm{i}kd_2(-1+A) - \mathrm{i}k\vartheta(1+A)\Big)\mathrm{e}^{\mathrm{i}kx_1d_1}.$$

  This vanishes for $d_2(-1+A) = \vartheta(1+A)$, i.e. $A = \frac{d_2+\vartheta}{d_2-\vartheta}$.

Summarising, the fields are

$$\begin{aligned} u^{\mathrm{Inc}}(\mathbf{x}) &= \mathrm{e}^{\mathrm{i}k(x_1d_1+x_2d_2)}, \\ u^{\mathrm{Scat}}(\mathbf{x}) &= A\mathrm{e}^{\mathrm{i}k(x_1d_1-x_2d_2)}, \\ u^{\mathrm{Tot}}(\mathbf{x}) &= \mathrm{e}^{\mathrm{i}k(x_1d_1+x_2d_2)} + A\mathrm{e}^{\mathrm{i}k(x_1d_1-x_2d_2)}, \end{aligned} \qquad \text{with } A = \begin{cases} -1 & \text{for sound-soft } \Gamma, \\ 1 & \text{for sound-hard } \Gamma, \\ \frac{d_2+\vartheta}{d_2-\vartheta} \in (-1,1) & \text{for impedance } \Gamma. \end{cases}$$

The plane waves reflected by sound-soft or sound-hard interfaces have the same amplitude of the incoming plane waves and either opposite (sound-soft) or the same (sound-hard) phase. On the other hand, since $|A| < 1$, the waves reflected by an impedance line have amplitude smaller than the incoming wave: the impedance boundary absorbs some of the wave energy. The amount of wave that is reflected and the amount that is absorbed depend on the direction of the incoming wave; in particular, if $\vartheta \leq 1$ the impedance boundary does not reflect (but absorbs completely) the impinging waves propagating in direction $\mathbf{d} = (\sqrt{1-\vartheta^2}, -\vartheta)$. An impedance boundary with $\vartheta = 1$ does not reflect the waves hitting perpendicularly.

For $\vartheta \to 0$ the impedance boundary condition converges to the sound-hard one, and consistently $A \to 1$; for $\vartheta \to \infty$ it converges to the sound-soft boundary condition and $A \to -1$.

Here we have considered a wave with infinite front hitting an infinite obstacle: clearly this is not a very realistic problem, but it helps to get an intuition of what happens when a wave hits an impenetrable obstacle.

Remark 4.1: (On unbounded domains: PDE + BCs $\not\Rightarrow$ BVP). When we solve a well-posed boundary value problem the solution is typically determined by the PDE and by the boundary conditions. Here we have used something more. Let us look for example at the sound-soft case. Given u^{Inc}, *any* combination $u^{\mathrm{Scat}}_\lambda(\mathbf{x}) = -\lambda\mathrm{e}^{\mathrm{i}k(x_1d_1+x_2d_2)} - (1-\lambda)\mathrm{e}^{\mathrm{i}k(x_1d_1-x_2d_2)}$ satisfies the Helmholtz equation and the boundary condition $u^{\mathrm{Scat}}_\lambda = -u^{\mathrm{Inc}}$ on Γ. We have chosen the case $\lambda = 0$ because from the law of reflection we expect the scattered field to propagate upwards. We will see that in all problems posed on unbounded domains we need to select the waves propagating in the desired direction, and that this is equivalent to imposing conditions "at infinity".

Remark 4.2: (General wave reflected by a straight line). The same reasoning made above shows that if u^{Inc} is any wave propagating downwards, then the reflected wave $u^{\mathrm{Scat}}_D(\mathbf{x}) = -u^{\mathrm{Inc}}(x_1, -x_2)$ propagates upwards and $u^{\mathrm{Inc}} + u^{\mathrm{Scat}}_D$ vanishes on Γ. So $u^{\mathrm{Tot}}_D = u^{\mathrm{Inc}} + u^{\mathrm{Scat}}_D$ is the total field in the presence of a sound-soft line.

You can see the reflection of the Herglotz function $u^{\mathrm{Inc}}(\mathbf{x}) = \int_{-\pi/3}^{-\pi/6} \mathrm{e}^{\mathrm{i}k(x_1\cos\varphi + x_2\sin\varphi)}\,\mathrm{d}\varphi$ by a sound-soft horizontal line in Figure 7 and in the animation on the course webpage.

Similarly, $u^{\mathrm{Scat}}_N(\mathbf{x}) = u^{\mathrm{Inc}}(x_1, -x_2)$ propagates upwards and $\frac{\partial}{\partial \mathbf{n}}(u^{\mathrm{Inc}} + u^{\mathrm{Scat}}_N)$ vanishes on Γ, so $u^{\mathrm{Tot}}_N = u^{\mathrm{Inc}} + u^{\mathrm{Scat}}_N$ is the reflection of any downward u^{Inc} by a sound-hard line.

For an impedance line Γ, the argument is slightly more complicated: since the reflection coefficient A depends on the direction of the incoming wave, to compute u^{Scat} we need to be able to decompose u^{Tot} in plane waves with different directions and reflect each one of them with its own coefficient. This is possible if u^{Inc} is a Herglotz function with kernel supported in the lower half of the unit circle ($g(\varphi) = 0$ for $0 < \varphi < \pi$).

Exercise 4.3: (Neumann reflection of Herglotz function). Plot the total field when the incoming field is the same Herglotz function as in Figure 7 and in Remark 4.2, but the horizontal line Γ acts as a Neumann boundary. Start from the Matlab file provided.

Repeat the same for an impedance boundary.

Exercise 4.4: (Neumann traces on sound-soft boundaries and vice versa). For the problem of a plane wave u^{Inc} impinging on $\Gamma = \{x_2 = 0\}$ as described above, show that, $\forall x \in \mathbb{R}$,

- if Γ is a sound-soft boundary then $\partial_{\mathbf{n}} u^{\mathrm{Scat}}(x, 0) = \partial_{\mathbf{n}} u^{\mathrm{Inc}}(x, 0)$ and $\partial_{\mathbf{n}} u^{\mathrm{Tot}}(x, 0) = 2\partial_{\mathbf{n}} u^{\mathrm{Inc}}(x, 0)$;
- if Γ is a sound-hard boundary then $u^{\mathrm{Scat}}(x, 0) = u^{\mathrm{Inc}}(x, 0)$ and $u^{\mathrm{Tot}}(x, 0) = 2u^{\mathrm{Inc}}(x, 0)$;
- if Γ is an impedance boundary then $(\partial_{\mathbf{n}} - \mathrm{i}k\widetilde{\vartheta})u^{\mathrm{Scat}} = (\partial_{\mathbf{n}} - \mathrm{i}k\widetilde{\vartheta})u^{\mathrm{Inc}}$ for $\widetilde{\vartheta} = \frac{(d_2)^2}{\vartheta}$.

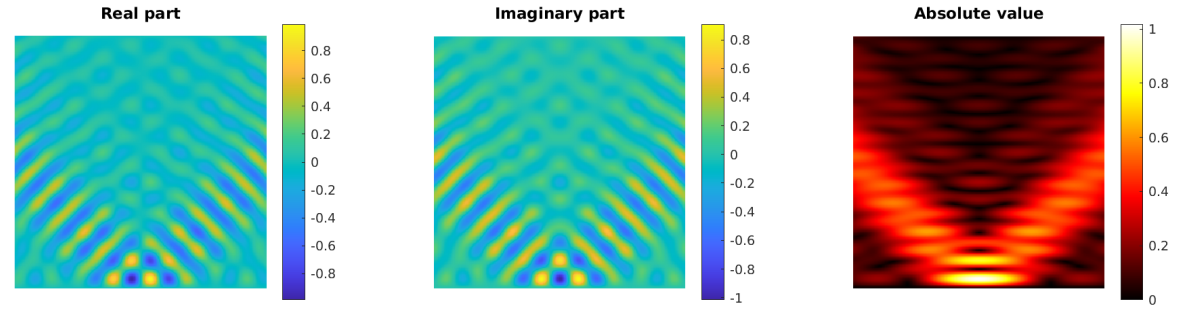


Figure 7: Reflection of the Herglotz function with kernel $g(\varphi) = \chi_{(-\frac{\pi}{3},-\frac{\pi}{6})}(\varphi)$ by the sound-soft line $\{x_2 = 0\}$. See Remark 4.2.

Exercise 4.5: (Reflection of vector plane waves). Consider a vector plane wave $\mathbf{A}\mathrm{e}^{\mathrm{i}k\mathbf{d}\cdot\mathbf{x}}$ that is solution of Maxwell's equations (10) as in Exercise 2.3. Compute the plane wave reflected by the horizontal plane $\{\mathbf{x} \in \mathbb{R}^3, x_3 = 0\}$ equipped with PEC boundary conditions $\mathbf{E} \times \mathbf{n} = \mathbf{0}$. Recall that the amplitude of the reflected wave must be orthogonal to its propagation direction.

The reflection of elastic waves is more complicated: an impinging pressure (or shear) wave generates a reflected waves that is sum of a pressure and a shear wave. This phenomenon is called "mode conversion" and is due to the boundary conditions, which involve all three Cartesian components of the fields.

4.2 Boundary value problems in bounded domains

Let $\Omega \subset \mathbb{R}^2$ be a bounded, open, Lipschitz set. The Dirichlet BVP for the Helmholtz equation is: given a source term f defined in Ω and a boundary datum g_D defined on $\partial\Omega$ find u on Ω such that

$$\Delta u + k^2 u = -f \quad \text{in } \Omega, \qquad \gamma u = g_D \quad \text{on } \partial\Omega. \tag{25}$$

We know that the Dirichlet problem for the Poisson equation (problem (25) with $k = 0$) is well-posed when the data f, g_D are sufficiently smooth. Despite the Helmholtz equation looks like an innocuous low-order perturbation of the Laplace equation, well-posedness of the Helmholtz Dirichlet problem is not guaranteed.

We start from a simple example. Let Ω be the rectangle $(0, L_1) \times (0, L_2)$. Then for all $j_1, j_2 \in \mathbb{N}$ the field $u_{j_1,j_2}(\mathbf{x}) = \sin(\frac{j_1\pi}{L_1}x_1)\sin(\frac{j_2\pi}{L_2}x_2)$ vanishes on $\partial\Omega$ and is solution of $\Delta u + k_{j_1,j_2}^2 u = 0$ with $k_{j_1,j_2}^2 = \frac{j_1^2\pi^2}{L_1^2} + \frac{j_2^2\pi^2}{L_2^2}$. So there are infinitely many values of k such that the homogeneous ($f = 0$ and $g_D = 0$) Helmholtz Dirichlet BVP admits non-trivial solutions. It follows that for these values of k the problem (25) is not well-posed: if there is a solution then it cannot be unique.

Solutions of the homogeneous Helmholtz Dirichlet problem are called **Dirichlet eigenfunctions** of the Laplacian with eigenvalue $\Lambda = k^2$, as they satisfy the eigenproblem $-\Delta u = \Lambda u$, $\gamma u = 0$.

As a second example, if $\Omega = B_R$ is the disc of radius R, then the circular waves $J_\ell(kr)\mathrm{e}^{\pm\mathrm{i}\ell\theta}$ are Dirichlet eigenfunctions for $\Lambda = k^2$ and k such that kR is a zero of the ℓth Bessel function J_ℓ (recall the plots in the left panel of Figure 4 and the top panel of Figure 5: each J_ℓ has infinitely many zeros).

In other domains we find the same situation as in the two examples described, even if we cannot compute eigenvalues and eigenfunctions explicitly; see [SBH19, §9] for the spectral theory of elliptic operators.

To understand the problem in more general bounded Lipschitz domains, we study the problem from a variational point of view. From Green's first identity (20), the variational problem for the Helmholtz Dirichlet BVP (25) with homogeneous boundary conditions $g_D = 0$ is

$$\text{find } u \in H_0^1(\Omega) \text{ such that} \quad \mathcal{A}(u, w) := \int_\Omega (\nabla u \cdot \nabla \overline{w} - k^2 u\overline{w})\,\mathrm{d}\mathbf{x} = \int_\Omega f\overline{w}\,\mathrm{d}\mathbf{x} =: \mathcal{F}(w) \quad \forall w \in H_0^1(\Omega). \tag{26}$$

The sesquilinear form $\mathcal{A}(\cdot,\cdot)$ and the linear functional $\mathcal{F}(\cdot)$ are continuous in $H_0^1(\Omega)$. On the other hand, $\mathcal{A}(\cdot,\cdot)$ is not coercive (for k sufficiently large), as the two terms ($\nabla u\nabla\overline{w}$ and $-k^2 u\overline{w}$) have opposite signs,

see Exercise 4.8 or [Spence14, Lemma 6.2]. However, it satisfies a Gårding inequality (24) with $\alpha = 1$ and $C_V = k^2 + 1$:

$$\Re\{\mathcal{A}(w,w)\} = \mathcal{A}(w,w) = \|w\|^2_{H^1(\Omega)} - (k^2+1)\,\|w\|^2_{L^2(\Omega)} \qquad \forall w \in H^1(\Omega).$$

Proposition 3.14, together with the compactness of $H^1_0(\Omega)$ in $L^2(\Omega)$, gives that the operator $A : H^1_0(\Omega) \to (H^1_0(\Omega))^*$, $\boxed{A : u \mapsto f, \text{ is Fredholm}}$.[12] Corollary 3.15 of Fredholm alternative then implies that, given Ω and k, only two situations can happen:

- If Ω and k are such that the homogeneous ($f = 0$) problem (26) admits only the trivial solution $u = 0$, then also problem (26) with any $f \in L^2(\Omega)$ (or more generally $\mathcal{F} \in (H^1_0(\Omega))^*$) admits a unique solution.
- On the other hand, if there is a $u \neq 0$ such that $\mathcal{A}(u,w) = 0\ \forall w \in H^1_0(\Omega)$, then the problem (26) is not well-posed for *any* f. It might have no solutions for some f, and many solutions for some other f.

We have proved part of the following proposition. To prove the remaining part (the existence, discreteness and divergence at infinity of the eigenvalues), one needs the spectral theory of self-adjoint compact operators, see e.g. [Brezis11, §6]. To treat inhomogeneous Dirichlet boundary conditions $g_D \neq 0$, one uses a "lifting", i.e. a $u_D \in H^1(\Omega)$ such that $\gamma u_D = g_D$ (which exists because of the surjectivity of the trace operator), and then solves for $u_0 = u - u_D \in H^1_0(\Omega)$.

Proposition 4.6: (Well-posedness of the Helmholtz–Dirichlet BVP). For a Lipschitz bounded domain Ω, there exist a sequence of positive numbers $k_1 < k_2 < \dots$, with $k_j \to \infty$, such that:

- If $k = k_j$ for some j, then the Dirichlet problem (25) is not well-posed.
  In particular, the homogeneous case with $f = 0$ and $g_D = 0$ admits non-trivial solutions.
- If $0 < k \neq k_j$ for all j, then the Dirichlet problem (25) is well-posed in $H^1_0(\Omega)$ for all $f \in L^2(\Omega)$ and $g_D \in H^{\frac{1}{2}}(\partial\Omega)$.

The smallest of the values k_j equals the inverse of the Poincaré constant of Ω: the smallest value C_P such that $\|u\|_{L^2(\Omega)} \leq C_P \|\nabla u\|_{L^2(\Omega)}$ for all $u \in H^1_0(\Omega)$ is $C_P = 1/k_1$ (this is an easy consequence of the existence of an orthogonal basis of $H^1_0(\Omega)$ made of eigenfunctions).

Exercise 4.7: (Helmholtz–Neumann BVP).

- What are the eigenvalues and the eigenfunctions for the Laplacian with Neumann boundary conditions $\partial_{\mathbf{n}} u = 0$ on the rectangle and on the disc?
- Show that the positive Neumann eigenvalues for a rectangle coincide with the Dirichlet eigenvalues but the eigenfunctions differ. (On the contrary, using subtle properties of the Bessel functions it is possible to see that there are no Neumann eigenvalues of the disc that are also Dirichlet eigenvalues.)
- Write the variational formulation of the Helmholtz–Neumann BVP with inhomogeneous conditions $\partial_{\mathbf{n}} u = g_N \in H^{-\frac{1}{2}}(\partial\Omega)$: the sesquilinear form coincides with that in (26) but the linear functional and the function space differ.

Exercise 4.8: (No coercivity for $k > k_1$). Let u_1, u_2 be Dirichlet eigenfunctions in Ω, associated to different eigenvalues k_1^2 and k_2^2 and normalised as $\|u_1\|_{L^2(\Omega)} = \|u_2\|_{L^2(\Omega)} = 1$. Let $k_1 < k < k_2$, and define $w = u_1 \pm \sqrt{\frac{k^2-k_1^2}{k_2^2-k^2}}\, u_2$. Show that $\mathcal{A}(w,w) = 0$.

Deduce that $\mathcal{A}(\cdot,\cdot)$ is not coercive for all k^2 larger than the first Dirichlet eigenvalue.

Hint: recall (or prove) that eigenfunctions associated to different eigenvalues are orthogonal both in $L^2(\Omega)$ and in $H^1(\Omega)$ norms.

Exercise 4.9: (Absorption gives well-posedness.). Show that the Dirichlet and the Neumann problems

$$\Delta u + k^2 u = -f \quad \text{in } \Omega, \qquad \gamma u = 0 \quad \text{on } \partial\Omega, \quad \text{or}$$

[12] If A is Fredholm, what are the invertible and the compact operators that sum to $A : H^1_0(\Omega) \to (H^1_0(\Omega))^*$?

We can split the sesquilinear form as $\mathcal{A}(u,w) = \mathcal{A}_0(u,w) + \mathcal{K}(u,w)$ with $\mathcal{A}_0(u,w) := \int_\Omega \nabla u \cdot \nabla \overline{w}\,\mathrm{d}\mathbf{x}$ and $\mathcal{K}(u,w) := (-k^2)\int_\Omega u\overline{w}\,\mathrm{d}\mathbf{x}$. Then $\mathcal{A}_0(\cdot,\cdot)$ is the sesquilinear form associated to the (well-posed) Laplace–Dirichlet BVP in $H^1_0(\Omega)$, so it is continuous and coercive (by Poincaré inequality), thus the operator A_0 associated, i.e. $\langle A_0 u, w\rangle_{(H^1_0(\Omega))^* \times H^1_0(\Omega)} = \mathcal{A}_0(u,w)$, is invertible. This operator is simply the Laplacian $A_0 = -\Delta : H^1_0(\Omega) \to (H^1_0(\Omega))^*$. The operator $K : H^1_0(\Omega) \to (H^1_0(\Omega))^*$ associated to the second sesquilinear form, i.e. $\langle Ku, w\rangle_{(H^1_0(\Omega))^* \times H^1_0(\Omega)} = \mathcal{K}(u,w)$, is $-k^2$ times the embedding of $H^1_0(\Omega)$ in $(H^1_0(\Omega))^*$, which is a compact operator (the dual is a space larger than $L^2(\Omega)$).

The trick, hidden in Proposition 3.14, to decompose the Helmholtz "solution-to-data" operator A is to write the analogous operator for the Laplace equation (A_0, invertible) and to verify that the remainder ($K = A - A_0$) is compact as it comes from the lower-order term in the PDE. We will use again the "Helmholtz = Laplace + compact low-order perturbation" trick.

$$\Delta u + k^2 u = -f \quad \text{in } \Omega, \qquad \partial_{\mathbf{n}} u = g_N \quad \text{on } \partial\Omega,$$

are well-posed if $\Im k > 0$, $f \in L^2(\Omega)$ and $g_N \in H^{-\frac{1}{2}}(\partial\Omega)$.

Hint: First write the two BVPs as variational problems $\mathcal{A}(u,w) = \mathcal{F}(w)$ in $H^1_0(\Omega)$ and $H^1(\Omega)$, respectively. Then use Lax–Milgram theorem. To prove the coercivity $|\mathcal{A}(w,w)| \geq c\|w\|^2_{H^1(\Omega)}$ of the sesquilinear form obtained, first control the L^2 norm of u, then the H^1 seminorm by using the triangle inequality.

Deduce a bound on $\|u\|_{H^1(\Omega)}$. The bounding constant C will blow up for $\Im k \searrow 0$.

If instead of sound-soft and sound-hard conditions we have impedance ones we obtain a different result. Consider the **impedance BVP**:

$$\Delta u + k^2 u = -f \quad \text{in } \Omega, \qquad \partial_{\mathbf{n}} u - \mathrm{i}k\vartheta\, \gamma u = g_I \quad \text{on } \partial\Omega, \tag{27}$$

for $\vartheta > 0$, $f \in L^2(\Omega)$, $g_I \in H^{-\frac{1}{2}}(\partial\Omega)$. Its variational form is: find $u \in H^1(\Omega)$ such that

$$\mathcal{A}_I(u,w) := \int_\Omega (\nabla u \cdot \nabla \overline{w} - k^2 u \overline{w})\,\mathrm{d}\mathbf{x} - \mathrm{i}k\vartheta \int_{\partial\Omega} \gamma u\, \gamma\overline{w}\,\mathrm{d}s = \int_\Omega f\overline{w}\,\mathrm{d}\mathbf{x} + \int_{\partial\Omega} g_I\, \gamma\overline{w}\,\mathrm{d}s =: \mathcal{F}_I(w) \quad \forall w \in H^1(\Omega). \tag{28}$$

As before, the sesquilinear form is continuous, coercive only for small k, and satisfies a Gårding inequality. So to check the well-posedness we only have to look at the homogeneous problem. If u_0 satisfies (28) with $\mathcal{F}_I = 0$ (i.e. $f = 0$ and $g_I = 0$), taking the imaginary part of $\mathcal{A}_I(u_0, u_0) = 0$, we see that $\gamma u_0 = 0$ on $\partial\Omega$, and from the boundary condition that also $\partial_{\mathbf{n}} u_0 = 0$ on $\partial\Omega$. We will see in Corollary 5.13 that this implies that the impedance BVP is always well-posed. A different proof can be found in [SBH19, §8.8].

Remark 4.10: (What kind of waves are the eigenfunctions?). We have observed in §4.1 that Dirichlet and Neumann boundary conditions reflect waves without losing energy. Roughly speaking, we can interpret Dirichlet and Neumann eigenfunctions as waves that bounce around in Ω forever, without any damping. At the right wavenumber the interference of the wave with itself is constructive (after a full round of bounces the wave has precisely the same phase it started with), so, in a sense, it can exist without a source; see Exercises 4.11–4.12. For example, if Ω is a disc, one can imagine a wave propagating along a regular polygon inscribed in Ω, reflected by $\partial\Omega$ at every corner of the polygon. These are called **creeping waves**, as they "crawl" around $\partial\Omega$ and are small in the centre of Ω, or **whispering gallery modes** (from some circular buildings where a whisper can be heard in any place close to the wall but not in the centre); see Figure 8.

On the other hand, impedance boundary conditions and complex wavenumbers entail some energy absorption: in this case the waves cannot propagate forever and there are no eigenfunctions for any $k > 0$.

For high frequencies $k \nearrow \infty$ Helmholtz solutions resemble more and more trajectories of particles, or billiard balls on a table, or light rays if the particles are photons. The study of the relationships between the dynamics of "billiard trajectories" and the properties of PDEs with a vanishingly small parameter ($h = k^{-1}$ in $h^2\Delta u + u = 0$) is the topic of "semiclassical analysis" (the name comes from the analogy with the relation *particles : waves = classical physics : quantum physics*).

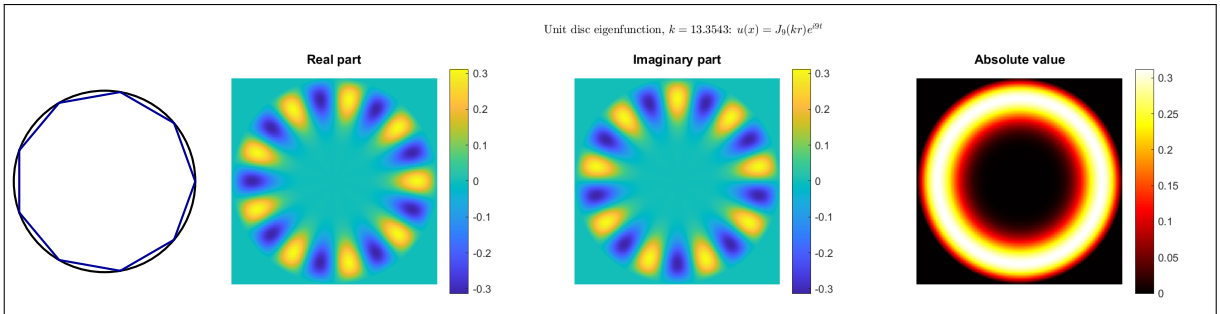
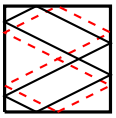


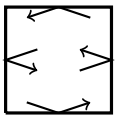
Figure 8: The Dirichlet eigenfunction $J_9(kr)\mathrm{e}^{\mathrm{i}9\theta}$ with $k = 13.3543$ on the unit disc. This is a typical "creeping wave" or "whispering gallery mode", as it is concentrated along the boundary of the domain.

Exercise 4.11: (Billiards and eigenfunctions). Show that for all pairs $j_1, j_2 \in \mathbb{N}$ there is a closed "billiard trajectory" in the unit square $\Omega = (0,1)^2$ whose length is an integer multiple of the wavelength $\lambda_{j_1,j_2} = \frac{2\pi}{k_{j_1,j_2}} = \frac{2}{\sqrt{j_1^2+j_2^2}}$ of the Dirichlet eigenfunction with indices j_1, j_2.

A "billiard trajectory" is the trajectory of a particle leaving from a point in Ω, moving in a straight line, that is reflected when it hits $\partial\Omega$ with the equal-angle law, and that never hits the corners of Ω. It is closed if it is a loop, i.e. the particle repeats it infinitely many time.



Exercise 4.12: (Plane waves and eigenfunctions of the square).
Fix $j_1, j_2 \in \mathbb{N}$ and define the plane wave $u_{\searrow}(\mathbf{x}) = \mathrm{e}^{\mathrm{i}\pi(j_1 x_1 - j_2 x_2)}$. Write the wave $u_{\nearrow}$ defined as the sound-soft reflection of $u_{\searrow}$ on the line $\{x_2 = 0\}$, using the rules learned in §4.1. Extend these rules to write the wave $u_{\nwarrow}$ defined as the reflection of $u_{\nearrow}$ on the line $\{x_1 = 1\}$, and $u_{\swarrow}$ as the reflection of $u_{\nwarrow}$ on the line $\{x_2 = 1\}$. Show that the reflection of $u_{\swarrow}$ on the line $\{x_1 = 0\}$ is $u_{\searrow}$.



We have found a wave that after four sound-soft reflections on the lines corresponding to the sides of the square $(0,1)^2$ is identical to itself.

Show that $u_{\searrow} + u_{\nearrow} + u_{\nwarrow} + u_{\swarrow}$ is (a multiple of) one of the eigenfunctions of $(0,1)^2$ described at the beginning of this section.

Exercise 4.13: (Domain scaling and eigenvalues). Let Ω be a bounded domain, $\rho > 0$ a constant factor, and $\widehat{\Omega} = \rho\Omega = \{\hat{\mathbf{x}} = \rho\mathbf{x} : \ \mathbf{x} \in \Omega\}$ be a dilation of Ω. Let Λ be a Dirichlet (or Neumann) eigenvalue of Ω. Show that $\widehat{\Lambda} = \Lambda/\rho$ is a Dirichlet (Neumann, respectively) eigenvalue of $\widehat{\Omega}$.

4.3 Exterior boundary value problems

A typical problem in computational wave propagation is that of **scattering**. In a scattering problem we want to compute how a given incoming wave is perturbed by the interaction with an obstacle. Here we consider only sound-soft obstacles.

We fix some notation. Let $\boxed{\Omega_-} \subset \mathbb{R}^2$ be a bounded Lipschitz domain, denote $\boxed{\Omega_+} := \mathbb{R}^2 \setminus \overline{\Omega_-}$ and $\boxed{\Gamma} = \partial\Omega_-$. We will always assume that Ω_+ is connected, i.e. Ω_- has no holes. We choose the unit normal vector field $\boxed{\mathbf{n}}$ on $\partial\Omega$ that points out of Ω_- and into Ω_+. We need to take traces of fields defined in Ω_- and in Ω_+: for clarity we write $\boxed{\gamma^{\pm}}$ and $\boxed{\partial_{\mathbf{n}}^{\pm}}$ for the traces taken from Ω_+ and Ω_-. If $u \in H^1_{\mathrm{loc}}(\mathbb{R}^2)$ then $\gamma^+ u = \gamma^- u$ (and we may write γu); if instead $u \in H^1_{\mathrm{loc}}(\Omega_- \cup \Omega_+)$ then $\gamma^+ u$ and $\gamma^- u$ might differ. The same holds for the Neumann traces $\partial_{\mathbf{n}}^{\pm}$ and the $H^1(\cdot;\Delta)$ spaces.

Let $\boxed{u^{\mathrm{Inc}}}$ be the incoming wave, or incident wave, a given Helmholtz solution which will be the datum of our scattering problem. We want to find the field $\boxed{u^{\mathrm{Scat}}}$ scattered by Ω_-, that is a Helmholtz solution in the exterior domain Ω_+ and such that $\gamma(u^{\mathrm{Tot}}) = 0$ on Γ, where $\boxed{u^{\mathrm{Tot}}} = u^{\mathrm{Inc}} + u^{\mathrm{Scat}}$. We see in the next section that these two conditions are not enough to determine u^{Scat}.

4.3.1 Example: scattering by a disc

Let us consider a simple example using separation of variables. Assume that (i) $\Omega_- = B_R$ is the disc of radius $R > 0$ centred at the origin, and (ii) the trace of u^{Inc} on Γ is a circular harmonic, in polar coordinates $(\gamma^+ u^{\mathrm{Inc}})(R,\theta) = \mathrm{e}^{\mathrm{i}\ell\theta}$ for some $\ell \in \mathbb{Z}$. From §2.3, we know that all fields in the form[13]

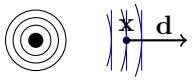
$$u_\lambda^{\mathrm{Scat}}(r,\theta) = -\lambda \frac{H_\ell^{(1)}(kr)}{H_\ell^{(1)}(kR)} \mathrm{e}^{\mathrm{i}\ell\theta} - (1-\lambda)\frac{H_\ell^{(2)}(kr)}{H_\ell^{(2)}(kR)} \mathrm{e}^{\mathrm{i}\ell\theta}, \qquad \lambda \in \mathbb{C},$$

are Helmholtz solutions in Ω_+ and satisfy $\gamma^+(u^{\mathrm{Inc}} + u_\lambda^{\mathrm{Scat}}) = 0$ on the circle Γ. Which value of λ should we choose?

The scattered field is produced by the interaction of the obstacle Ω_- and the incoming field u^{Inc}. So it should look like a wave propagating away from Ω_- towards infinity.

First of all, we would like $|u(\mathbf{x})|$ to decrease to zero for $r \to \infty$ (the further we are from a sound source, the weaker the sound we hear). In particular if $|u(\mathbf{x})|^2 \sim r^{-1}$ for $r \to \infty$, then the "energy" $\int_{\{|\mathbf{x}|=R\}} |u|^2 \,\mathrm{d}s$ is bounded for $R \to \infty$. All Fourier–Bessel and Fourier–Hankel functions decay as $\sqrt{2/\pi kr}$ for $r \to \infty$, so this does not help choosing λ.

If we plot u^{Scat} in a position $\mathbf{x}$ very far from Ω_-, we expect it to point away from Ω_-, i.e. radially towards infinity. We would like u^{Scat} close to $\mathbf{x}$ to look like a plane wave pointing away from Ω_-, i.e. in the direction $\mathbf{d} = \frac{\mathbf{x}}{r}$:



$$u^{\mathrm{Scat}}(\mathbf{x}) \approx A\mathrm{e}^{\mathrm{i}k\mathbf{x}\cdot\frac{\mathbf{x}}{r}} = A\mathrm{e}^{\mathrm{i}kr} \qquad \mathbf{x} = (r\cos\theta, r\sin\theta).$$

Here $A \in \mathbb{C}$ includes the amplitude, proportional to $\frac{1}{\sqrt{r}}$, and the phase of the wave.

[13] Here we have chosen the Hankel functions, as opposed to J_ℓ and Y_ℓ, because they are different from 0 for all values of kR, so we can normalise as written.

We recall that Bessel functions with large arguments can be approximated by the following formulas, [CK2, (3.82)]:

$$J_\ell(z) = \sqrt{\frac{2}{\pi z}}\cos\Big(z - \frac{\ell\pi}{2} - \frac{\pi}{4}\Big)\Big(1+\mathcal{O}\Big(\frac{1}{z}\Big)\Big), \quad Y_\ell(z) = \sqrt{\frac{2}{\pi z}}\sin\Big(z - \frac{\ell\pi}{2} - \frac{\pi}{4}\Big)\Big(1+\mathcal{O}\Big(\frac{1}{z}\Big)\Big), \quad z\to\infty.$$

Recalling the relations (16) between Bessel and Hankel functions, we can write the scattered field as

$$u_\lambda^{\mathrm{Scat}}(\mathbf{x}) \approx -\sqrt{\frac{2}{\pi kr}}\left(\lambda\frac{\mathrm{e}^{\mathrm{i}(kr-\frac{\ell\pi}{2}-\frac{\pi}{4}+\ell\theta)}}{H_\ell^{(1)}(kR)} + (1-\lambda)\frac{\mathrm{e}^{\mathrm{i}(-kr+\frac{\ell\pi}{2}+\frac{\pi}{4}+\ell\theta)}}{H_\ell^{(2)}(kR)}\right)\left(1+\mathcal{O}\Big(\frac{1}{kr}\Big)\right). \tag{29}$$

We see that the $H_\ell^{(1)}$ term gives a factor $\mathrm{e}^{\mathrm{i}kr}$, while $H_\ell^{(2)}$ term gives a factor $\mathrm{e}^{-\mathrm{i}kr}$. This means that only the first component is propagating outward, while the second one is directed towards the origin. So we want to keep the first term only and choose $\lambda = 1$.

Another way to see that the $\mathrm{e}^{\mathrm{i}kr}$ terms are outgoing is to recall the meaning of time-harmonic waves, as described in §1.2. For $u(\mathbf{x}) = r^{-1/2}\mathrm{e}^{\mathrm{i}kr}$, the time-dependent field (7) is $U(\mathbf{x},t) = r^{-1/2}\cos(kr-\omega t) = r^{-1/2}\cos(k(r-ct))$ which spreads (in time) radially from the origin. Conversely, $\tilde u(\mathbf{x}) = r^{-1/2}\mathrm{e}^{-\mathrm{i}kr} = \overline{u}(\mathbf{x})$ gives $\tilde U(\mathbf{x},t) = r^{-1/2}\cos(kr+\omega t) = U(\mathbf{x},-t)$ which moves towards the origin. See also Figure 9 for another way of reading the direction of propagation of a Fourier–Hankel function.

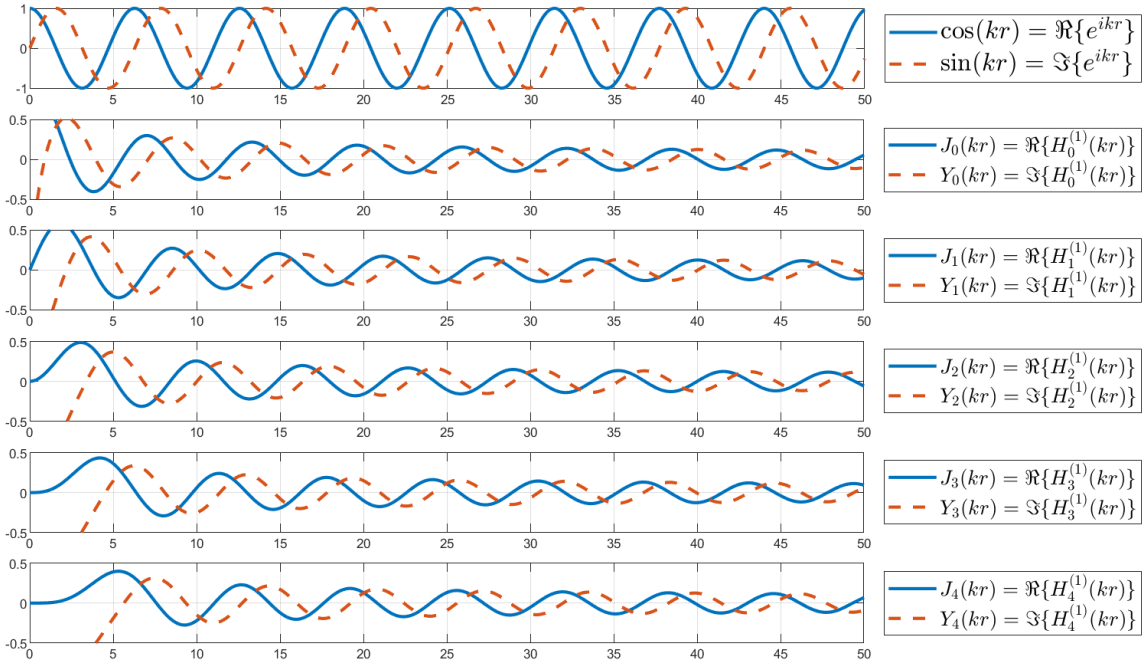


Figure 9: How can we see that $u(\mathbf{x}) = H_\ell^{(1)}(kr)\mathrm{e}^{\mathrm{i}\ell\theta}$ moves outwards from its definition? In this figure we compare the real part (blue continuous lines) and the imaginary part (red dashed line) of these circular waves for $\ell = 1,2,3,4$ and $k = 1$ along a ray from the origin $(0,0)$ (left) to the point $(50,0)$ (right). We see that the imaginary part is always slightly "ahead" of the real part. Recall Exercise 1.9: the time-dependent wave $U(\mathbf{x},t)$, after taking value $\Re u(x)$ will take value $\Im u(x)$. From the figures we see that this means that u is moving from left to right, i.e. from the origin towards infinity.
For comparison, in the first row we see the real and the imaginary parts of a plane wave propagating to the right.

Exercise 4.14: (Circular wave motion). Using the Matlab code provided, make time-harmonic animations of different combinations of Fourier–Bessel and Fourier–Hankel functions and observe in which direction they propagate.

For a general u^{Inc}, we can expand its trace on the circle Γ in circular harmonics as $u^{\mathrm{Inc}}(R,\theta) = \sum_{\ell\in\mathbb{Z}} a_\ell \mathrm{e}^{\mathrm{i}\ell\theta}$. The scattered field and the total field are then

$$u^{\mathrm{Scat}}(r,\theta) = -\sum_{\ell\in\mathbb{Z}} a_\ell \frac{H_\ell^{(1)}(kr)}{H_\ell^{(1)}(kR)}\mathrm{e}^{\mathrm{i}\ell\theta}, \qquad u^{\mathrm{Tot}}(\mathbf{x}) = u^{\mathrm{Inc}}(\mathbf{x}) + u^{\mathrm{Scat}}(\mathbf{x}).$$

This choice ensures that (i) u^{Scat} is Helmholtz solution in Ω_+, (ii) $\gamma^+ u^{\mathrm{Tot}} = 0$ on $\partial\Omega$, and (iii) u^{Scat} is made of outgoing waves only. This is an example of "Mie series".

You can see an example of scattered field computed with this formula in Figure 10 and on the webpage.

Exercise 4.15: (Scattering of a plane wave by a disc). Let u^{Inc} be a plane wave with direction $\mathbf{d}$.

Compute the field scattered by a disc of radius R using Jacobi–Anger formula $\mathrm{e}^{\mathrm{i}z\cos\alpha} = \sum_{\ell\in\mathbb{Z}} \mathrm{i}^\ell J_\ell(z)\mathrm{e}^{\mathrm{i}\ell\alpha}$. Reproduce the plots in Figure 10.

The sum over $\ell \in \mathbb{Z}$ has to be truncated at $|\ell| \le \ell_{\max}$, with $\ell_{\max}$ (slightly) larger than kR. Plot the magnitude of the coefficients $\hat{u}_\ell$ in $u^{\mathrm{Scat}}(\mathbf{x}) = \sum_{\ell\in\mathbb{Z}} \hat{u}_\ell H_\ell^{(1)}(kr)\mathrm{e}^{\mathrm{i}\ell\theta}$ to justify this choice.

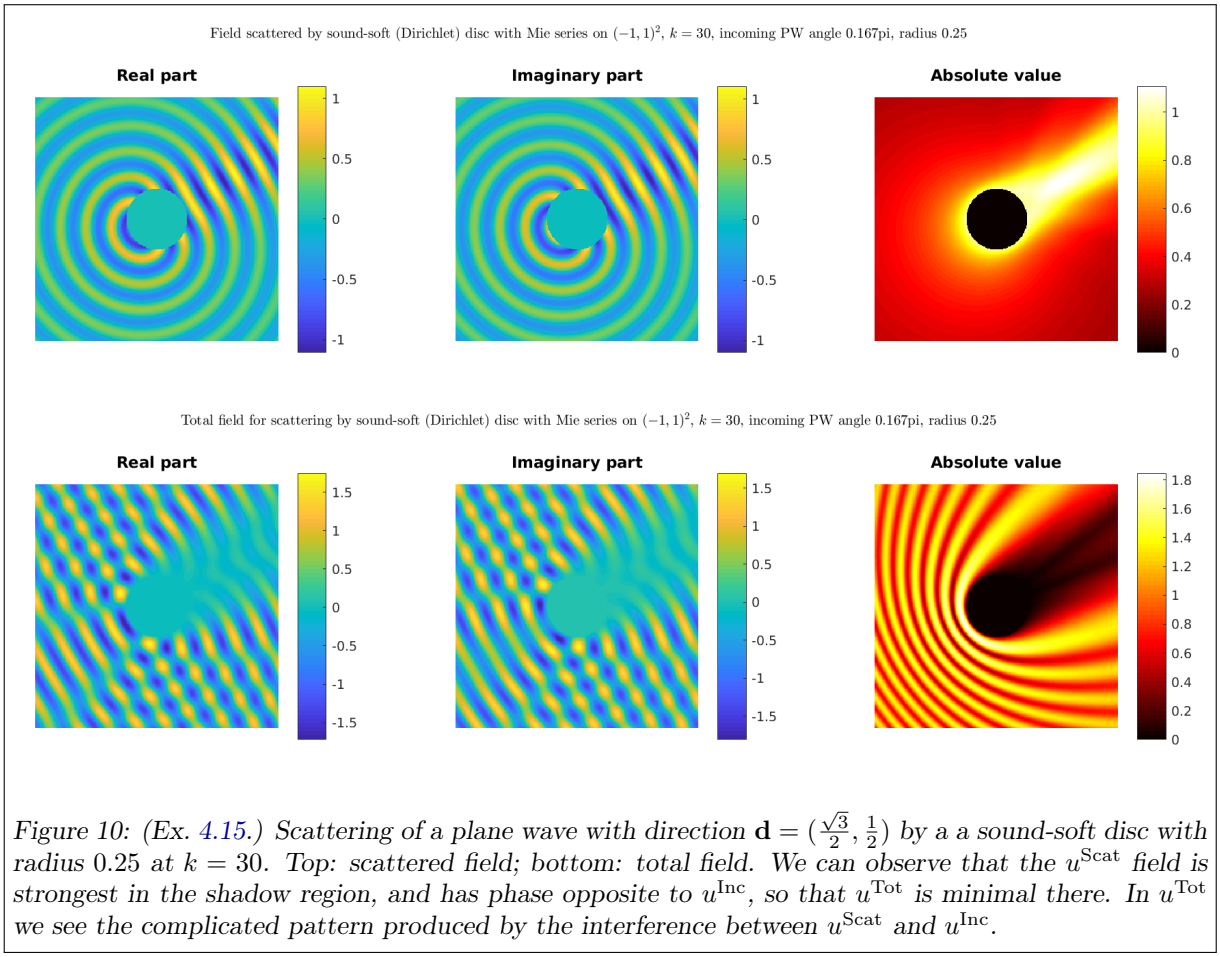


Figure 10: (Ex. 4.15.) Scattering of a plane wave with direction $\mathbf{d} = (\frac{\sqrt{3}}{2}, \frac{1}{2})$ *by a a sound-soft disc with radius* 0.25 *at* $k = 30$. *Top: scattered field; bottom: total field. We can observe that the* u^{Scat} *field is strongest in the shadow region, and has phase opposite to* u^{Inc}, *so that* u^{Tot} *is minimal there. In* u^{Tot} *we see the complicated pattern produced by the interference between* u^{Scat} *and* u^{Inc}.

Exercise 4.16: (Scattering of a circular wave by a disc).

1. Using the "Graf's addition theorem" in [DLMF, 10.23.7], derive the formula

$$H_0^{(1)}(k|\mathbf{x}-\mathbf{y}|) = \sum_{\ell\in\mathbb{Z}} H_\ell^{(1)}(k|\mathbf{y}|)\mathrm{e}^{-\mathrm{i}\ell\theta_{\mathbf{y}}}\, J_\ell(k|\mathbf{x}|)\mathrm{e}^{\mathrm{i}\ell\theta_{\mathbf{x}}} \qquad \forall \mathbf{x},\mathbf{y}\in\mathbb{R}^2, \quad |\mathbf{x}| < |\mathbf{y}|. \tag{30}$$

   Here $\theta_{\mathbf{x}}$ and $\theta_{\mathbf{y}}$ are the angular coordinates of $\mathbf{x}$ and $\mathbf{y}$.

2. Let $u^{\mathrm{Inc}}(\mathbf{x}) = H_0^{(1)}(k|\mathbf{x}-\mathbf{y}|)$ be a 0-order Fourier–Hankel function centred at $\mathbf{y} \in \mathbb{R}^2 \setminus \overline{B_R}$, $R > 0$. Show that, for this incoming field, the field scattered by the disc B_R, is

$$u^{\mathrm{Scat}}(r,\theta) = \sum_{\ell\in\mathbb{Z}} \frac{J_\ell(kR)\; H_\ell^{(1)}(k|\mathbf{y}|)\; \mathrm{e}^{-\mathrm{i}\ell\theta_{\mathbf{y}}}}{H_\ell^{(1)}(kR)}\, H_\ell^{(1)}(kr)\mathrm{e}^{\mathrm{i}\ell\theta} \qquad r > R.$$

3. Plot the scattered and the total fields.

4.3.2 Sound-soft scattering problems

We have seen how to select "outgoing" waves using the expansion in polar coordinates. How to do the same when this expansion is not available, namely when Ω_- is not a circle?

The radial dependence of all the outgoing terms in the circular wave approximation (29) is $\frac{1}{\sqrt{r}}\mathrm{e}^{\mathrm{i}kr}$ (ignoring high-order terms). Deriving with respect to the radial direction r we have $\partial_r(\frac{1}{\sqrt{r}}\mathrm{e}^{\mathrm{i}kr}) = \mathrm{i}k\frac{1}{\sqrt{r}}\mathrm{e}^{\mathrm{i}kr} - \frac{1}{2}r^{-3/2}\mathrm{e}^{\mathrm{i}kr}$. So, if u^{Scat} is a linear combination of $H_\ell^{(1)}(kr)\mathrm{e}^{\mathrm{i}\ell\theta}$ for different ℓs, then it satisfies $\partial_r u^{\mathrm{Scat}} - \mathrm{i}ku^{\mathrm{Scat}} = \mathcal{O}(r^{-3/2})$. On the other hand, the bad terms $H_\ell^{(2)}(kr)\mathrm{e}^{\mathrm{i}\ell\theta}$ satisfy only the condition with the opposite sign $\partial_r u + \mathrm{i}ku = \mathcal{O}(r^{-3/2})$ (recall that $H_\ell^{(2)} = \overline{H_\ell^{(1)}}$). This suggests the following classical definition (for the history of the Sommerfeld radiation condition see [Schot, 1992]).

Definition 4.17: (Radiating/outgoing solution). Let u be an $H^1_{\mathrm{loc}}(\mathbb{R}^2 \setminus B_R)$ solution of the Helmholtz equation in the complement of a ball. We say that u is **radiating**, or **outgoing**, if it satisfies the **Sommerfeld radiation condition**:

$$|\partial_r u - \mathrm{i}ku| = o(r^{-1/2}) \qquad r \to \infty. \tag{31}$$

Sommerfeld condition is meant to hold uniformly in all directions, namely

$$\lim_{r\to\infty} \sup_{\theta\in[0,2\pi]} \sqrt{r}\left|\partial_r u(r,\theta) - \mathrm{i}ku(r,\theta)\right| = 0.$$

Since radiating solutions can be expanded in series of Fourier–Hankel functions, Sommerfeld condition (31) is equivalent (for Helmholtz solutions only) to the apparently stronger condition

$$\exists C, R > 0 \quad \text{such that} \quad \left|\partial_r u(r,\theta) - \mathrm{i}ku(r,\theta)\right| \le Cr^{-3/2} \qquad \forall r > R, \quad \theta \in [0, 2\pi].$$

Sommerfeld condition also implies that $\lim_{R\to\infty}\int_{\partial B_R} |\partial_r u - \mathrm{i}ku|^2 \,\mathrm{d}s = 0$, $\partial B_R = \{\mathbf{x} : |\mathbf{x}| = R\}$. The approximation (29) implies that all linear combinations of $H_\ell^{(1)}(kr)\mathrm{e}^{\mathrm{i}\ell\theta}$ that converge outside of some ball are radiating. On the the other hand, no $H_\ell^{(2)}(kr)\mathrm{e}^{\mathrm{i}\ell\theta}$ term is allowed in radiating functions. The complex-conjugate of a radiating function is not radiating.

A more rigorous derivation of the Sommerfeld radiation condition can be done using the "limiting absorption principle": first consider the problem with absorption, i.e. $\Im k > 0$, where the $\mathrm{e}^{\mathrm{i}kr}$ behaviour corresponds to solutions decaying towards infinity, then study the limit for $\Im k \searrow 0$.

We can now define the class of exterior boundary value problems that we will consider in the following.

Definition 4.18: (Exterior Dirichlet problem—EDP). Let Ω_- be a bounded Lipschitz domain, $k > 0$ and $g_D \in H^{\frac{1}{2}}(\Gamma)$. We say that $u \in H^1_{\mathrm{loc}}(\overline{\Omega_+})$ satisfies the exterior Helmholtz Dirichlet problem if

$$\begin{aligned} \Delta u + k^2 u &= 0 && \text{in } \Omega_+, \\ \gamma^+ u &= g_D && \text{on } \Gamma, \\ u &\text{ is radiating.} \end{aligned} \tag{32}$$

In the language of scattering theory:

Definition 4.19: (Sound-soft scattering problem—SSSP). Let Ω_- a bounded Lipschitz domain, $k > 0$ and u^{Inc} is a Helmholtz solution in a neighbourhood of Γ. We say that $u^{\mathrm{Scat}} \in H^1_{\mathrm{loc}}(\overline{\Omega_+})$ satisfies the sound-soft scattering problem if

$$\begin{aligned} \Delta u^{\mathrm{Scat}} + k^2 u^{\mathrm{Scat}} &= 0 && \text{in } \Omega_+, \\ \gamma^+(u^{\mathrm{Scat}} + u^{\mathrm{Inc}}) &= 0 && \text{on } \Gamma, \\ u^{\mathrm{Scat}} &\text{ is radiating.} \end{aligned} \tag{33}$$

The sound-soft scattering problem is an exterior Dirichlet problem with $u = u^{\mathrm{Scat}}$ and $g_D = -\gamma^+ u^{\mathrm{Inc}}$.

We will see in §4.4 that problems (32) and (33) are well-posed.

In Definition 4.19 we have assumed that u^{Inc} is defined only in a neighbourhood of the scatterer's boundary. If the incoming wave is a plane wave, then of course it is defined in the whole of $\mathbb{R}^2$, but this definition allows to include more realistic incoming waves such as "point sources", i.e. Fourier–Hankel functions centred at some point of Ω_+.

The EDP and the SSSP are defined and analysed in details in, e.g., [CK2, §3.2]. However, [CK2] considers the 3D case (so the powers of r in Sommerfeld condition are different) and C^m spaces rather than Sobolev spaces.

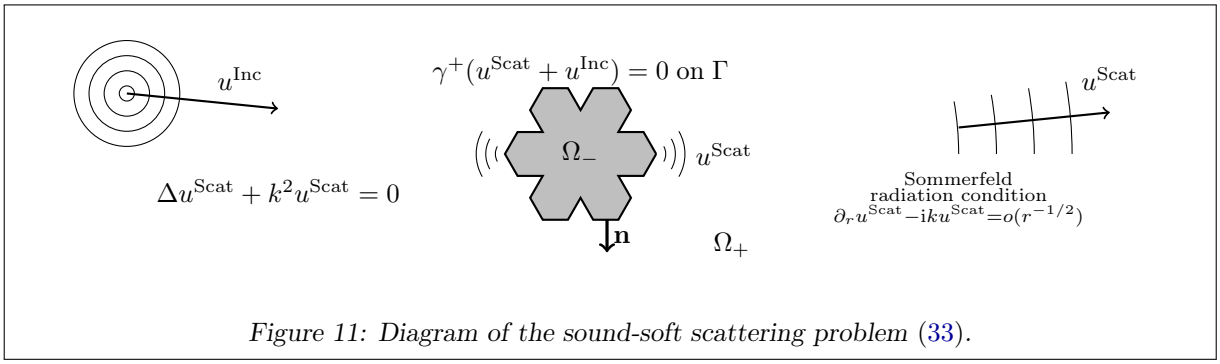


Figure 11: Diagram of the sound-soft scattering problem (33).

Remark 4.20: (Truncated problems). Often one does not want to deal with BVPs posed on unbounded domains such as in Definition 4.19, for example because one wants to approximate the solution with a finite element method. A possibility to reduce this problem to one posed on a bounded domain is to choose a large ball B_R (or a different shape) with $\Omega_- \subset B_R$ and solve a Helmholtz BVP on the truncated domain $B_R \cap \Omega_+$. On Γ we impose the Dirichlet condition as above. On the artificial boundary $\Gamma_R = \partial B_R$ one has to impose some artificial boundary condition that mimics the Sommerfeld radiation condition. The simplest choice is to choose impedance conditions $\partial_r u^{\mathrm{Scat}} - \mathrm{i}k u^{\mathrm{Scat}} = 0$ (compare with (31)). Many more efficient and more complicated boundary conditions exist. They are called absorbing, non-reflecting, radiation, generalised-impedance boundary conditions (ABC, NRBC, GIBC...). The quality of an artificial boundary condition depends on the ability to absorb the waves coming from the domain and to not reflect them back. See some examples in [Ihl98, Ch. 3].

Remark 4.21: (Far-field pattern). It is possible to prove (e.g. [CK2, eq. (3.86)]) that if u is a radiating Helmholtz solution, then it satisfies

$$u(\mathbf{x}) = \frac{\mathrm{e}^{\mathrm{i}kr}}{\sqrt{r}}\Big(u_\infty(\theta) + \mathcal{O}(r^{-1})\Big) \qquad \text{for } r = |\mathbf{x}| \to \infty,$$

for a function $u_\infty \in C^\infty(\mathbb{S}^1)$ (recall that $\mathbb{S}^1$ is the unit circle, and that θ denotes the angular polar coordinate of $\mathbf{x}$). This means that, up to factoring out the phase factor $\mathrm{e}^{\mathrm{i}kr}$ and the decay factor $\frac{1}{\sqrt{r}}$, when we move towards infinity along a straight line in the direction θ, a radiating field converges to a given value $u_\infty(\theta)$. The function u_∞ is called *far-field pattern*. If u is defined in Ω_+ and admits Dirichlet and Neumann traces on Γ, the far-field pattern can be computed using the formula

$$u_\infty(\theta) = \frac{\mathrm{e}^{\mathrm{i}\frac{\pi}{4}}}{\sqrt{8\pi k}} \int_\Gamma \Big(\gamma^+ u(\mathbf{y}) \partial_{\mathbf{n}}^+ \mathrm{e}^{-\mathrm{i}k\mathbf{y}\cdot\mathbf{d}} - \partial_{\mathbf{n}}^+ u(\mathbf{y}) \mathrm{e}^{-\mathrm{i}k\mathbf{y}\cdot\mathbf{d}}\Big)\, \mathrm{d}s(\mathbf{y}) \qquad \mathbf{d} = (\cos\theta, \sin\theta). \tag{34}$$

The far-field pattern is one of the main quantities of interest in remote-sensing applications, for example to quantify the amount of radiation "backscattered" by an obstacle when it is hit by a wave. You can see the far-field pattern of the wave scattered by a triangle in Figure 14.

Exercise 4.22: (Far-field pattern of a plane wave scattered by a disc). Compute (as a circular harmonic expansion) and plot with Matlab the far-field pattern of the field scattered by a disc hit by a plane wave. Use the Fourier expansion computed in Exercise 4.15. Study how the far-field pattern vary with k, R and the propagation direction of u^{Inc}. Denote $u_\infty(\theta, \xi)$ the far-field for $u^{\mathrm{Inc}}(\mathbf{x}) = \mathrm{e}^{\mathrm{i}k(x_1\cos\xi + x_2 \sin\xi)}$: can you find any symmetry between the two angles? See [CK2, Thm. 3.15].

Remark 4.23: (Direct and inverse scattering). The SSSP is a *direct* scattering problem: we know the incoming wave, we know the obstacle, we want to compute the scattered field. In applications (such as medical imaging, oil retrieval, seismic and atmospheric remote sensing, fault detection in materials, radar and sonar...) it is very important to consider also *inverse* scattering problem: given the scattered field or the far-field (typically from measurements), one wants to compute the obstacle and/or the incoming wave. Inverse problems are ill-posed and much harder than direct ones, both theoretically and computationally. Most numerical methods for the approximation of inverse problems require the numerical solution of many direct problems: this is one of the main motivations for the study of efficient methods to simulate direct scattering problems. Most of the book [CK2] is devoted to inverse scattering problems.

Remark 4.24: (Radiation conditions in 3D, electromagnetism and elasticity). In three space dimensions, the Sommerfeld condition requires a slightly faster decay of u: $|\partial_r u - \mathrm{i}ku| = o(r^{-1})$, $r \to \infty$, [CK2, Def. 2.4].

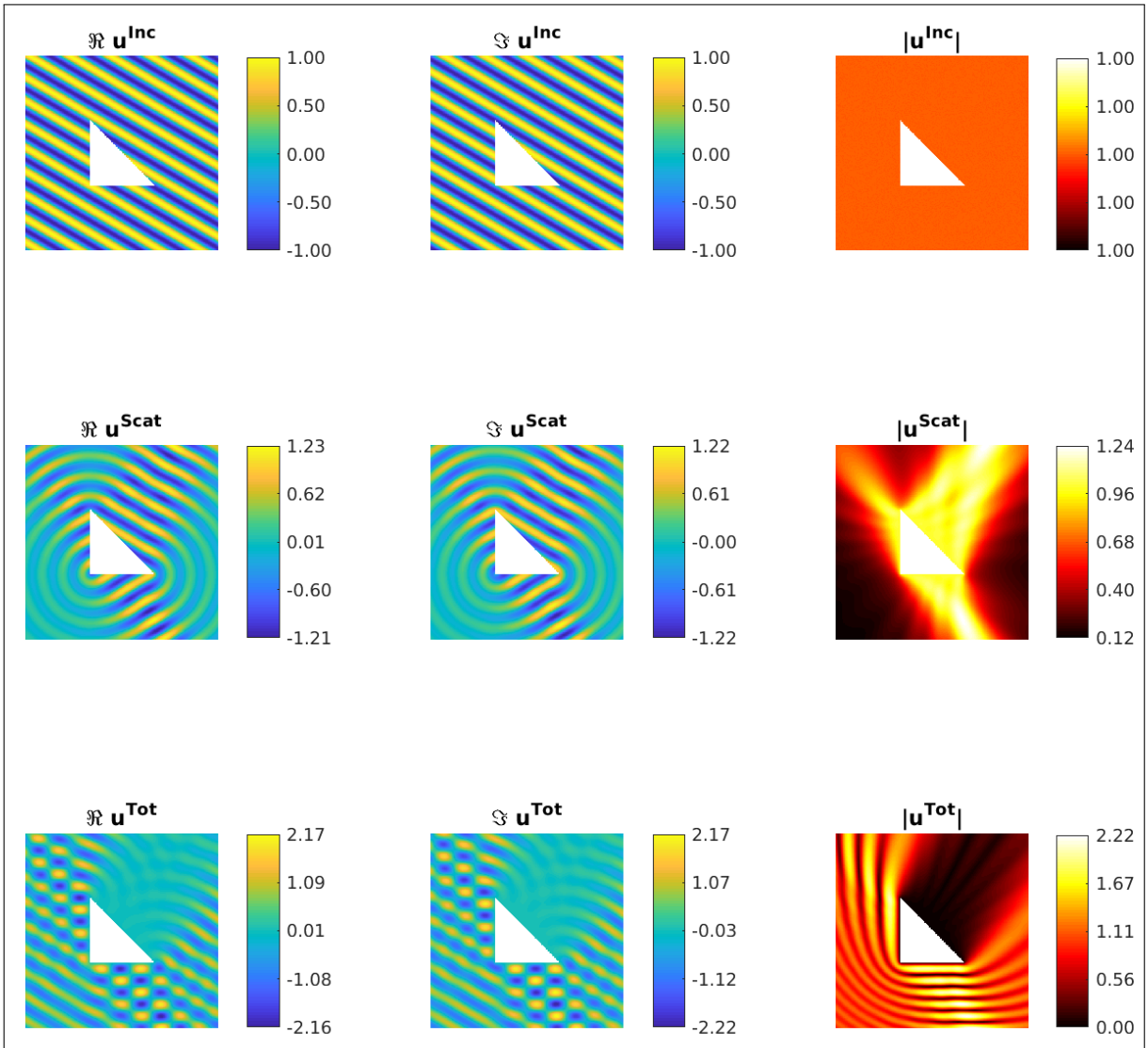


Figure 12: Scattering of a plane wave with direction $\mathbf{d} = (\frac{1}{2}, \frac{\sqrt{3}}{2})$ *by a sound-soft triangle. Here* $k = 20$ *(so* $\lambda = \frac{\pi}{10}$*), the cathetus of the triangle has length 1 and the fields are plotted on a square of side 3.*

A pair $(\mathbf{E}, \mathbf{H})$ that solves the time-harmonic Maxwell's equations (9) (with $\sigma = 0$) in the complement of a bounded domain of $\mathbb{R}^3$ is outgoing if it satisfies one of the **Silver–Müller** radiation conditions [CK2, Def. 6.6]:

$$|\mathbf{H} \times \mathbf{x} - r\mathbf{E}| = o(1) \quad \text{or} \quad |\mathbf{E} \times \mathbf{x} + r\mathbf{H}| = o(1) \qquad r \to \infty.$$

This is equivalent to assuming that all the Cartesian components (which are Helmholtz solutions) of the fields satisfy the Sommerfeld condition, [CK2, Thm. 6.8].

A displacement field $\mathbf{u}$ that solves Navier's equations (12) in the complement of a bounded domain is outgoing if it satisfies the **Kupradze** condition: the scalar and vector potentials χ and $\boldsymbol{\psi}$ of (13) satisfy the Sommerfeld condition with wavenumber k_P and k_S, respectively. So all Cartesian components of radiating pressure and shear waves are radiating Helmholtz solutions.

4.4 Well-posedness of the exterior Dirichlet problem (EDP)

The most common proof of the well-posedness of the EDP (32) relies on properties of BIOs and BIEs, e.g. [CK1, Thm. 3.21] and [CK2, Thm. 3.11]. Here instead we prove well-posedness using a variational formulation on a truncated domain and the "DtN map". However, both proofs rely on the same main tools: Fredholm theory and a "Rellich lemma", which ensures uniqueness.

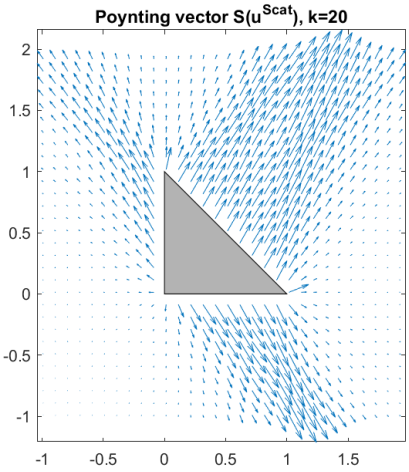


Figure 13: The Poynting vector $\mathbf{S}(u^{\mathrm{Scat}})$ *for the scattering problem in Figure 12. See Exercise 2.13.*

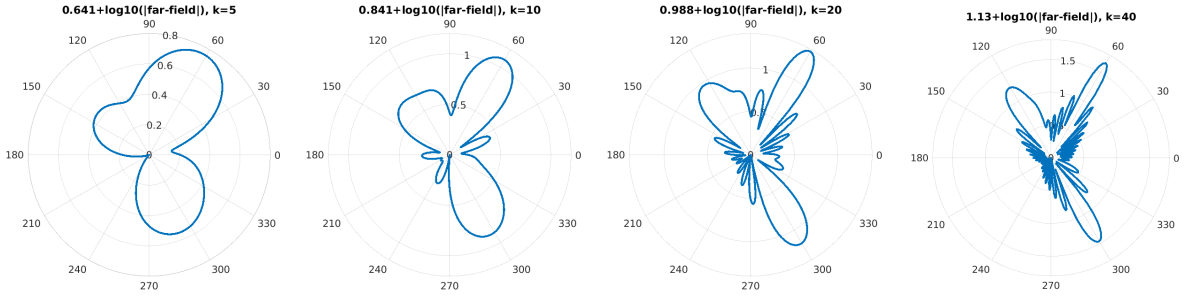


Figure 14: Some polar logarithmic plots (Matlab's `polarplot` *command) of the magnitudes of far-field patterns* $\log_{10}|u_\infty|$*. In this scattering problem, a plane wave with direction* $\frac{\pi}{3}$ *hits a sound-soft triangular scatterer* ◣ *with vertices* $(0,0)$*,* $(1,0)$ *and* $(0,1)$*, as in Figure 12. Each plot corresponds to a different wavenumber* ($k = 5, 10, 20, 40$)*: for increasing frequencies the far field becomes more complex and focused in few directions. The far-field pattern has maximal intensity in the direction* $\frac{\pi}{3}$ *of the incoming wave (up right), where the triangle projects its shadow. Two other peaks are in directions* $-\frac{\pi}{3}$ *and* $\frac{2\pi}{3}$*, corresponding to the wave reflected by the two illuminated sides. The field in all other directions is due to the diffraction by the corners. The far-field patterns were computed numerically with the BEM described in §5.2.*

4.4.1 DtN map

Let $R > 0$ be the radius of an open ball B_R centred at the origin such that $\Gamma \subset B_R$. We define the operator DtN, which acts on functions defined on ∂B_R by multiplying each terms in their Fourier expansion by the ratio of the radial derivative of the corresponding Fourier–Hankel function and the value of the Fourier–Hankel function itself:

$$\mathsf{DtN}(v) = \mathsf{DtN}\Big(\sum_{\ell\in\mathbb{Z}} \hat{v}_\ell \mathrm{e}^{\mathrm{i}\ell\theta}\Big) = \sum_{\ell\in\mathbb{Z}} T_\ell \hat{v}_\ell \mathrm{e}^{\mathrm{i}\ell\theta}, \qquad \text{for } T_\ell := \frac{kH_\ell^{(1)\prime}(kR)}{H_\ell^{(1)}(kR)}. \tag{35}$$

This operator is called **Dirichlet-to-Neumann (DtN) map** or **capacity operator** (see [Néd01, (2.6.92)] for the 3D version).

If u is a radiating solution in Ω_+, then in $\mathbb{R}^2 \setminus B_R$ it can[14] be expanded as $u(\mathbf{x}) = \sum_{\ell\in\mathbb{Z}} a_\ell H_\ell^{(1)}(kr)\mathrm{e}^{\mathrm{i}\ell\theta}$. Then its traces on ∂B_R are

$$\gamma^+_{\partial B_R} u = \sum_{\ell\in\mathbb{Z}} a_\ell H_\ell^{(1)}(kR)\mathrm{e}^{\mathrm{i}\ell\theta} \quad \text{and} \quad \partial^+_{\mathbf{n}\,\partial B_R}(u) = \sum_{\ell\in\mathbb{Z}} a_\ell k H_\ell^{(1)\prime}(kR)\mathrm{e}^{\mathrm{i}\ell\theta} \quad \text{thus} \quad \boxed{\mathsf{DtN}(\gamma^+_{\partial B_R} u) = \partial^+_{\mathbf{n}\,\partial B_R}(u).}$$

[14] We haven't proved that *all* radiating solutions can be expanded in Fourier–Hankel series; see [CK1, Thm. 3.6] for a proof (in the 3D case) which relies on Green's representation (which we will see later).

In words: **the DtN operator maps the Dirichlet trace** (on ∂B_R) **of a radiating solution to its Neumann trace** (on ∂B_R).

From the formulas ([DLMF, 10.6i and 10.19E2]) for the derivative and the large-index asymptotics of the Hankel functions

$$H_\ell^{(1)\prime}(z) = \frac{H_{\ell-1}^{(1)}(z) - H_{\ell+1}^{(1)}(z)}{2}, \qquad H_\ell^{(1)}(z) \sim -\mathrm{i}\sqrt{\frac{2}{\pi}}\Big(\frac{2}{\mathrm{e}z}\Big)^\ell \ell^{\ell-\frac12} \quad \text{for } \ell\to\infty$$

we have

$$T_\ell = k\frac{H_{\ell-1}^{(1)}(kR) - H_{\ell+1}^{(1)}(kR)}{H_\ell^{(1)}(kR)} \sim k\Big(\frac{\mathrm{e}kR}{2\ell} - \frac{2\ell}{\mathrm{e}kR}\Big) = \mathcal{O}(\ell) \qquad \ell\to\infty.$$

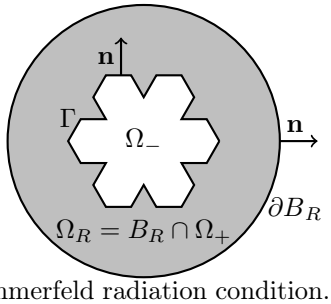
From the definition of the fractional norms (19) on the circle we have that DtN is continuous as an operator $\mathsf{DtN} : H^s(\partial B_R) \to H^{s-1}(\partial B_R)$ for any $s\in\mathbb{R}$: for some $C>0$,

$$\|\mathsf{DtN} v\|^2_{H^{s-1}(\partial B_R)} = 2\pi\sum_{\ell\in\mathbb{Z}} |\hat v_\ell|^2 \underbrace{|T_\ell|^2}_{\sim\ell^2}(1+\ell^2)^{s-1} \le C\sum_{\ell\in\mathbb{Z}} |\hat v_\ell|^2(1+\ell^2)^s \le C\,\|v\|^2_{H^s(\partial B_R)}.$$

4.4.2 Truncated problem

The EDP (32) is equivalent to the following problem on the truncated domain $\Omega_R := B_R\cap\Omega_+$:

$$\begin{aligned} \Delta u + k^2 u &= 0 && \text{in } \Omega_R,\\ \gamma u &= g_D && \text{on } \Gamma, \\ \mathsf{DtN}(\gamma u) - \partial_{\mathbf{n}} u &= 0 && \text{on } \partial B_R. \end{aligned} \tag{36}$$



The last condition on the exterior boundary ∂B_R is equivalent to the Sommerfeld radiation condition. To write this BVP as a variational problem we define the space $H^1_{0,R}(\Omega_R) := \{u\in H^1(\Omega_R) :\ \gamma u = 0 \text{ on } \Gamma\}$ (the H^1 functions whose trace vanishes on the interior boundary Γ but not necessarily on ∂B_R).

From the surjectivity of the trace operator (Theorem 3.7) there exists a lifting $u_D\in H^1(\Omega_R)$ such that $\gamma^+ u_D = g_D$; it is possible to choose $u_D\in H^1(\Omega_R;\Delta)$ (e.g. by solving an auxiliary Laplace BVP). If we can solve the problem

$$\begin{aligned} \Delta u_0 + k^2 u_0 &= -f && \text{in } \Omega_R,\\ \gamma u_0 &= 0 && \text{on } \Gamma, \\ \mathsf{DtN}(\gamma u_0) - \partial_{\mathbf{n}} u_0 &= g_R && \text{on } \partial B_R, \end{aligned} \qquad f := -\Delta u_D - k^2 u_D, \quad g_R := -\mathsf{DtN}(\gamma u_D) + \partial_{\mathbf{n}} u_D,$$

then $u = u_D + u_0$ would solve (36). (We could also choose u_D such that $g_R = 0$.) Using Green's first identity it is easy to deduce a variational problem for u_0:

$$\text{find} \quad u_0\in H^1_{0,R}(\Omega_R) \quad \text{such that} \quad \mathcal{A}_R(u_0,w) = \mathcal{F}_R(w) \quad \forall w\in H^1_{0,R}(\Omega_R) \quad \text{where} \tag{37}$$
$$\mathcal{A}_R(u_0,w) := \int_{\Omega_R} (\nabla u_0\nabla\overline{w} - k^2 u_0\overline{w})\,\mathrm{d}\mathbf{x} - \int_{\partial B_R} (\mathsf{DtN}\gamma u_0)(\gamma\overline{w})\,\mathrm{d}s,$$
$$\mathcal{F}_R(w) := \int_\Omega f\overline{w}\,\mathrm{d}\mathbf{x} - \int_{\partial B_R} g_R\gamma\overline{w}\,\mathrm{d}s.$$

Using the continuity of the trace operator ($\gamma : H^1_{0,R}(\Omega_R)\to H^{\frac12}(\partial B_R)$) and the DtN map ($\mathsf{DtN} : H^{\frac12}(\partial B_R)\to H^{-\frac12}(\partial B_R)$) we deduce the continuity of $\mathcal{A}_R$ and $\mathcal{F}_R$ in $H^1_{0,R}(\Omega_R)$.

4.4.3 Gårding inequality

We first look at the real part of the boundary sesquilinear form associated to DtN: for all $v(\mathbf{x}) = \sum_{\ell\in\mathbb{Z}}\hat v_\ell \mathrm{e}^{\mathrm{i}\ell\theta}\in H^{\frac12}(\partial B_R)$ we have

$$\Re\int_{\partial B_R} (\mathsf{DtN} v)\overline{v}\,\mathrm{d}s(\mathbf{x}) = R\,\Re\int_0^{2\pi}\Big(\sum_{\ell\in\mathbb{Z}} T_\ell\hat v_\ell \mathrm{e}^{\mathrm{i}\ell\theta}\Big)\Big(\sum_{\ell\in\mathbb{Z}}\overline{\hat v_\ell}\mathrm{e}^{-\mathrm{i}\ell\theta}\Big)\,\mathrm{d}\theta = 2\pi R\sum_{\ell\in\mathbb{Z}}|\hat v_\ell|^2\Re\{T_\ell\}.$$

For any complex-valued differentiable function f of a real variable we have

$$\Re\Big\{\frac{f'(t)}{f(t)}\Big\} = \Re\Big\{\frac{f'(t)\overline{f}(t)}{|f(t)|^2}\Big\} = \frac{f'(t)\overline{f}(t)+\overline{f'}(t)f(t)}{2|f(t)|^2} = \frac{1}{2|f(t)|^2}\frac{\partial\big(f(t)\overline{f}(t)\big)}{\partial t} = \frac{1}{2|f(t)|^2}\frac{\partial(|f(t)|^2)}{\partial t}.$$

Choosing $f(r) = H_\ell^{(1)}(kr)$ we have

$$\Re\{T_\ell\} = \Re\Big\{\frac{f'(R)}{f(R)}\Big\} = \frac{1}{2|H_\ell^{(1)}(kR)|^2}\frac{\partial(|H_\ell^{(1)}(kr)|^2)}{\partial r}\Big|_{r=R} < 0$$

since the absolute value of the Hankel function is monotonically decreasing. Combining all these ingredients we see that the sesquilinear form satisfies a Gårding inequality:

$$\begin{aligned}\Re\{\mathcal{A}_R(w,w)\} &= \int_{\Omega_R}(|\nabla w|^2 - k^2|w|^2)\,\mathrm{d}\mathbf{x} - \Re\int_{\partial B_R}(\mathsf{DtN}\gamma w)\gamma\overline{w}\,\mathrm{d}s\\ &= \|\nabla w\|^2_{L^2(\Omega_R)^2} - k^2\,\|w\|^2_{L^2(\Omega_R)} - 2\pi R\sum_{\ell\in\mathbb{Z}}|\hat{w}_\ell|^2\underbrace{\Re\{T_\ell\}}_{\le 0}\\ &\ge \|\nabla w\|^2_{L^2(\Omega_R)^2} - k^2\,\|w\|^2_{L^2(\Omega_R)} = \|w\|^2_{H^1(\Omega_R)} - (k^2+1)\,\|w\|^2_{L^2(\Omega_R)},\end{aligned}$$

where we have expanded $(\gamma w)(\theta) = \sum_{\ell\in\mathbb{Z}}\hat{w}_\ell \mathrm{e}^{\mathrm{i}\ell\theta}$ on ∂B_R. Moreover, $H^1_{0,R}(\Omega_R)$ is compactly embedded in $L^2(\Omega_R)$. From Corollary 3.15 we have that if the homogeneous version of the variational problem (37) (find $u_0 \in H^1_{0,R}(\Omega_R)$ such that $\mathcal{A}_R(u_0,w) = 0$ for all $w \in H^1_{0,R}(\Omega_R)$) admits only the trivial solution $u_0 = 0$, then (37) is well-posed for any right-hand side.

4.4.4 Uniqueness

We first prove the following important result, [CK1, Thm. 3.12].

Theorem 4.25: (Rellich's lemma). If u is a radiating Helmholtz solution in Ω_+ then

$$\Im\int_{\partial B_R}\partial_{\mathbf{n}}u\;\gamma\overline{u}\,\mathrm{d}s \le 0 \qquad\Rightarrow\qquad u = 0 \quad \text{in } \Omega_+.$$

Proof. As before, we assume that u admits a Fourier–Hankel expansion $u = \sum_{\ell\in\mathbb{Z}}\hat{u}_\ell \mathrm{e}^{\mathrm{i}\ell\theta}H_\ell^{(1)}(kr)$, [CK1, Thm. 3.6]. With this normalisation, on ∂B_R we have

$$\begin{aligned}\Im\int_{\partial B_R}\partial_{\mathbf{n}}u\;\gamma\overline{u}\,\mathrm{d}s &= R\,\Im\int_0^{2\pi}\Big(\sum_{\ell\in\mathbb{Z}}\hat{u}_\ell\mathrm{e}^{\mathrm{i}\ell\theta}kH_\ell^{(1)\prime}(kr)\Big)\Big(\sum_{\ell\in\mathbb{Z}}\overline{\hat{u}_\ell}\mathrm{e}^{-\mathrm{i}\ell\theta}\overline{H_\ell^{(1)}(kr)}\Big)\,\mathrm{d}\theta\\ &= 2\pi R\sum_{\ell\in\mathbb{Z}}|\hat{u}_\ell|^2\Im\{kH_\ell^{(1)\prime}(kR)\overline{H_\ell^{(1)}}(kR)\}\\ &= 2\pi R\sum_{\ell\in\mathbb{Z}}|\hat{u}_\ell|^2 k\Big(Y'(kR)J(kR) - Y(kR)J'(kR)\Big) = 2\pi R\sum_{\ell\in\mathbb{Z}}|\hat{u}_\ell|^2\frac{2}{\pi R} = 4\sum_{\ell\in\mathbb{Z}}|\hat{u}_\ell|^2 \ge 0,\end{aligned}$$

from the Wronskian identity [DLMF, eq. 10.5.E2]. Since all terms in the series are positive, if the series is ≤ 0 then it is 0 and $\hat{u}_\ell = 0$ for all $\ell\in\mathbb{Z}$, so $u = 0$. □

If we choose u_0 to be the solution of the homogeneous variational problem (37) with $\mathcal{F}_R = 0$, then

$$0 = \Im\{\mathcal{F}_R(u_0)\} = \Im\{\mathcal{A}_R(u_0,u_0)\} = -\Im\int_{\partial B_R}(\mathsf{DtN}\gamma u_0)\gamma\overline{u}_0\,\mathrm{d}s = -\Im\int_{\partial B_R}\partial_{\mathbf{n}}u_0\;\gamma\overline{u}_0\,\mathrm{d}s$$

and $u_0 = 0$ by Rellich's lemma 4.25.

We can now complete the proof of the well-posedness. The homogeneous variational problem ((37) with $\mathcal{F}_R = 0$) admits only the trivial solution $u_0 = 0$ (§4.4.4) and its sesquilinear form satisfies a Gårding inequality (§4.4.3). By Corollary 3.15 the problem (37) is well-posed for all f and g_R. Thanks to the equivalence with (36), also the EDP (32) and the special case of the SSSP (33) are well-posed.

5 Boundary integral equations and the boundary element method

5.1 Single-layer potential, operator and the first BIE

[15]We define the 2D Helmholtz **fundamental solution**:

$$\boxed{\Phi_k(\mathbf{x},\mathbf{y}) := \frac{\mathrm{i}}{4}H_0^{(1)}(k|\mathbf{x}-\mathbf{y}|), \qquad \mathbf{x}\neq\mathbf{y}\in\mathbb{R}^2.} \tag{38}$$

For a given point $\mathbf{y}$, this is a radiating Fourier–Hankel function of order 0 centred at $\mathbf{y}$. It is a smooth Helmholtz solution in $\mathbb{R}^2\setminus\{\mathbf{y}\}$ and has a logarithmic singularity at $\mathbf{y}$. The roles of $\mathbf{x}$ and $\mathbf{y}$ are symmetric. It represents the field produced by a point source located in $\mathbf{y}$. The value of $\Phi_k(\mathbf{x},\mathbf{y})$ only depends on the distance $|\mathbf{x}-\mathbf{y}|$ between the arguments (more precisely: it depends on the number of wavelengths contained in that distance: $\frac{|\mathbf{x}-\mathbf{y}|}{\lambda} = \frac{k|\mathbf{x}-\mathbf{y}|}{2\pi}$).

We will see in Exercise 5.21 that the normalisation factor $\frac{\mathrm{i}}{4}$ in (38) gives that, for all $\mathbf{y}\in\mathbb{R}^2$, $\Delta\Phi_k(\cdot,\mathbf{y})+k^2\Phi_k(\cdot,\mathbf{y}) = \delta_{\mathbf{y}}$, in the sense of distributions, where $\delta_{\mathbf{y}}$ is the Dirac delta centred at $\mathbf{y}$. Moreover this coefficient will allow to write a simple Green's integral representation in §5.3.

Any linear combination $\sum_j \psi_j\Phi_k(\cdot,\mathbf{y}_j)$ of fundamental solutions centred at points $\mathbf{y}\in\overline{\Omega_-}$ satisfies the Helmholtz equation in Ω_+ and is radiating.[16] We can also take a continuous linear combination of fundamental solutions, which we write as

$$\boxed{(\mathcal{S}\psi)(\mathbf{x}) := \int_\Gamma \Phi_k(\mathbf{x},\mathbf{y})\psi(\mathbf{y})\,\mathrm{d}s(\mathbf{y}) \qquad \mathbf{x}\in\Omega_+,} \tag{39}$$

where ψ is a function on Γ. We can think at ψ as the density of acoustic sources[17] generating the field $\mathcal{S}\psi$. The function $\mathbf{x}\mapsto\Phi_k(\mathbf{x},\mathbf{y})\psi(\mathbf{y})$ belongs to $C^\infty(\Omega_+)$ for any given $\mathbf{y}\in\Gamma$. Thus, by the differentiation under integral sign theorem, the function $\mathcal{S}\psi$ belongs to $C^\infty(\Omega_+)$, is radiating and is a solution of the Helmholtz equation, [CGLS12, Thm. 2.14]. The operator $\mathcal{S}$ is called (acoustic) **single-layer potential** or, sometimes, **simple-layer potential**. It is possible to prove that the single-layer potential is continuous as a mapping $\mathcal{S}: H^{-\frac{1}{2}}(\partial\Omega)\to H^1_{\mathrm{loc}}(\overline{\Omega_+})$, [CGLS12, Thm. 2.15].

This suggests to look for a solution of the EDP (32) in the form $u(\mathbf{x}) = (\mathcal{S}\psi)(\mathbf{x})$ for some "density" ψ. But, how can we find ψ? We need to relate $\mathcal{S}\psi$ to the boundary condition.

We first introduce the **single-layer operator** S:

$$\boxed{(S\psi)(\mathbf{x}) := \int_\Gamma \Phi_k(\mathbf{x},\mathbf{y})\psi(\mathbf{y})\,\mathrm{d}s(\mathbf{y}) \qquad \mathbf{x}\in\Gamma.} \tag{40}$$

The only difference between the single-layer potential $\mathcal{S}$ and operator S is that the former is evaluated in points *off* the boundary, and the latter *on* the boundary Γ. When $\psi\in C^0(\Gamma)$, then the evaluation of $(\mathcal{S}\psi)(\mathbf{x})$ is the integral of a continuous function. On the other hand, no matter the regularity of ψ, the evaluation of $(S\psi)(\mathbf{x})$ is a singular integral, because of the singularity of $\Phi_k(\mathbf{x},\mathbf{y})$ at $\mathbf{x}=\mathbf{y}$. The single-layer operator is a first example of **boundary integral operator (BIO)**, in particular it is a **weakly singular integral operator**, as the singularity of Φ_k is logarithmic.

The Dirichlet trace operator relates $\mathcal{S}$ and S:

$$S\psi = \gamma^+(\mathcal{S}\psi). \tag{41}$$

This fact seems obvious from the definitions (39)–(40) of single-layer potential and operator, but its justification requires some care because of the singularity of the fundamental solution Φ_k, which appears in the

[15]This section closely follows [Sayas06, §3]. However we use the notation $\mathcal{S}, S$ of [Spence14], while [Sayas06] uses $\mathcal{S}_\Gamma, V_\Gamma$.

[16]This suggests a numerical method consisting in choosing N points $\mathbf{y}_1,\dots,\mathbf{y}_N\in\Omega_-$ and in searching the coefficients $\psi_1,\dots,\psi_N$ that minimise $\left\|\sum_{j=1}^N\psi_j\Phi_k(\cdot,\mathbf{y}_j)-g_D(\cdot)\right\|_{L^2(\Gamma)}$ (or some other norm on Γ). This is a well-known scheme called the "method of fundamental solutions" (MFS). It can give extremely good accuracy but has some drawbacks: it is very sensitive with respect to the location of the $\mathbf{y}_j$s, the minimisation generally leads to ill-conditioned linear systems, it struggles to approximate solutions for scatterers with corners.

[17]The terminology ("potential", "layer", "density", ...) comes from electrostatics, i.e. the Laplace equation case $k=0$. In this case ψ represents the surface density of electric charges generating, by Coulomb's law, the electrostatic potential $\mathcal{S}\psi$. For a simple and very brief summary see https://cims.nyu.edu/~oneil/courses/sp19-math2840/electrostatics.pdf

definition of $\mathcal{S}$ and S. [18] We will see in §5.4 (equation (57)) another boundary integral potential/operator pair defined by similar formulas, where the operator is not the trace of the potential.

If we are able to find ψ on Γ such that

$$\boxed{S\psi = g_D \quad \text{on } \Gamma,} \tag{42}$$

where g_D is the Dirichlet datum of the EDP (32), then

$$\boxed{u = \mathcal{S}\psi \quad \text{in } \Omega_+} \tag{43}$$

is a radiating Helmholtz solution in Ω_+ with $\gamma^+ u = \gamma^+ \mathcal{S}\psi = S\psi = g_D$ by (41), thus u itself is a **solution of the EDP** (32).

Equation (42) is the first example of **boundary integral equation** (BIE) and (43) is the corresponding **representation formula**. The unknown of the BIE is ψ, which is a distribution supported on Γ and does not need to have a physical meaning: for this reason this is called **indirect method**. If we could solve the BIE and compute ψ, then the solution u of the EDP could be obtained from the representation formula, which amounts to the computation of an integral on Γ for each point $\mathbf{x} \in \Omega_+$ where we want to evaluate u.

We will see in §6.1 that the BIE (42) is well-posed under some conditions on Γ and k.

From the continuity of the single-layer potential $\mathcal{S} : H^{-\frac{1}{2}}(\Gamma) \to H^1_{\mathrm{loc}}(\overline{\Omega_+})$, the trace formula (41) and the trace theorem 3.7, it follows that the single-layer operator is continuous as a mapping $S : H^{-\frac{1}{2}}(\Gamma) \to H^{\frac{1}{2}}(\Gamma)$.

Remark 5.1: (Continuity of the single-layer operator on the circle). We can verify the continuity $S : H^{-\frac{1}{2}}(\Gamma) \to H^{\frac{1}{2}}(\Gamma)$ for a circular boundary $\Gamma = \partial B_R = \{\mathbf{x} : |\mathbf{x}| = R\}$ for $R > 0$, where these norms can be computed from Fourier coefficients. Let $v(\theta) = \sum_{\ell\in\mathbb{Z}} \hat{v}_\ell \mathrm{e}^{\mathrm{i}\ell\theta}$ be a function (or distribution) defined on Γ and $\mathbf{x} = (R\cos\theta_{\mathbf{x}}, R\sin\theta_{\mathbf{x}}) \in \partial B_R$. By Graf's addition formula (30), the orthogonality of the circular harmonics, the large-index asymptotics [DLMF, 10.19.1–10.19.2] for Bessel and Hankel functions

$$\begin{aligned}(Sv)(\mathbf{x}) &= \int_{\partial B_R} \Phi_k(\mathbf{x},\mathbf{y}) v(\mathbf{y})\,\mathrm{d}s(\mathbf{y}) = R\int_0^{2\pi} \left(\sum_{\ell\in\mathbb{Z}} \frac{\mathrm{i}}{4} H_\ell^{(1)}(kR) J_\ell(kR) \mathrm{e}^{\mathrm{i}\ell\theta_{\mathbf{x}}} \mathrm{e}^{-\mathrm{i}\ell\theta_{\mathbf{y}}}\right)\left(\sum_{\ell\in\mathbb{Z}} \hat{v}_\ell \mathrm{e}^{\mathrm{i}\ell\theta_{\mathbf{y}}}\right) \mathrm{d}\theta_{\mathbf{y}} \\ &= 2\pi R \frac{\mathrm{i}}{4} H_\ell^{(1)}(kR) J_\ell(kR)\, \hat{v}_\ell\, \mathrm{e}^{\mathrm{i}\ell\theta_{\mathbf{x}}}\end{aligned}$$

[18] We show this (intuitively very plausible) fact when $\psi \in C^0(\Gamma)$ and Γ is a polygon.

Using [DLMF, eq. 10.8.E2] and the smoothness of J_0, one can show that $\Phi_k(\mathbf{x},\mathbf{y}) = -\frac{1}{2\pi}\log(k|\mathbf{x}-\mathbf{y}|) + \mathcal{R}(|\mathbf{x}-\mathbf{y}|)$ for a function $\mathcal{R} \in C^1(\mathbb{R})$. We denote by $L_{\mathcal{R}} := \sup_{t\in[0,\operatorname{diam}\Gamma]} |\mathcal{R}'(t)|$ the Lipschitz constant of $\mathcal{R}$.

We fix a point $\mathbf{x} \in \Gamma$ and a sequence $(\mathbf{x}_j)_{j\in\mathbb{N}} \subset \Omega_+ \cap B_{1/4}(\mathbf{x})$ with $\mathbf{x}_j \to \mathbf{x}$. We denote $\epsilon_j := 2|\mathbf{x}-\mathbf{x}_j|^{1/2}$. Then

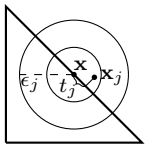
$$\begin{aligned}&|S\psi(\mathbf{x}) - \mathcal{S}\psi(\mathbf{x}_j)| = \left|\int_\Gamma [\Phi_k(\mathbf{x},\mathbf{y}) - \Phi_k(\mathbf{x}_j,\mathbf{y})]\psi(\mathbf{y})\,\mathrm{d}s(\mathbf{y})\right| \\ &\le \|\psi\|_{L^\infty(\Gamma)} \left(\int_\Gamma \left|\mathcal{R}(|\mathbf{x}-\mathbf{y}|) - \mathcal{R}(|\mathbf{x}_j-\mathbf{y}|)\right| \mathrm{d}s(\mathbf{y}) + \frac{1}{2\pi}\int_\Gamma \left|\log(k|\mathbf{x}-\mathbf{y}|) - \log(k|\mathbf{x}_j-\mathbf{y}|)\right| \mathrm{d}s(\mathbf{y}))\right) \\ &\le \|\psi\|_{L^\infty(\Gamma)} \left(|\Gamma| L_{\mathcal{R}} |\mathbf{x}-\mathbf{x}_j| + \int_{\Gamma\setminus B_{\epsilon_j}(\mathbf{x})} \left|\log|\mathbf{x}-\mathbf{y}| - \log|\mathbf{x}_j-\mathbf{y}|\right| \mathrm{d}s(\mathbf{y}) + \int_{\Gamma\cap B_{\epsilon_j}(\mathbf{x})} \left|\log|\mathbf{x}-\mathbf{y}| - \log|\mathbf{x}_j-\mathbf{y}|\right| \mathrm{d}s(\mathbf{y})\right).\end{aligned}$$

The first term clearly vanishes in the limit $\mathbf{x}_j \to \mathbf{x}$. Using that $|\log t - \log s| \le \frac{|t-s|}{\min\{t,s\}}$ for all $t, s > 0$ also the second term vanishes in the same limit:

$$\int_{\Gamma\setminus B_{\epsilon_j}(\mathbf{x})} \left|\log|\mathbf{x}-\mathbf{y}| - \log|\mathbf{x}_j-\mathbf{y}|\right| \mathrm{d}s(\mathbf{y}) \le \int_{\Gamma\setminus B_{\epsilon_j}(\mathbf{x})} \frac{\big||\mathbf{x}-\mathbf{y}| - |\mathbf{x}_j-\mathbf{y}|\big|}{\min\{|\mathbf{x}-\mathbf{y}|, |\mathbf{x}_j-\mathbf{y}|\}} \mathrm{d}s(\mathbf{y}) \le \frac{2|\Gamma|}{\epsilon_j}|\mathbf{x}-\mathbf{x}_j| = |\Gamma||\mathbf{x}-\mathbf{x}_j|^{1/2}.$$

Now assume that $\mathbf{x}$ belongs to a side of Γ (which was assumed to be a polygon) and that ϵ_j is smaller than the distance between $\mathbf{x}$ and the closest vertex of Γ (which is true for sufficiently large j). Let $t_j \in (0, \epsilon_j/2)$ be the distance between $\mathbf{x}$ and the point of Γ closest to $\mathbf{x}_j$. Then the last integral term can be bounded as

$$\begin{aligned}&\int_{\Gamma\cap B_{\epsilon_j}(\mathbf{x})} \left|\log|\mathbf{x}-\mathbf{y}| - \log|\mathbf{x}_j-\mathbf{y}|\right| \mathrm{d}s(\mathbf{y}) \\ &\le \int_{-\epsilon_j}^{\epsilon_j} \left|\log|t|\right| + \left|\log|t-t_j|\right| \mathrm{d}t \\ &= 2\epsilon_j|\log\epsilon_j - 1| + (\epsilon_j + t_j)|\log(\epsilon_j + t_j) - 1| + (\epsilon_j - t_j)|\log(\epsilon_j - t_j) - 1| \;\le\; C\epsilon_j|\log\epsilon_j|.\end{aligned}$$



(If $\mathbf{x}$ is instead a vertex of Γ one can proceed in a similar way treating separately the two segments of $\Gamma \cap B_{\epsilon_j}(\mathbf{x})$.) From the definition of ϵ_j, all three terms in the bound on $|\mathcal{S}\psi(\mathbf{x}_j) - S\psi(\mathbf{x})|$ converge to 0 for $j \to \infty$, so $\lim_{j\to\infty} \mathcal{S}\psi(\mathbf{x}_j) = S\psi(\mathbf{x})$. We have also shown that $\mathbf{x} \mapsto \int_\Gamma \Phi_k(\mathbf{x},\mathbf{y})\psi(\mathbf{y})\,\mathrm{d}s(\mathbf{y})$ is a continuous function on $\Omega_+ \cup \Gamma$ (actually it is continuous on $\mathbb{R}^2$). See [SS11, Thm. 3.3.5] for a more general proof.

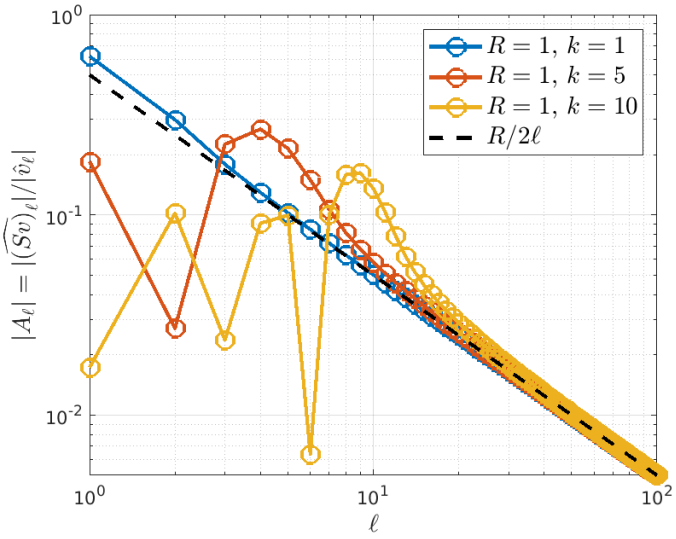


Figure 15: A log–log plot of the multipliers $A_\ell = \frac{\widehat{(Sv)}_\ell}{\hat{v}_\ell}$ *in the circular-harmonic expansion of the single-layer operator* S *on the boundary of a circle. They decay proportionally to* $\frac{1}{\ell}$*, demonstrating the continuity of* $S : H^s(\Gamma) \to H^{s+1}(\Gamma)$*. Each colour correspond to a wavenumber* k*. See Remark 5.1.*

$$\Rightarrow \qquad \widehat{(Sv)}_\ell = A_\ell \hat{v}_\ell, \qquad A_\ell := \frac{\pi R \mathrm{i}}{2} H_\ell^{(1)}(kR) J_\ell(kR) \sim \frac{R}{2\ell} \qquad \ell \to \infty.$$

We say that S "diagonalises" in the Fourier basis: expanding the argument v in the $\mathrm{e}^{\mathrm{i}\ell\theta}$ basis, the action of S corresponds to a multiplication of each coefficient $\hat{v}_\ell$ by a factor A_ℓ (exactly as when we multiply a diagonal matrix and a vector). In Figure 15 we show the log–log plot of the factors A_ℓ for $0 \le \ell \le 100$, $R = 1$ and different values of k. The important observation is that the coefficients decay as $A_\ell \sim \ell^{-1}$.

Recalling the definition of the fractional Sobolev norms (19) on the circle, this gives that $\|Sv\|_{H^{s+1}(\partial B_R)} \le C \|v\|_{H^s(\partial B_R)}$ for all $s \in \mathbb{R}$ and all $v \in H^s(\Gamma)$, or equivalently $S : H^s(\partial B_R) \to H^{s+1}(\partial B_R)$. This bound can be proved rigorously for all boundaries Γ and for a range of s that depends on the boundary regularity.

The continuity property $S : H^{-\frac{1}{2}}(\Gamma) \to H^{\frac{1}{2}}(\Gamma)$ lets us write a variational form of the BIE (42). Recall from §3.3.3 that $H^{\frac{1}{2}}(\Gamma)$ is the anti-dual of $H^{-\frac{1}{2}}(\Gamma)$. So we can seek $\psi \in H^{-\frac{1}{2}}(\Gamma)$ such that

$$\mathcal{A}(\psi, \xi) := \langle S\psi, \xi \rangle_{H^{\frac{1}{2}}(\Gamma) \times H^{-\frac{1}{2}}(\Gamma)} = \langle g_D, \xi \rangle_{H^{\frac{1}{2}}(\Gamma) \times H^{-\frac{1}{2}}(\Gamma)} =: \mathcal{F}(\xi) \qquad \forall \xi \in H^{-\frac{1}{2}}(\Gamma). \tag{44}$$

$\mathcal{A}(\cdot,\cdot)$ and $\mathcal{F}(\cdot)$ are the sesquilinear and the antilinear form of the variational problem. When $\xi \in L^2(\Gamma)$ the duality products $\langle \cdot, \cdot \rangle_\Gamma$ in (44) can be written as integrals over Γ. Recall that when we deal with BVPs for PDEs we obtain sesquilinear forms from integration by parts; with BIEs we simply multiply the equation by a test function and integrate (using the duality product) over Γ.

Exercise 5.2: (Quasi-self-adjointness of S). Formula [CK2, (3.84)] for the singularity of Φ_k implies that there is a constant $C > 0$ such that $\sup_{\mathbf{x},\mathbf{y} \in \Gamma} |\Phi_k(\mathbf{x},\mathbf{y}) - \frac{1}{2\pi} \log \frac{1}{|\mathbf{x}-\mathbf{y}|}| \le C$.

- (Difficult!) Deduce that the function $\Gamma \to \mathbb{R}$, $\mathbf{x} \mapsto \|\Phi_k(\mathbf{x}, \cdot)\|_{L^2(\Gamma)}$ is bounded.

  You can use the definition of Lipschitz domain.
- (Easier.) Show that the single-layer operator S is "quasi-self-adjoint" [CGLS12, p. 120], i.e. for $\psi, \phi \in L^2(\Gamma)$ it holds

  $$\int_\Gamma (S\psi) \overline{\phi} \, \mathrm{d}s(\mathbf{y}) = \int_\Gamma \psi (S\overline{\phi}) \, \mathrm{d}s(\mathbf{y}).$$

  Careful: to use Fubini theorem you need to verify that the integrand is integrable (L^1) on a suitable Cartesian product set.
- Show that S is not self-adjoint, i.e., in general, $\int_\Gamma (S\psi) \, \overline{\phi} \, \mathrm{d}s(\mathbf{y}) \ne \int_\Gamma \psi \, \overline{(S\phi)} \, \mathrm{d}s(\mathbf{y})$.

5.2 Piecewise-constant BEM for the single-layer BIE

We have seen that if we were able to find a solution ψ to the BIE (42) then we would have a solution $u = \mathcal{S}\psi$ of the EDP. In general we cannot solve the BIE analytically, thus we resort to a numerical method.

The **boundary element method** (BEM) consists of choosing an N-dimensional space $V_N \subset H^{-\frac{1}{2}}(\Gamma)$ and looking for a $\psi_N \in V_N$ that approximately solves the BIE (42). There are two ways of imposing the BIE.

- **Collocation-BEM**. We choose N points $\mathbf{x}_1, \dots, \mathbf{x}_N$ on Γ and look for

$$\psi_N \in V_N \text{ such that } (S\psi_N)(\mathbf{x}_j) = g_D(\mathbf{x}_j), \quad j = 1, \dots, N.$$

- **Galerkin-BEM**. We restrict the variational form (44) to the finite-dimensional space V_N, as in (23). In practice, all discrete functions we may want to consider belong to $L^2(\Gamma)$, so we can write the Galerkin-BEM as: find $\psi_N \in V_N$ such that

$$\mathcal{A}(\psi_N, \xi_N) = \int_\Gamma (S\psi_N)\overline{\xi_N}\,\mathrm{d}s = \int_\Gamma g_D\overline{\xi_N}\,\mathrm{d}s = \mathcal{F}(\xi_N) \qquad \forall \xi_N \in V_N.$$

We recall that $H^{-\frac{1}{2}}(\Gamma)$ is a space larger than $L^2(\Gamma)$, so it accommodates discontinuous functions. This makes the construction of the discrete space V_N simpler. The simplest choice of V_N is the following: we partition the curve Γ in a mesh $\mathcal{T}_N(\Gamma)$ of N (possibly curvilinear) segments $K_1, \dots, K_N \subset \Gamma$ (with $\bigcup_{j=1}^N \overline{K_j} = \Gamma$ and $K_j \cap K_{j'} = \emptyset$ for $j \neq j'$) and choose V_N to be the space of **piecewise constant** functions on $\mathcal{T}_N(\Gamma)$.

The obvious **basis** $\{\varphi_j\}_{j=1}^N$ of V_N is defined by $\varphi_j(\mathbf{x}) = 1$ if $\mathbf{x} \in K_j$ and $\varphi_j(\mathbf{x}) = 0$ if $\mathbf{x} \in \Gamma \setminus K_j$.

We expand the BEM solution in coordinates as $\boxed{\psi_N = \sum_{j=1}^N \Psi_j \varphi_j}$, where $\Psi_j = (\boldsymbol{\Psi})_j$ is the jth element of the vector $\boldsymbol{\Psi} \in \mathbb{C}^n$. So $\psi_N(\mathbf{x}) = \Psi_j$ if $\mathbf{x} \in K_j$.

With this discrete space and basis, in the collocation-BEM we choose the **collocation nodes** $\mathbf{x}_j$ such that $\mathbf{x}_j \in K_j$ for all $j = 1, \dots, N$ and obtain the linear system $\boxed{\underline{\underline{\mathbf{A}}}^{\mathrm{Col}}\boldsymbol{\Psi} = \mathbf{F}^{\mathrm{Col}}}$, where

$$\boxed{A^{\mathrm{Col}}_{j,m} := (S\varphi_m)(\mathbf{x}_j) = \int_\Gamma \Phi_k(\mathbf{x}_j, \mathbf{y})\varphi_m(\mathbf{y})\,\mathrm{d}s(\mathbf{y}) = \int_{K_m} \Phi_k(\mathbf{x}_j, \mathbf{y})\,\mathrm{d}s(\mathbf{y}), \qquad F^{\mathrm{Col}}_j := g_D(\mathbf{x}_j).} \tag{45}$$

Indeed, by the linearity of S, if ψ_N is the collocation-BEM solution then its coefficient vector $\boldsymbol{\Psi}$ solves $\underline{\underline{\mathbf{A}}}^{\mathrm{Col}}\boldsymbol{\Psi} = \mathbf{F}^{\mathrm{Col}}$:

$$F^{\mathrm{Col}}_j = g_D(\mathbf{x}_j) = (S\psi_N)(\mathbf{x}_j) = \Big(S\sum_{m=1}^N \Psi_m\varphi_m\Big)(\mathbf{x}_j) = \sum_{m=1}^N \Psi_m (S\varphi_m)(\mathbf{x}_j) = (\underline{\underline{\mathbf{A}}}^{\mathrm{Col}}\boldsymbol{\Psi})_j.$$

To be able to evaluate the collocation-BEM right-hand side vector $\mathbf{F}^{\mathrm{Col}}$ we need $g_D \in C^0(\Gamma)$, which is typically satisfied if the EDP comes from a scattering problem such as (33).

Similarly, if ψ_N is the Galerkin-BEM solution, its coefficient vector $\boldsymbol{\Psi}$ solves $\boxed{\underline{\underline{\mathbf{A}}}^{\mathrm{Gal}}\boldsymbol{\Psi} = \mathbf{F}^{\mathrm{Gal}}}$ with

$$\boxed{\begin{aligned} A^{\mathrm{Gal}}_{j,m} &:= \mathcal{A}(\varphi_m, \varphi_j) = \langle S\varphi_m, \varphi_j\rangle_\Gamma = \int_\Gamma (S\varphi_m)(\mathbf{x})\overline{\varphi_j}(\mathbf{x})\,\mathrm{d}s(\mathbf{x}) = \int_\Gamma\int_\Gamma \Phi_k(\mathbf{x},\mathbf{y})\varphi_m(\mathbf{y})\overline{\varphi_j}(\mathbf{x})\,\mathrm{d}s(\mathbf{y})\,\mathrm{d}s(\mathbf{x}) \\ F^{\mathrm{Gal}}_j &:= \mathcal{F}(\varphi_j) = \langle g_D, \varphi_j\rangle_\Gamma = \int_\Gamma g_D(\mathbf{x})\overline{\varphi_j}(\mathbf{x})\,\mathrm{d}s(\mathbf{x}). \end{aligned}} \tag{46}$$

Choosing piecewise-constant basis functions on the elements K_j, these expressions simplify to

$$\boxed{A^{\mathrm{Gal}}_{j,m} = \int_{K_j}\int_{K_m} \Phi_k(\mathbf{x},\mathbf{y})\,\mathrm{d}s(\mathbf{y})\,\mathrm{d}s(\mathbf{x}), \qquad F^{\mathrm{Gal}}_j = \int_{K_j} g_D(\mathbf{x})\,\mathrm{d}s(\mathbf{x}).}$$

Multiplying each row of $\underline{\underline{\mathbf{A}}}^{\mathrm{Col}}$ and $\mathbf{F}^{\mathrm{Col}}$ by the length of the corresponding element, we see that the collocation-BEM is a Galerkin-BEM with a simple 1-point quadrature. Formally, one can also think at the collocation method as a Petrov–Galerkin method with delta functions as test functions, but this is not mathematically correct in the $H^{\pm\frac{1}{2}}(\Gamma)$ framework because delta functions do not belong to $H^{-\frac{1}{2}}(\Gamma)$.

Exercise 5.3: (Testing against constants is taking means). Show that the Galerkin-BEM with piecewise-constant functions is equivalent to impose that the *integral averages* of $S\psi_N$ and g_D coincide on each element.

Both matrices $\underline{\underline{\mathbf{A}}}^{\mathrm{Col}}$ and $\underline{\underline{\mathbf{A}}}^{\mathrm{Gal}}$ are **dense**: this is a major difference between the BEM and the finite element method (FEM). The Galerkin matrix is also complex-symmetric, but not Hermitian.

This shortcoming of the BEM with respect to the FEM is compensated by a **dimensional reduction**: to solve a 2D problem we only need to mesh a 1D object, the boundary Γ. Thus typically BEM requires much fewer degrees of freedom (DOFs) than FEM for comparable problems and accuracies.

Another advantage of BEM is that it deals with a BVP posed on the unbounded domain Ω_+ by discretising only a bounded object, Γ. To treat the EDP with FEM one has to truncate Ω_+ as in Remark 4.20, introducing additional errors, [Ihl98, §3].

The collocation-BEM is simpler to implement than the Galerkin-BEM. However, in many situations the choice of the collocation nodes adversely affects the performance of the method. To improve the numerical stability of the collocation method, often one chooses M collocation nodes with $M > N$ (oversampling) and solves an overdetermined rectangular linear system in the least-squares sense. If the datum g_D is discontinuous (which is possible since $H^{\frac{1}{2}}(\Gamma) \not\subset C^0(\Gamma)$), then the computation of $\mathbf{F}^{\mathrm{Col}}$ is not well-defined; on the other hand, for typical scattering problems g_D is continuous and piecewise smooth. Moreover, the stability and convergence theory for the Galerkin-BEM is much more complete.

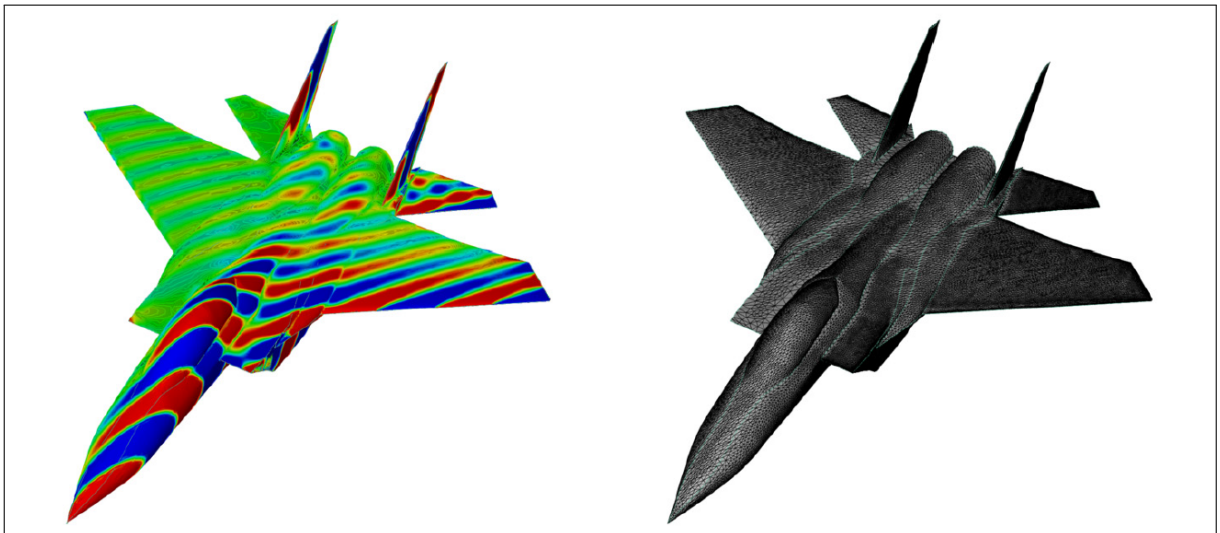


Figure 16: Scattering of a plane wave by an F15 figther aircraft, from [Chaillat, Groth, Loseille, 2018]. Left: the solution of (the 3d version of) the single-layer BIE (42) approximated with a piecewise-linear collocation-BEM. Right: the mesh used, with more than 160 000 triangles.

Exercise 5.4: (Least-squares BEM). A third version of the BEM, after collocation and Galerkin, is a least-squares formulation. In this case, one looks for $\psi_N \in V_N$ such that $\int_\Gamma (S\psi_N)\overline{(S\xi_N)}\,\mathrm{d}s = \int_\Gamma g_D\overline{(S\xi_N)}\,\mathrm{d}s$ for all $\xi_N \in V_N$. Compute the entries of the matrix $\underline{\underline{\mathbf{A}}}^{\mathrm{LS}}$ and the right-hand side $\mathbf{F}^{\mathrm{LS}}$ of the corresponding linear system. How many integrations are needed for each of them? Show that $\underline{\underline{\mathbf{A}}}^{\mathrm{LS}}$ is Hermitian and, if (42) is uniquely solvable, positive definite.

5.2.1 BEM and quadrature

From (45) and (46) we see that to compute each entry of the system matrix we need to compute an **integral of the fundamental solution**: it is a single integral on a mesh element for the collocation-BEM and a double integral on the Cartesian product of two elements for the Galerkin-BEM. To compute each entry of the right-hand side vectors, in the collocation-BEM we only need to evaluate the boundary datum g_D while for the Galerkin-BEM we need an integral over an element.

All these integrals require accurate **quadrature** formulas: these are among the main difficulties in a BEM implementation. In particular, for both matrices, the **diagonal entries require the approximation of singular integrals**, because of the (logarithmic) singularity of $\Phi_k(\mathbf{x},\mathbf{y})$ at $\mathbf{x}=\mathbf{y}$.

Let us assume that Ω_- is a connected **polygon** and each mesh element is a **straight segment**. For $j = 1,\dots,N$, the element K_j has endpoints $\mathbf{p}_j$ and $\mathbf{p}_{j+1}$ and length $h_j := |\mathbf{p}_{j+1} - \mathbf{p}_j|$ (of course

$\mathbf{p}_{N+1} = \mathbf{p}_1$). The element is parametrised by $\mathbf{X}_j : (0, h_j) \to K_j$, $\mathbf{X}_j(s) := \mathbf{p}_j + s\boldsymbol{\tau}_j$, where $\boldsymbol{\tau}_j := \frac{\mathbf{p}_{j+1}-\mathbf{p}_j}{|\mathbf{p}_{j+1}-\mathbf{p}_j|}$ is the unit tangent vector to K_j.

Recalling the definition of the fundamental solution (38), the entries of the BEM matrices and vectors are then computed as integrals over intervals and rectangles:

$$A^{\mathrm{Col}}_{j,m} = \frac{\mathrm{i}}{4}\int_0^{h_m} H_0^{(1)}(k|\mathbf{p}_m + s\boldsymbol{\tau}_m - \mathbf{x}_j|)\,\mathrm{d}s, \qquad F^{\mathrm{Col}}_j = g_D(\mathbf{x}_j), \qquad F^{\mathrm{Gal}}_j = \int_0^{h_j} g_D(\mathbf{p}_j + s\boldsymbol{\tau}_j)\,\mathrm{d}s,$$

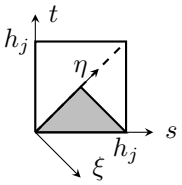
$$A^{\mathrm{Gal}}_{j,m} = \int_0^{h_j}\left(\int_0^{h_m} \Phi_k(\mathbf{p}_j + t\boldsymbol{\tau}_j, \mathbf{p}_m + s\boldsymbol{\tau}_m)\,\mathrm{d}s\right)\mathrm{d}t = \frac{\mathrm{i}}{4}\int_0^{h_j}\left(\int_0^{h_m} H_0^{(1)}\big(k|\mathbf{p}_j - \mathbf{p}_m + t\boldsymbol{\tau}_j - s\boldsymbol{\tau}_m|\big)\,\mathrm{d}s\right)\mathrm{d}t.$$

Each diagonal entry of the collocation-BEM matrix is a singular integral, as $\mathbf{x}_j \in K_j$ for all j. A simple recipe to compute $A^{\mathrm{Col}}_{j,j}$ is to split the element in the two components of $K_j \setminus \{\mathbf{x}_j\}$ and apply Gauss quadrature on each side. To this purpose a Gauss–Legendre quadrature is a good choice, while Gauss–Lobatto rules cannot be used as they involve the value of the integrand at the endpoints, where the singularity is located. The obvious choice for the collocation nodes is to take the element midpoints: $\mathbf{x}_j = \frac{1}{2}(\mathbf{p}_j + \mathbf{p}_{j+1})$.

The jth diagonal entry of the Galerkin-BEM matrix is a double integral on the square $(0, h_j)^2 = \{0 < s, t < h_j\}$, whose integrand has a singularity along the diagonal $s = t$. One can use a quadrature formula for triangles (e.g. based on Duffy transform) on each half of the square. Alternatively, splitting the square in four triangle, exploiting the symmetries, and using the isometric change of variables $\xi = \frac{s-t}{\sqrt{2}}$, $\eta = \frac{s+t}{\sqrt{2}}$ ($s = \frac{\xi+\eta}{\sqrt{2}}$, $t = \frac{\eta-\xi}{\sqrt{2}}$), we obtain

$$A^{\mathrm{Gal}}_{j,j} = \int_0^{h_j}\int_0^{h_j} \frac{\mathrm{i}}{4}H_0^{(1)}\big(k|s-t|\big)\,\mathrm{d}s\,\mathrm{d}t = 4\int_0^{\frac{h_j}{\sqrt{2}}}\left(\int_\xi^{\frac{h_j}{\sqrt{2}}} \frac{\mathrm{i}}{4}H_0^{(1)}(\sqrt{2}k\xi)\,\mathrm{d}\eta\right)\mathrm{d}\xi$$

$$= \int_0^{\frac{h_j}{\sqrt{2}}} \mathrm{i}\Big(\frac{h_j}{\sqrt{2}} - \xi\Big)H_0^{(1)}(\sqrt{2}k\xi)\,\mathrm{d}\xi \overset{(h_j\zeta=\sqrt{2}\xi)}{=} \frac{\mathrm{i}h_j^2}{2}\int_0^1 (1-\zeta)H_0^{(1)}(kh_j\zeta)\,\mathrm{d}\zeta.$$

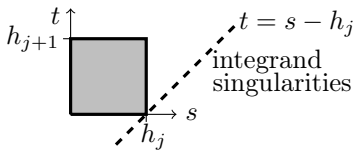


This is a one-dimensional integral with a weak singularity at the endpoint $\zeta = 0$. Its approximation with a Gauss–Legendre quadrature rule gives good results.

If all elements are identical, e.g. straight segments of the same length, then also the diagonal terms are identical and they need to be computed only once: $h_j = h_{j'} \Rightarrow A_{j,j} = A_{j',j'}$.

In the Galerkin-BEM, also when two element share an endpoint we have a singularity. For example, if the segments K_j and K_{j+1} are aligned ($\boldsymbol{\tau}_j = \boldsymbol{\tau}_{j+1}$), we have

$$A^{\mathrm{Gal}}_{j,j+1} = \int_0^{h_j}\left(\int_0^{h_{j+1}} \frac{\mathrm{i}}{4}H_0^{(1)}\big(k|h_j - s + t|\big)\,\mathrm{d}t\right)\mathrm{d}s.$$



This is a double integral on the rectangle $\{0 < s < h_j, 0 < t < h_{j+1}\}$ with a logarithmic singularity at the vertex $s = h_j$, $t = 0$. Again, since the since the weak singularity is at the boundary of the domain of integration, Gauss rules can be used.

Exercise 5.5: (BEM on non-polygonal Γ). Write the entries of the collocation- and Galerkin-BEM when Ω_- is not a polygon but a general Lipschitz domain and its boundary is defined by a parametrisation $\mathbf{X} : [0, L] \to \Gamma$.

Remark 5.6: (Singularity extraction quadrature). A typical technique to compute the singular integrals is the "singularity extraction". The small-argument asymptotics of the Hankel function ($H_0^{(1)}(z) \sim \frac{2\mathrm{i}}{\pi}\log z$ for $z \searrow 0$) give $\Phi_k(\mathbf{x}, \mathbf{y}) = -\frac{1}{2\pi}\log|\mathbf{x} - \mathbf{y}| + \mathcal{R}(\mathbf{x}, \mathbf{y})$ for a reminder $\mathcal{R}$ of class C^1. Inserting this in the expression of $A^{\mathrm{Col/Gal}}_{j,j}$ one can compute analytically the terms coming from the log and use a standard quadrature for the remainder $\mathcal{R}$.

Once we have assembled and solved the BEM linear system, we have an approximation $\psi_N \in V_N$ of the solution $\psi \in H^{-\frac{1}{2}}(\Gamma)$ of the BIE (42). However to approximate the solution of the EDP (32)/SSSP (33), we need to approximate u in the unbounded domain Ω_+. Recalling the representation formula $u = \mathcal{S}\psi$ (43), the BEM approximation of u is

$$u_N(\mathbf{x}) := (\mathcal{S}\psi_N)(\mathbf{x}) = \int_\Gamma \Phi_k(\mathbf{x}, \mathbf{y})\psi_N(\mathbf{y})\,\mathrm{d}s(\mathbf{y}) = \sum_{j=1}^N \Psi_j \int_{K_j} \Phi_k(\mathbf{x}, \mathbf{y})\,\mathrm{d}s(\mathbf{y}) \qquad \mathbf{x} \in \Omega_+.$$

Again, each term in this sum is an integral that needs to be approximated with a quadrature formula. For all $\mathbf{x} \in \Omega_+$ the integrand is C^∞. However, if $\mathbf{x}$ lies very close to Γ the accurate evaluation of $u_N(\mathbf{x})$ requires a careful use of the quadrature as the integral is near-singular.

The function u_N computed with the BEM is an exact solution of the Helmholtz equation and exactly satisfies the Sommerfeld radiation condition. It is an approximation of the solution u of the exterior Dirichlet problem (32) only because it satisfies the boundary condition on Γ approximately.

Remark 5.7: (Resolution of oscillations). Another difficulty is due to the **oscillatory behaviour** of both Φ_k and the solution for large values of the wavenumber k.

A first issue is that to approximate the solution one needs to use more DOFs for larger values of k. See Figure 17 for a simple visual justification of this statement. A typical recipe for "engineering accuracy" (a few percent relative errors) is to use at least 10 DOFs per wavelength λ. This means that the length of each element should not exceed $\frac{\lambda}{10} = \frac{\pi}{5k}$. (Sometimes the rule of thumb is to use 6 DOFs per wavelength, corresponding roughly to $kh_j \leq 1$, for $10\%/15\%$ error.) This implies that the number of DOFs must grow like $N = \mathcal{O}(k)$ for increasing k: high-frequency problem are computationally very expensive. On the other hand, a 2D FEM needs at least $\mathcal{O}(k^2)$ DOFs as the wavelength has to be resolved in two dimensions, which is a much stronger requirement; the difficulties encountered by FEM in solving high-frequency problems are described e.g. in [Ihl98].

The fundamental solution oscillates with wavelength close to $\lambda = \frac{2\pi}{k}$. Thus, even for elements K_j, K_m far from each other, in the assembly of the matrices $\underline{\underline{\mathbf{A}}}^{C/G}$ one has to use a sufficiently accurate quadrature formula to take into account the oscillations.

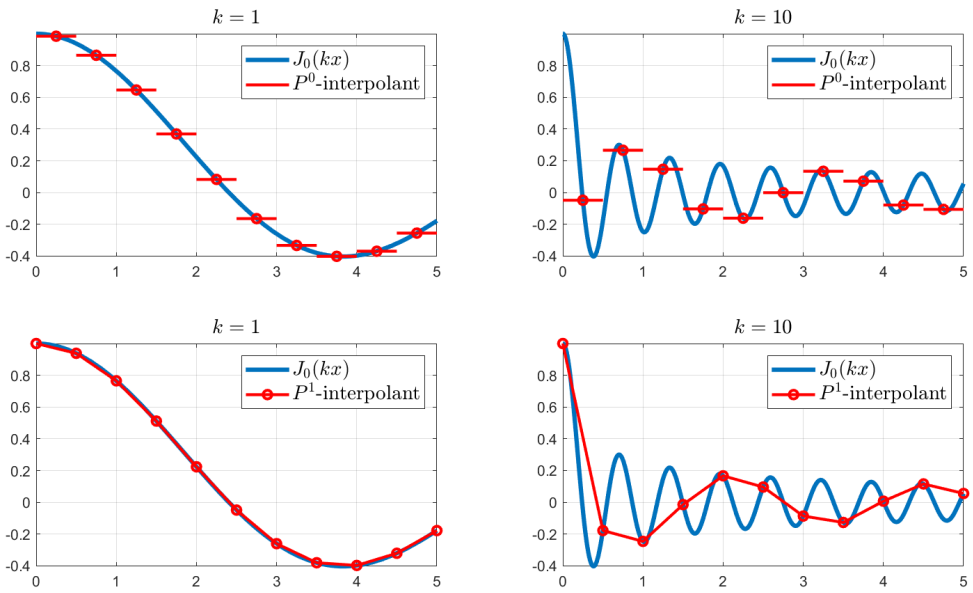


Figure 17: The piecewise-constant (top) and the continuous piecewise-linear (bottom) interpolants of the Bessel function $J_0(kx)$ for different wavenumbers $k = 1$ (left) and $k = 10$ (right). In all examples, the domain $(0, 5)$ is divided in a uniform mesh made of 10 segments. It is clear that at low frequencies ($k = 1$) the interpolant is able to describe the shape of the function, while for high frequencies ($k = 10$) it completely misses it. This tells us that when solving problems at higher frequencies with the BEM we need to refine the mesh (at least) proportionally to k, see Remark 5.7.

Obvious improvements of the piecewise-constant BEM use piecewise-polynomial discrete spaces of higher order. They can be discontinuous, or $C^0(\Gamma)$ or of higher continuity (at least away from corners of Γ) such as splines. The choice of basis functions, collocation points and quadrature rules is in general non-trivial. When Γ is smooth one can use also global functions such as mapped trigonometric functions, in the spirit of spectral methods. If Γ is a polygon (possibly with curvilinear sides), the BIE solution is usually singular close to its corners: to approximate it efficiently one can use a graded mesh, i.e. a mesh whose elements are smaller the closer they are to a corner.

5.2.2 BEM CODING PROJECT

Implement the collocation-BEM method for the scattering of a plane wave by a polygon Ω_-.
Use a discrete space V_N of piecewise-constant functions. Choose a mesh such that the elements on a given side of Γ have equal length. Plot the scattered field u_N and the total field on a portion of Ω_+.
You can use the quadrature routine provided.

A possible suggestion for the main steps in the code:

1. **Initialise** the geometric data structure. Given the vertices of the polygon, decide the number of elements on each side, and generate (for each element K_j) the endpoint $\mathbf{p}_j$, the length h_j, the tangent vector $\boldsymbol{\tau}_j$, and the collocation point $\mathbf{x}_j$, chosen as the element midpoint. Choose elements with roughly the same size.

   (In Matlab it might be convenient to treat points and vectors in the plane as complex numbers with the usual identification $\mathbb{C} \sim \mathbb{R}^2$; then the `abs` function allows for immediate computation of distances.)

2. **Assemble** the matrix $\underline{\underline{\mathbf{A}}}^{\mathrm{Col}}$ (being careful with the quadrature) and the right-hand side $\mathbf{F}^{\mathrm{Col}}$.

   You can use the Gauss–Legendre quadrature routine provided.

3. **Solve** the linear system.

4. **Evaluate** the near-field u_N on a grid of points in 2D using the representation formula and plot it.

   To generate the grid of points for the plot use `meshgrid`. To plot the field you can use one of the Matlab commands `pcolor` (used for Figure 12), `surf`, `mesh` or `contour`. To hide the grid points that lie inside Ω_- you can locate them with `inpolygon` and set them to 0 or `NaN` (or use the command `patch`).

   Plot the total field $u_N^{\mathrm{Tot}} = u_N + u^{\mathrm{Inc}}$: if it doesn't vanish approaching Γ then the code is not correct.

Plenty of interesting extensions are possible:

- Choose as incident wave u^{Inc} a **fundamental solution** centred in some $\mathbf{x}_0 \in \Omega_+$. (The total field cannot be plotted close to $\mathbf{x}_0$.) You can also try more exotic incident waves, such as Herglotz functions.
- Implement the problem of the scattering by **multiple polygons**, i.e. with Ω_- made of several components. Here Matlab's "cell arrays" are useful to store vectors with different lengths in a single variable, e.g., the coordinates of the vertices of polygons with different numbers of sides.
- Implement the **Galerkin-BEM** for the same problem, on the same mesh. The only difference is in the assembly of the matrix and the right-hand side. Compare the solutions obtained with the two versions of the scheme.
- Generate a time-harmonic **animation** of the scattered and total fields.

  You can generate a .gif file using the script provided or export a video in various formats such as .avi.
- Implement the **singularity-extraction** quadrature of Remark 5.6.
- Use meshes that are **locally refined** towards the corners of Γ to better approximate the solution singularities.
- Estimate and plot the Helmholtz **Poynting vector** $\mathbf{S}(u^{\mathrm{Scat}})$ defined in Exercise 2.13 (see Figure 13).

  The Matlab command `quiver` may help. Approximate the gradient ∇u^{Scat} with finite differences.
- Study the dependence of the results on the **parameters**. You can see how the plots and the norms of the solution vary when you change the number N of DOFs, the wavenumber k, the product kh, the quality of the quadrature, the shape of Ω_-...

  You can approximate the **error** committed by the scheme e.g. as $\|\psi_N - \psi_{\mathrm{ref}}\|_{L^2(\Gamma)}$ or $\|u_N - u_{\mathrm{ref}}\|_{L^2(\Omega_*)}$ where Ω_* is (a subset of) the portion of Ω_+ where you plot the near-field and $\psi_{\mathrm{ref}}/u_{\mathrm{ref}}$ are the reference solutions obtained with a very fine mesh. You can then plot the convergence of the error against N or study its dependence on the other parameters. For instance: for increasing k, if the mesh is refined keeping kh constant, how does the error behave?

  Another way to estimate the accuracy of the numerical solution is provided by Remark 5.27.

  How does the condition number depends on N?

- Test the code (for a square scatterer) by computing the near-field error against the **reference solution** provided.[19] This was computed with MPSpack (`https://github.com/ahbarnett/mpspack`).
- Implement the collocation-BEM with **oversampling**, i.e. with M collocation points, N degrees of freedom, and $M > N$. For instance take $M = cN$ for some fixed $c > 1$. Solve the rectangular linear system in the least-square sense (with the backslash command). Does the accuracy of the method improve? How does the cost increase? How does the position (equispaced, random,...) of the collocation points impact the results?
- Solve the BEM linear system using an appropriate **iterative method**, e.g. GMRES. Study its performance (number of iterations to reach a given accuracy, computing time...) in dependence of the problem parameters (N, k, the shape of Ω_-,...).
- Extend the code to **curvilinear** polygons and/or smooth scatterers.

  The use of the arclength parametrisation of the boundary makes the implementation simpler. Make sure that the quadrature points lie on the curve Γ.

  You can test your code for a circular obstacle against the analytic solution obtained in Exercise 4.15.
- Implement a **higher-order** version of the BEM with ($C^0(\Gamma)$) piecewise-polynomial basis functions of degree $p > 0$.
- Extend the code to discrete spaces made of continuous **splines** of maximal regularity, i.e. piecewise-polynomials of degree p with continuity $C^{p-1}(\Gamma)$. In the collocation version of the scheme, for even p choose the collocation nodes as the element mid-points (as we have done for $p = 0$), for odd p choose them as the element endpoints. With this choice, the collocation BEM is well-posed and quasi-optimal for sufficiently small h; see [Arnold, Wendland, 1985].
- Implement a **spectral BEM** on a smooth curvilinear scatterer.

  Choose a smooth scatterer defined by a parametrisation $\mathbf{X} : [0, 2\pi] \to \Gamma$. E.g. a shape often used as example for scattering problems is the kite $\mathbf{X}(t) = (\cos t + 0.65(\cos 2t - 1),\ 1.5 \sin t)$, [CK2, Fig. 3.1]. For $L \in \mathbb{N}$ and $N = 2L + 1$, choose as basis functions the mapped complex exponentials $\varphi_\ell(\mathbf{x}) = \mathrm{e}^{\mathrm{i}\ell\mathbf{X}^{-1}(\mathbf{x})}$, $-L \le \ell \le L$, or the corresponding trigonometric functions $\sin(\ell\mathbf{X}^{-1}(\mathbf{x})), \cos(\ell\mathbf{X}^{-1}(\mathbf{x}))$.
- ...

Exercise 5.8: (Build your own quasi-resonance). When a plane wave impinges on a scatterer Ω_- with a simple shape, for example convex or star-shaped, it is possible to prove that some norm of the scattered field is stable (does not grow) for $k \to \infty$. On the contrary, if Ω_- has a cavity (part of Ω_+, which is connected) that can "trap" a wave, the scattered field in the cavity can be extremely large.

Using your BEM code, try to find a polygonal scatterer that traps an incoming plane wave. Recall the discussion of Dirichlet eigenfunctions (which exist only for bounded regions) in §4.2. With a simple polygon with 12 sides it is not difficult to obtain a scattered field with magnitude $|u^{\mathrm{Scat}}| > 30$ in some region (with $\|u^{\mathrm{Inc}}\|_{L^\infty(\mathbb{R}^2)} = 1$).

Careful: for a given domain only some wavenumbers will do. The computations in §4.2 might help you choosing the right k for your trapping domain.

Exercise 5.9: (BEM vs FEM). Recall the approximation of the SSSP/EDP described in Remark 4.20 and Exercise 5.14: Ω_+ is truncated to $\Omega_R = B_R \setminus \overline{\Omega_-}$ and the homogeneous impedance boundary condition $\partial_{\mathbf{n}} u^{\mathrm{Scat}} - \mathrm{i}k u^{\mathrm{Scat}} = 0$ is imposed on ∂B_R.

Use the 2D FEM code you have implemented in the finite element course to discretise this BVP. Careful: FEM basis functions (e.g. tent functions) are real, but the right-hand side vector and the coefficients of the discrete solution are complex. Compare the results against those obtained with BEM.

Extra: Instead of the impedance condition, impose the DtN map (35) using circular harmonics. This is harder and involves the assembly of a dense block in the matrix.

Remark 5.10: (Fast Matlab code). Often, most of the computational time in a simple BEM code is spent evaluating the fundamental solution for the assembly of $\underline{\underline{\mathbf{A}}}^{\mathrm{Col}}$ and the evaluation of u_N. You can check how much time is spent on each part of your Matlab code using the `profile` command, for instance

[19] Load the .mat file provided. The vectors `gx`, `gy` contain the coordinates of the nodes of a Cartesian grid in the square $(-1.5, 1.5)^2$. The matrix `ui` contains the values of an incoming plane wave with direction $\mathbf{d} = \frac{1}{\sqrt{2}}(1, -1)$ and wavenumber $k = 20$, in the the grid points exterior to the unit circle, and 0 otherwise. The entries of the matrix `u` are: the values of the scattered field $u = u^{\mathrm{Scat}}$ in the grid points exterior to the unit circle, the values of the total field u^{Tot} in the points interior to the unit circle but exterior to the scatterer $(-0.5, 0.5)^2$, and not-a-number `nan` values inside the scatterer. You can plot the total field with the commands

```
load MPSpackBenchmarkSquareScatt.mat; pcolor(gx,gy,real(ui+u)); shading flat; axis square;
```

```
profile on;    MyFunctionDoingWhatever;    profile viewer
```

A simple implementation of the piecewise-constant collocation-BEM assembles the matrix $\underline{\underline{\mathbf{A}}}^{\mathrm{Col}}$ using a triple loop: over rows, columns and quadrature points. This requires $N^2 q$ calls to the function `besselh`, where q is the number of quadrature nodes per element (ignoring the different treatment needed for the diagonal terms). However, we know that Matlab is most efficient if we reduce as much as possible the use of nested loops and apply functions such as `besselh` to few vectors (as opposed to many scalars). You can observe this with

```
n = 5000; A=rand(n);
B=ones(n)*1i; C=ones(n)*1i;      % Initialise complex matrices
tic; B=besselh(0,1,A); toc
tic; for j=1:n; for jj=1:n; C(j,jj)=besselh(0,1,A(j,jj)); end; end; toc
```

Keep this in mind to write a fast code: avoid loops and minimise the number of calls to the fundamental solution.

5.3 Green's integral representation

The Green's integral representation, or Green's third identity, is an important tool to derive new BIEs and BEMs, and to understand the properties of the corresponding BVPs. It extends Green's second identity (21). We write it and prove it for bounded and unbounded Lipschitz domains.

We keep using the notation introduced in §4.3. In particular, we recall that the unit normal $\mathbf{n}$ on the boundary of the bounded domain Ω_- is defined to point outwards, into the complement Ω_+; this enters the definition of the Neumann trace $\partial_{\mathbf{n}}$. Since the fundamental solution Φ_k depends on two variables, we write $\frac{\partial \Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{x})}$ and $\frac{\partial \Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{y})}$ to make clear which is the variable with respect to which we derive and take the trace. We recall that when we write $\int_\Gamma \varphi\overline{\psi}\,\mathrm{d}s$ for $\varphi \in H^{-\frac{1}{2}}(\partial\Omega)$ and $\psi \in H^{\frac{1}{2}}(\partial\Omega)$ (or vice versa) we mean the duality product $\langle \varphi, \psi\rangle_\Gamma$.

Theorem 5.11: (Green's representation in Ω_-). Let Ω_- be a bounded Lipschitz domain and $u \in H^1(\Omega_-;\Delta) \cap C^2(\Omega_-)$ be a Helmholtz solution in Ω_-. Then:

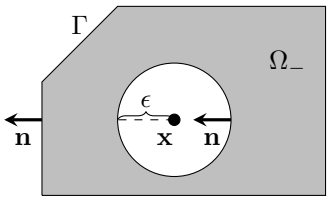
$$\int_\Gamma \left(\partial_{\mathbf{n}}^- u(\mathbf{y}) \Phi_k(\mathbf{x},\mathbf{y}) - \gamma^- u(\mathbf{y}) \frac{\partial \Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{y})} \right) \mathrm{d}s(\mathbf{y}) = \begin{cases} u(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega_-, \\ 0 & \text{if } \mathbf{x} \in \Omega_+. \end{cases} \tag{47}$$

Proof. If $\mathbf{x} \in \Omega_+$, then both u and Φ_k are Helmholtz solution in Ω_- thus the volume integral in Green's second identity (21) vanishes and what is left is (47).

If $\mathbf{x} \in \Omega_-$ we take $\epsilon > 0$ such that the ball $B_\epsilon(\mathbf{x}) := \{\mathbf{z} \in \mathbb{R}^2 : |\mathbf{z}-\mathbf{x}| < \epsilon\} \subset \Omega_-$. Applying again Green's second identity in $\Omega_- \setminus \overline{B_\epsilon(\mathbf{x})}$ we have

$$\int_\Gamma \left(\partial_{\mathbf{n}}^- u(\mathbf{y}) \Phi_k(\mathbf{x},\mathbf{y}) - \gamma^- u(\mathbf{y}) \frac{\partial \Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{y})} \right) \mathrm{d}s(\mathbf{y}) = - \int_{\partial B_\epsilon(\mathbf{x})} \left(\partial_{\mathbf{n}} u(\mathbf{y}) \Phi_k(\mathbf{x},\mathbf{y}) - \gamma u(\mathbf{y}) \frac{\partial \Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{y})} \right) \mathrm{d}s(\mathbf{y}),$$

where $\mathbf{n}$ points into the ball (this is why we have a minus sign in front of $\int_{\partial B_\epsilon(\mathbf{x})}$). We now want to take the limit for $\epsilon \searrow 0$ of the right-hand side, and see that the first term vanish, while the second converges to $u(\mathbf{x})$. The advantage of using a ball centred at $\mathbf{x}$, is that on its boundary the value of $\Phi_k(\mathbf{x},\cdot)$ and its normal derivative are constant.



To this purpose, we need some properties of Hankel functions: the formula for the first derivative of $H_0^{(1)}$ and the asymptotics for small arguments (from [DLMF, §10.7]):

$$\frac{\partial}{\partial z} H_0^{(1)}(z) = -H_1^{(1)}(z), \qquad H_0^{(1)}(z) \sim \frac{2\mathrm{i}}{\pi} \log z, \qquad H_\ell^{(1)}(z) \sim -\frac{\mathrm{i}}{\pi}(\ell-1)! \frac{2^\ell}{z^\ell}, \qquad \ell \in \mathbb{N}, \qquad z \to 0. \tag{48}$$

(Here $a(z) \sim b(z)$ for $z \to 0$ means that $\lim_{z \searrow 0} \frac{a(z)}{b(z)} = 1$.)

We take the limit $\lim_{\epsilon \to 0}$ of the first term using the expression of the fundamental solution, the divergence theorem, $\operatorname{div} \nabla = \Delta$, $\Delta u = -k^2 u$, the asymptotics (48), the mean value theorem (in the form

$\lim_{\epsilon\to 0}\frac{1}{\pi\epsilon^2}\int_{B_\epsilon(\mathbf{x})} f(\mathbf{y})\,\mathrm{d}\mathbf{y} = f(\mathbf{x})$), the boundedness of u in $\mathbf{x}$:

$$\begin{aligned}-\int_{\partial B_\epsilon(\mathbf{x})} \partial_{\mathbf{n}} u(\mathbf{y})\Phi_k(\mathbf{x},\mathbf{y})\,\mathrm{d}s(\mathbf{y}) &= -\frac{\mathrm{i}}{4}H_0^{(1)}(k\epsilon)\int_{\partial B_\epsilon(\mathbf{x})}\partial_{\mathbf{n}}u(\mathbf{y})\,\mathrm{d}s(\mathbf{y})\\ &= \frac{\mathrm{i}}{4}H_0^{(1)}(k\epsilon)\int_{B_\epsilon(\mathbf{x})}\Delta u(\mathbf{y})\,\mathrm{d}\mathbf{y}\\ &= -\frac{\mathrm{i}}{4}H_0^{(1)}(k\epsilon)k^2\int_{B_\epsilon(\mathbf{x})} u(\mathbf{y})\,\mathrm{d}\mathbf{y} \sim -\frac{\mathrm{i}}{4}\Big(\frac{2\mathrm{i}}{\pi}\log k\epsilon\Big)k^2\pi\epsilon^2 u(\mathbf{x}) \xrightarrow{\epsilon\searrow 0} 0.\end{aligned}$$

Using the expression of the radial derivative of Φ_k, the asymptotics (48), the mean value theorem ($\lim_{\epsilon\to 0}\frac{1}{2\pi\epsilon}\int_{\partial B_\epsilon(\mathbf{x})} f(\mathbf{y})\,\mathrm{d}s(\mathbf{y}) = f(\mathbf{x})$) we get

$$\begin{aligned}\int_{\partial B_\epsilon(\mathbf{x})} u(\mathbf{y})\frac{\partial\Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{y})}\,\mathrm{d}s(\mathbf{y}) &= k\frac{\mathrm{i}}{4}\,H_1^{(1)}(k\epsilon)\int_{\partial B_\epsilon(\mathbf{x})} u(\mathbf{y})\,\mathrm{d}s(\mathbf{y})\\ &\sim k\frac{\mathrm{i}}{4}\Big(\frac{-2\mathrm{i}}{\pi k\epsilon}\Big)\int_{\partial B_\epsilon(\mathbf{x})} u(\mathbf{y})\,\mathrm{d}s(\mathbf{y}) \sim \Big(\frac{1}{2\pi\epsilon}\Big)2\pi\epsilon u(\mathbf{x}) = u(\mathbf{x}).\end{aligned}$$

(The first of the two limits $\lim_{\epsilon\to 0}\int_{\partial B_\epsilon(\mathbf{x})}\dots$ vanishes, while the second one is non-zero: this is because the singularity of the first derivative of the fundamental solution Φ_k is stronger than that of Φ_k itself, so it compensates the infinitesimal length of the circle on which it is integrated.) □

Green's representation formula allows to compute any Helmholtz solution from its two traces.

Representation (47) also explains why we chose the coefficient $\frac{\mathrm{i}}{4}$ in the definition (38) of Φ_k: it allows to write a simple Green's representation.

From Green's representation it follows immediately that:

Corollary 5.12: (At least one trace is non-zero). If u is a Helmholtz solution in Ω_- and $\gamma^- u = \partial_{\mathbf{n}}^- u = 0$ then $u = 0$.
In particular, $u \neq 0$ is not simultaneously Dirichlet and Neumann eigenfunction in Ω_- for the Laplacian.

We have seen at the end of §4.2 that the solution of the impedance BVP (27) in Ω_- satisfies a Fredholm-type variational problem. We have also seen that the solution u_0 of the homogeneous impedance BVP has zero traces: $\gamma^- u_0 = \partial_{\mathbf{n}}^- u_0 = 0$. From Green's representation (47), $u_0 = 0$, so the operator associated to the sesquilinear form $\mathcal{A}_I(\cdot,\cdot)$ is injective. Since the operator is Fredholm, by Theorem 3.12 it is also invertible and the following corollary holds.

Corollary 5.13: (Well-posedness of interior impedance BVP). The impedance BVP (27) (equivalently, (28)) in Ω_- is well-posed.

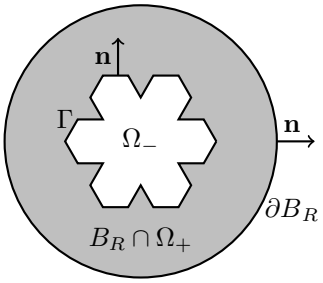
The Fredholm alternative implies that the solution of the impedance BVP (27) exists and is unique, and also that its $H^1(\Omega_-)$ norm is controlled by the norm of the data (f and g_I). However, the bounding constant is not explicit as in the situations where we can use Lax–Milgram theorem.

Exercise 5.14: (Truncated BVP). Let $\Omega_R = B_R \setminus \overline{\Omega_-}$, where B_R is an open ball containing $\overline{\Omega_-}$. Write a truncation of the EDP (32) to Ω_R, as described in Remark 4.20 imposing impedance boundary conditions on ∂B_R and sound-soft conditions on Γ. Show that the BVP obtained is well-posed.

Remark 5.15: (Green's representation with volume term). If u in (47) were not Helmholtz solution we would need to add to the left-hand side of Green's representation the volume integral term $-\int_{\Omega_-}\big(\Delta u(\mathbf{y}) + k^2 u(\mathbf{y})\big)\Phi_k(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y}$. When $\mathbf{x}\in\Omega_-$ the integrand is weakly singular, see (48), and the integral is well-defined.

Remark 5.16: (Green's representation in 3D). In the 3-dimensional case all the arguments are similar. The fundamental solution has the simpler expression $\Phi_k(\mathbf{x},\mathbf{y}) = \frac{\mathrm{e}^{\mathrm{i}k|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}$, which does not involve Bessel and Hankel function, so checking the limits for $\epsilon\searrow 0$ is simpler.

To write a Green's representation formula for unbounded domains, we denote by B_R the ball of radius $R>0$ centred at $\mathbf{0}$ and fix $\mathbf{n} = \frac{\mathbf{x}}{r}$ on ∂B_R. We work in the bounded region $B_R\cap\Omega_+$, for sufficiently large R and use Sommerfeld condition to take the limit $R\to\infty$. The unit normal $\mathbf{n}$ points into $B_R\cap\Omega_+$ on the inner boundary Γ and out of $B_R\cap\Omega_+$ on the outer boundary ∂B_R.



Lemma 5.17: (Properties of radiating solutions). Let Ω_- be a bounded Lipschitz domain, $\Gamma = \partial\Omega_-$, $\Omega_+ = \mathbb{R}^2 \setminus \overline{\Omega_-}$, and $u, w \in H^1_{\mathrm{loc}}(\overline{\Omega_+};\Delta) \cap C^2(\Omega_+)$ be two *radiating* Helmholtz solutions in Ω_+. Then:

$$\lim_{R\to\infty} \int_{\partial B_R} |u|^2 \,\mathrm{d}s < \infty, \qquad \Im \int_\Gamma \partial_{\mathbf{n}} u \, \gamma\overline{u} \,\mathrm{d}s \geq 0 \qquad \lim_{R\to\infty} \int_{\partial B_R} (\partial_{\mathbf{n}} u \, w - u \partial_{\mathbf{n}} w) \,\mathrm{d}s = 0. \quad (49)$$

Proof. We first prove the boundedness of the limit of $\|u\|_{L^2(\partial B_R)}$. The imaginary part of Green's first identity (20) with $w = \overline{u}$ in $B_R \cap \Omega_+$ gives

$$\begin{aligned}\Im \int_\Gamma \partial_{\mathbf{n}} u \, \gamma\overline{u} \,\mathrm{d}s &= \Im \int_{\partial B_R} \partial_{\mathbf{n}} u \, \gamma\overline{u} \,\mathrm{d}s + \underbrace{\Im \int_{B_R\cap\Omega_+} (k^2|u|^2 - |\nabla u|^2) \,\mathrm{d}\mathbf{x}}_{=0,\text{ imaginary part of real value}} \\ &= \frac{1}{2k} \int_{\partial B_R} \left(k^2|u|^2 + |\partial_{\mathbf{n}} u|^2 - |\partial_{\mathbf{n}} u - \mathrm{i}ku|^2\right) \mathrm{d}s,\end{aligned}$$

where we have used the identity $|a - \mathrm{i}b|^2 = |a|^2 + |b|^2 - 2\Re\{a\,\overline{\mathrm{i}b}\} = |a|^2 + |b|^2 - 2\Im\{a\bar{b}\}$, which holds for all $a, b \in \mathbb{C}$, applied to $a = \partial_{\mathbf{n}} u$, $b = ku$. Taking the limit for $R \to \infty$, the term $\int_{\partial B_R} |\partial_{\mathbf{n}} u - \mathrm{i}ku|^2 \,\mathrm{d}s \to 0$, by the Sommerfeld condition (31). The left-hand side is independent of R, thus

$$\lim_{R\to\infty} \frac{1}{2k} \left(k^2 \|u\|^2_{L^2(\partial B_R)} + \|\partial_{\mathbf{n}} u\|^2_{L^2(\partial B_R)}\right) = \Im \int_\Gamma \partial_{\mathbf{n}} u \, \gamma\overline{u} \,\mathrm{d}s < \infty.$$

Since the norms are non-negative and the limit is finite, each of them is bounded, which is the desired inequality. In particular $u = \mathcal{O}(r^{-1/2})$ for $r \to \infty$. Moreover the left-hand side is non-negative, so also the second inequality is proved.

The Sommerfeld condition, together with $u, w = \mathcal{O}(r^{-1/2})$, gives the identity involving w:

$$\begin{aligned}\int_{\partial B_R} (\partial_{\mathbf{n}} u \, w - w \partial_{\mathbf{n}} u) \,\mathrm{d}s &= \int_{\partial B_R} \left(\left(\mathrm{i}ku + o(R^{-1/2})\right) w - u\left(\mathrm{i}kw + o(R^{-1/2})\right)\right) \mathrm{d}s \\ &= \int_{\partial B_R} \left(o(R^{-1/2})\mathcal{O}(R^{-1/2}) - \mathcal{O}(R^{-1/2})o(R^{-1/2})\right) \mathrm{d}s = \int_{\partial B_R} o(R^{-1}) \,\mathrm{d}s \xrightarrow{R\to\infty} 0.\end{aligned}$$

□

Recall that we have already seen in Theorem 4.25 (Rellich's lemma) a stronger version of the second inequality in (49): if this integral is 0 then $u = 0$ (see the first formula in the proof to relate the integrals on Γ and on ∂B_R).

The Sommerfeld radiation condition can be extended to problems with complex wavenumbers with $\Im k \geq 0$, Lemma 5.17 holds also in this case, see [CK1, Thm. 3.3].

Theorem 5.18: (Green's representation in Ω_+). Let Ω_- be a bounded Lipschitz domain, $\Gamma = \partial\Omega_-$, $\Omega_+ = \mathbb{R}^2 \setminus \overline{\Omega_-}$, and $u \in H^1_{\mathrm{loc}}(\overline{\Omega_+};\Delta) \cap C^2(\Omega_+)$ be a *radiating* Helmholtz solution in Ω_+. Then:

$$-\int_\Gamma \left(\partial^+_{\mathbf{n}} u(\mathbf{y}) \Phi_k(\mathbf{x},\mathbf{y}) - \gamma^+ u(\mathbf{y}) \frac{\partial \Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{y})}\right) \mathrm{d}s(\mathbf{y}) = \begin{cases} 0 & \text{if } \mathbf{x} \in \Omega_-, \\ u(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega_+. \end{cases} \quad (50)$$

Proof. Let $R > 0$ be the radius of a ball such that $\Gamma \subset B_R$ and, if $\mathbf{x} \in \Omega_+$ also $\mathbf{x} \in B_R$. Then Green's representation (47) applied in $\Omega_+ \cap B_R$ gives

$$\left(\int_{\partial B_R} - \int_\Gamma\right)\left(\partial^+_{\mathbf{n}} u(\mathbf{y}) \Phi_k(\mathbf{x},\mathbf{y}) - \gamma^+ u(\mathbf{y}) \frac{\partial \Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{y})}\right) \mathrm{d}s(\mathbf{y}) = \begin{cases} 0 & \text{if } \mathbf{x} \in \Omega_-, \\ u(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega_+. \end{cases}$$

Taking the limit for $R \to \infty$, the integral over ∂B_R vanishes by the identity in (49) and we conclude. □

Corollary 5.19: (There are no entire radiating solutions). Let $u_+ \in H^1_{\mathrm{loc}}(\overline{\Omega_+};\Delta)$ be a radiating Helmholtz solution and $u_- \in H^1(\Omega_-;\Delta)$ a Helmholtz solution. If they have the same traces on Γ, i.e. $\gamma^+ u_+ = \gamma^- u_-$ and $\partial^+_{\mathbf{n}} u_+ = \partial^-_{\mathbf{n}} u_-$, then $u_+ = 0$ and $u_- = 0$.
In particular, the only radiating Helmholtz solution in the whole of $\mathbb{R}^2$ is $u = 0$.

In this corollary, the assumption that the solution is radiating is necessary: plane waves and smooth Fourier–Bessel functions are Helmholtz solutions in $\mathbb{R}^2$, but not radiating. The radiating Fourier–Hankel functions are not defined in the whole of $\mathbb{R}^2$ as they are singular at the origin. Intuitively, the corollary says that all radiating solutions "go towards infinity" and they need to have a source somewhere; on the other hand, plane waves are not radiating and have "a source at infinity".

We have evaluated the integral in Green's representation in $\mathbf{x} \in \Omega_- \cup \Omega_+ = \mathbb{R}^2 \setminus \Gamma$. What about $\mathbf{x} \in \Gamma$? We extend Green's representation (47)/(50) to this case.

For $\mathbf{x} \in \Gamma$ define

$$\sigma(\mathbf{x}) := \lim_{\epsilon \to 0} \frac{1}{2\pi\epsilon} \int_{\mathbf{y} \in \Omega_-, |\mathbf{y}-\mathbf{x}|=\epsilon} \mathrm{d}s.$$

If Γ is C^1 in $\mathbf{x}$ then $\sigma(\mathbf{x}) = \frac{1}{2}$; if Γ forms an angle with opening α at $\mathbf{x}$ then $\sigma(\mathbf{x}) = \frac{\alpha}{2\pi}$. By Rademacher theorem (Lipschitz functions are differentiable a.e.), for a Lipschitz Γ, $\sigma = \frac{1}{2}$ almost everywhere on Γ.

Lemma 5.20: (Green's representation on Γ). Let $u \in H^1(\Omega_-; \Delta) \cap C^0(\overline{\Omega_-})$ be a Helmholtz solution. Then

$$\int_\Gamma \left(\partial_{\mathbf{n}}^- u(\mathbf{y}) \Phi_k(\mathbf{x}, \mathbf{y}) - \gamma^- u(\mathbf{y}) \frac{\partial \Phi_k(\mathbf{x}, \mathbf{y})}{\partial n(\mathbf{y})} \right) \mathrm{d}s(\mathbf{y}) = \sigma(\mathbf{x}) u(\mathbf{x}) \qquad \mathbf{x} \in \Gamma. \tag{51}$$

If $u \in H^1_{\mathrm{loc}}(\overline{\Omega_+}; \Delta) \cap C^0(\overline{\Omega_+})$ is a radiating Helmholtz solution then

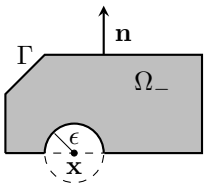
$$-\int_\Gamma \left(\partial_{\mathbf{n}}^+ u(\mathbf{y}) \Phi_k(\mathbf{x}, \mathbf{y}) - \gamma^+ u(\mathbf{y}) \frac{\partial \Phi_k(\mathbf{x}, \mathbf{y})}{\partial n(\mathbf{y})} \right) \mathrm{d}s(\mathbf{y}) = \big(1 - \sigma(\mathbf{x})\big) u(\mathbf{x}) \qquad \mathbf{x} \in \Gamma. \tag{52}$$

Proof. We prove the first identity, the second one is analogous.

For simplicity we denote the integrand with the shorthand

$$\bigstar := \partial_{\mathbf{n}}^- u(\mathbf{y}) \Phi_k(\mathbf{x}, \mathbf{y}) - \gamma^- u(\mathbf{y}) \frac{\partial \Phi_k(\mathbf{x}, \mathbf{y})}{\partial n(\mathbf{y})}.$$

For $\mathbf{x} \in \Gamma$, we apply Green's second identity on $\Omega_- \setminus \overline{B_\epsilon(\mathbf{x})}$ (see figure). Its boundary is decomposed in $\Gamma \setminus B_\epsilon(\mathbf{x})$ and $\partial B_\epsilon(\mathbf{x}) \cap \Omega_-$; on both parts we choose $\mathbf{n}$ pointing outwards:



$$0 = \int_{\partial(\Omega_- \setminus \overline{B_\epsilon(\mathbf{x})})} \bigstar \, \mathrm{d}s(\mathbf{y}) = \int_{\Gamma \setminus B_\epsilon(\mathbf{x})} \bigstar \, \mathrm{d}s(\mathbf{y}) + \int_{\partial B_\epsilon(\mathbf{x}) \cap \Omega_-} \bigstar \, \mathrm{d}s(\mathbf{y}).$$

The limit for $\epsilon \to 0$ of the integral over $\Gamma \setminus B_\epsilon(\mathbf{x})$ is exactly the integral of the same integrand over the whole Γ, the left-hand side of the assertion (understood as a Cauchy principal value):

$$\int_\Gamma \bigstar \, \mathrm{d}s(\mathbf{y}) = \lim_{\epsilon \to 0} \int_{\Gamma \setminus B_\epsilon(\mathbf{x})} \bigstar \, \mathrm{d}s(\mathbf{y}) = -\lim_{\epsilon \to 0} \int_{\partial B_\epsilon(\mathbf{x}) \cap \Omega_-} \bigstar \, \mathrm{d}s(\mathbf{y}).$$

Proceeding as in the proof of Theorem 5.11 we see that $\lim_{\epsilon \to 0} \int_{\partial B_\epsilon(\mathbf{x}) \cap \Omega_-} \partial_{\mathbf{n}}^- u(\mathbf{y}) \Phi_k(\mathbf{x}, \mathbf{y}) \ \mathrm{d}s(\mathbf{y}) = 0$, while the last term gives:

$$\begin{aligned} \int_{\partial B_\epsilon(\mathbf{x}) \cap \Omega_-} \gamma^- u(\mathbf{y}) \frac{\partial \Phi_k(\mathbf{x}, \mathbf{y})}{\partial n(\mathbf{y})} \, \mathrm{d}s(\mathbf{y}) &= k \frac{\mathrm{i}}{4} \, H_1^{(1)}(k\epsilon) \int_{\partial B_\epsilon(\mathbf{x}) \cap \Omega_-} u(\mathbf{y}) \, \mathrm{d}s(\mathbf{y}) \\ &\sim k \frac{\mathrm{i}}{4} \Big(\frac{-2\mathrm{i}}{\pi k \epsilon} \Big) \int_{\partial B_\epsilon(\mathbf{x}) \cap \Omega_-} u(\mathbf{y}) \, \mathrm{d}s(\mathbf{y}) \sim \Big(\frac{1}{2\pi\epsilon} \Big) 2\pi\sigma(\mathbf{x}) \epsilon u(\mathbf{x}) = \sigma(\mathbf{x}) u(\mathbf{x}). \end{aligned}$$

□

We summarise Green's representations (47), (50), (51) and (52) as follows. With u as in Lemma 5.20:

$\Delta u + k^2 u = 0$ in Ω_- :

$$\int_\Gamma \left(\partial_{\mathbf{n}}^- u(\mathbf{y}) \Phi_k(\mathbf{x}, \mathbf{y}) - \gamma^- u(\mathbf{y}) \frac{\partial \Phi_k(\mathbf{x}, \mathbf{y})}{\partial n(\mathbf{y})} \right) \mathrm{d}s(\mathbf{y}) = \begin{cases} u(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega_-, \\ \sigma(\mathbf{x}) u(\mathbf{x}) & \text{if } \mathbf{x} \in \Gamma, \\ 0 & \text{if } \mathbf{x} \in \Omega_+, \end{cases}$$

$\Delta u + k^2 u = 0$ in Ω_+, radiating:

$$-\int_\Gamma \left(\partial_{\mathbf{n}}^+ u(\mathbf{y})\Phi_k(\mathbf{x},\mathbf{y}) - \gamma^+ u(\mathbf{y})\frac{\partial\Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{y})}\right)\mathrm{d}s(\mathbf{y}) = \begin{cases} 0 & \text{if } \mathbf{x}\in\Omega_-,\\ \big(1-\sigma(\mathbf{x})\big)u(\mathbf{x}) & \text{if } \mathbf{x}\in\Gamma,\\ u(\mathbf{x}) & \text{if } \mathbf{x}\in\Omega_+.\end{cases}$$

Exercise 5.21: (Fundamental solution and delta function). Show that the fundamental solution deserves its name: for any $\mathbf{y}\in\mathbb{R}^2$, it satisfies $-\Delta\Phi_k(\cdot,\mathbf{y}) - k^2\Phi_k(\cdot,\mathbf{y}) = \delta_{\mathbf{y}}$ in distributional sense, where $\delta_{\mathbf{y}}$ is the Dirac delta at $\mathbf{y}$.

This means that $\int_{\mathbb{R}^2}\Phi_k(\mathbf{x},\mathbf{y})\big(\Delta\rho(\mathbf{x})+k^2\rho(\mathbf{x})\big)\,\mathrm{d}\mathbf{x} = -\rho(\mathbf{y})$ for all $\rho\in\mathscr{D}(\mathbb{R}^2)$ and $\mathbf{y}\in\mathbb{R}^2$.

Hint: integrate by parts in $\mathbb{R}^2\setminus B_\epsilon(\mathbf{y})$ and take the limit for $\epsilon\to 0$ using the technique of the proof of Theorem 5.11.

Show that for $f\in\mathcal{D}(\mathbb{R}^2)$, $u(\mathbf{y}) := \int_{\mathbb{R}^2}\Phi_k(\mathbf{x},\mathbf{y})f(\mathbf{x})\,\mathrm{d}\mathbf{x}$ is solution of the inhomogeneous Helmholtz equation $-\Delta u - k^2u = f$ and is radiating.

The operator $V: L^2(D)\to H^1_{\mathrm{loc}}(\mathbb{R}^2)$, $Vf(\mathbf{y}) := \int_{\mathbb{R}^2}\Phi_k(\mathbf{x},\mathbf{y})f(\mathbf{x})\,\mathrm{d}\mathbf{x}$, for a bounded $D\subset\mathbb{R}^2$, is called "volume potential" and is used to construct the "volume integral equations"; see [CK2, Ch. 8].

Exercise 5.22: (Discrete-valued field). Find a field $u\in C^\infty(\mathbb{R}^2)$ and a bounded, Lipschitz, connected domain Ω_- such that the function $Z:\mathbf{x}\mapsto\left|\int_\Gamma\left(\partial_{\mathbf{n}}^- u(\mathbf{y})\Phi_k(\mathbf{x},\mathbf{y}) - \gamma^- u(\mathbf{y})\frac{\partial\Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{y})}\right)\mathrm{d}s(\mathbf{y})\right|$ takes exactly the values $0,1,2,4$ for $\mathbf{x}\in\mathbb{R}^2$.

Despite the function Z is defined by the same formula for all $\mathbf{x}\in\mathbb{R}^2$, it takes exactly 4 different values.

5.4 Double-layer potential and operator

Green's representation formulas (47) and (50) mean that all Helmholtz solutions in Ω_- and all radiating Helmholtz solutions in Ω_+ can be written as boundary integrals over Γ. The integral of $\partial_{\mathbf{n}}u\ \Phi_k$ is the single-layer potential $\mathcal{S}\psi$ for $\psi=\partial_{\mathbf{n}}^\pm u$ we already know from (39). Now we are evaluating $(\mathcal{S}\psi)(\mathbf{x})$ for both $\mathbf{x}\in\Omega_+$ and $\mathbf{x}\in\Omega_-$, so we are extending the definition of the single-layer potential (39) to all points in the complement of Γ:

$$(\mathcal{S}\psi)(\mathbf{x}) := \int_\Gamma \Phi_k(\mathbf{x},\mathbf{y})\psi(\mathbf{y})\,\mathrm{d}s(\mathbf{y}) \qquad \mathbf{x}\in\Omega_+\cup\Omega_-.$$

In particular, for any $\psi\in H^{-\frac12}(\Gamma)$, $\mathcal{S}\psi$ is a Helmholtz solution both in Ω_- and Ω_+ and is radiating. With the same reasoning as in (41) (see footnote 18), the Dirichlet traces from both sides of Γ coincide:

$$\gamma^+(\mathcal{S}\psi) = \gamma^-(\mathcal{S}\psi) = S\psi.$$

In particular, if $\psi\in C^0(\Gamma)$, then $\mathcal{S}\psi\in C^\infty(\mathbb{R}^2\setminus\Gamma)$ admits an extension in $C^0(\mathbb{R}^2)$ but not in $C^1(\mathbb{R}^2)$ because, as we will see, the normal derivative of $\mathcal{S}\psi$ is discontinuous across Γ.

The second term in (47) and (50) is a new potential:

$$\boxed{(\mathcal{D}\psi)(\mathbf{x}) := \int_\Gamma \frac{\partial\Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{y})}\psi(\mathbf{y})\,\mathrm{d}s(\mathbf{y}) \qquad \mathbf{x}\in\Omega_-\cup\Omega_+.} \tag{53}$$

This is called acoustic **double-layer potential**.[20] For a function ψ on Γ, sufficiently smooth, $\mathcal{S}\psi$ is a smooth Helmholtz solution in both Ω_- and Ω_+, and satisfies the radiation condition. It can be proved that it is continuous as mapping $\mathcal{D}: H^{\frac12}(\Gamma)\to H^1_{\mathrm{loc}}(\Omega_-\cup\Omega_+)$, [CGLS12, Thm. 2.15] (here we need the density to be in $H^{\frac12}(\Gamma)$, while for the single-layer potential $H^{-\frac12}(\Gamma)$ was enough, this is because the singularity of $\nabla\Phi_k$ is stronger than that of Φ_k).

Then Green's representation can be written as:

$$\boxed{\begin{aligned} \text{if } u \text{ is Helmholtz solution in } \Omega_-: &\quad u = \mathcal{S}\partial_{\mathbf{n}}^- u - \mathcal{D}\gamma^- u &&\text{in } \Omega_-,\\ \text{if } u \text{ is radiating Helmholtz solution in } \Omega_+: &\quad u = -\mathcal{S}\partial_{\mathbf{n}}^+ u + \mathcal{D}\gamma^+ u &&\text{in } \Omega_+.\end{aligned}} \tag{54}$$

This means that any radiating Helmholtz solution is known once we know the "Cauchy data" γu and $\partial_{\mathbf{n}}u$, i.e. its Dirichlet and Neumann traces.

[20] The double-layer potential has this name because it can be thought as the acoustic potential generated by two "sheets" of charges with opposite signs, parallel to Γ, in the limit when the distance between the sheets decreases to 0 (from $\frac{\partial\Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{y})} = \lim_{\delta\to0}\frac{1}{2\delta}(\Phi_k(\mathbf{x},\mathbf{y}+\delta\mathbf{n}) - \Phi_k(\mathbf{x},\mathbf{y}-\delta\mathbf{n}))$. In electrostatics (Laplace equation) this would be a layer of dipoles.

As we did for the single-layer, we can define the **double-layer operator**:

$$\boxed{(D\psi)(\mathbf{x}) := \int_\Gamma \frac{\partial \Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{y})}\psi(\mathbf{y})\,\mathrm{d}s(\mathbf{y}) \qquad \mathbf{x}\in\Gamma.} \tag{55}$$

If Γ is of class C^2 and $\psi\in C^0(\Gamma)$, then $D\psi$ is well-defined as a standard (weakly singular) integral and $S\psi\in C^0(\Gamma)$. On the other hand, if Γ is only Lipschitz and $\psi\in L^2(\Gamma)$, then $D\psi$ must be understood as Cauchy principal value:

$$(D\psi)(\mathbf{x}) = \lim_{\epsilon\to 0}\int_{\Gamma\cap\{\mathbf{y}:|\mathbf{y}-\mathbf{x}|>\epsilon\}} \frac{\partial \Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{y})}\psi(\mathbf{y})\,\mathrm{d}s(\mathbf{y}) \qquad \text{a.e. } \mathbf{x}\in\Gamma.$$

Then $D\psi\in L^2(\Gamma)$.

Exercise 5.23: (Double-layer operator on straight segments). Assume that $\Gamma_*\subset\Gamma$ is a straight segment and that $\psi\in C^0(\Gamma)$ is supported in Γ_*. Show that $(D\psi)(\mathbf{x})=0$ for all for all $\mathbf{x}\in\Gamma_*$.

Hint: you do not need the precise value of Φ_k but only that it depends only on $|\mathbf{x}-\mathbf{y}|$.

We have seen in (41) that $S=\gamma\mathcal{S}$. One might expect that $D=\gamma\mathcal{D}$, but this is not the case.

The assertions of Lemma 5.20, Green's representation on the boundary, can be written in terms of layer operators as (compare against (54))

$$\boxed{S\partial_{\mathbf{n}}^- u - D\gamma^- u = \sigma\gamma^- u, \qquad D\gamma^+ u - S\partial_{\mathbf{n}}^+ u = (1-\sigma)\gamma^+ u.} \tag{56}$$

Now let u be a Helmholtz solution in Ω_- and denote $\psi := \gamma^- u\in H^{\frac12}(\Gamma)$. We have

$$\psi = \gamma^- u \overset{\text{Green repr. (47)}}{=} \gamma^-\mathcal{S}\partial_{\mathbf{n}}^- u - \gamma^-\mathcal{D}\psi \overset{\text{single-layer trace (41)}}{=} S\partial_{\mathbf{n}}^- u - \gamma^-\mathcal{D}\psi \qquad \text{and} \qquad \sigma\psi \overset{(56)}{=} S\partial_{\mathbf{n}}^- u - D\psi$$

Taking the difference between these two equations we have the Dirichlet **trace formula for the double-layer potential**:

$$\boxed{\gamma^-\mathcal{D}\psi = D\psi - (1-\sigma)\psi, \qquad \gamma^+\mathcal{D}\psi = D\psi + \sigma\psi.} \tag{57}$$

The second of these equations is obtained similarly using (50) in place of (47) and (52) in place of (51). Recall that $\sigma=\frac12$ in all smooth points, so (57) reads almost everywhere as

$$\gamma^\pm\mathcal{D}\psi = D\psi \pm \frac12\psi, \quad \text{or, in operator form, } \gamma^\pm\mathcal{D} = D\pm\frac12 I,$$

where I is the identity operator. All these formulas tell us that the Dirichlet trace of $\mathcal{D}$ is not simply D but a correction term is needed, due to the singular behaviour of $\frac{\partial\Phi_k}{\partial n}$. Taking the difference between the two equations in (57), the correction terms $\pm\frac12\psi$ give the **jump relation**:

$$[\![\gamma\mathcal{D}\psi]\!] := \gamma^+\mathcal{D}\psi - \gamma^-\mathcal{D}\psi = \psi.$$

Given a ψ on the boundary Γ, the double-layer potential $\mathcal{D}\psi$ is a radiating Helmholtz solution in the complement of Γ, whose jump on Γ is ψ itself.

For simplicity, in the following we write $\frac12$ instead of σ and $1-\sigma$, with the implicit convention that equalities on Γ hold almost everywhere (everywhere except at corners).

5.5 Neumann traces of the potentials: two more BIOs and jump relations

In the previous section we have learned how to construct fields in $\Omega_-\cup\Omega_+$ from distributions defined on Γ:

$$\forall\psi\in H^{-\frac12}(\Gamma), \quad \forall\varphi\in H^{\frac12}(\Gamma), \qquad \mathcal{S}\psi,\ \mathcal{D}\varphi\in H^1(\Omega_-;\Delta)\times H^1_{\mathrm{loc}}(\Omega_+;\Delta) \tag{58}$$

are radiating Helmholtz solution in the complement of the boundary Γ. Moreover, the Dirichlet traces of $\mathcal{S}\psi$ coincide: $\gamma^+\mathcal{S}\psi=\gamma^-\mathcal{S}\psi$, so $\mathcal{S}\psi\in H^1_{\mathrm{loc}}(\mathbb{R}^2;\Delta)$, while $\mathcal{D}\varphi$ is discontinuous on Γ, (57).

We now want to look at the Neumann traces of $\mathcal{S}\psi$, $\mathcal{D}\varphi$. To this purpose, we need to introduce two more BIOs (the last ones!): the **adjoint double-layer operator** D' and the **hypersingular operator** H:

$$\boxed{(D'\varphi)(\mathbf{x}) := \int_\Gamma \frac{\partial\Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{x})}\varphi(\mathbf{y})\,\mathrm{d}s(\mathbf{y}), \qquad (H\varphi)(\mathbf{x}) := \frac{\partial}{\partial n(\mathbf{x})}\int_\Gamma \frac{\partial\Phi_k(\mathbf{x},\mathbf{y})}{\partial n(\mathbf{y})}\varphi(\mathbf{y})\,\mathrm{d}s(\mathbf{y}), \quad \mathbf{x}\in\Gamma.} \tag{59}$$

Note that D' differs from D only in that the normal derivation is taken with respect to a different variable of Φ_k. If Γ is not C^2 or $\varphi \notin C^0(\Gamma)$, the adjoint double-layer operator has to be understood as a principal value integral, in the same way as D. Its name and notation are due to the identity $\int_\Gamma (D\varphi)\psi \,\mathrm{d}s = \int_\Gamma \varphi(D'\psi)\,\mathrm{d}s$, valid for all $\varphi, \psi \in L^2(\Gamma)$, which can be proved using Fubini theorem and with some complications due to the singularity of the integrand, [CGLS12, eq. (2.37)].

The hypersingular operator is more complicated: it has to be understood as a limit $(H\varphi)(\mathbf{x}) = \lim_{\mathbf{z}\to\mathbf{x}} \mathbf{n}(\mathbf{x})\cdot\nabla(\mathcal{D}\varphi)(\mathbf{z})$, for a suitable choice of the points $\mathbf{z}$. In its definition we are not allowed to move $\frac{\partial}{\partial n(\mathbf{x})}$ inside the integral because the second derivatives of Φ_k are not integrable.

The main properties of the BIOs are their relations with the traces of the two layer potentials:

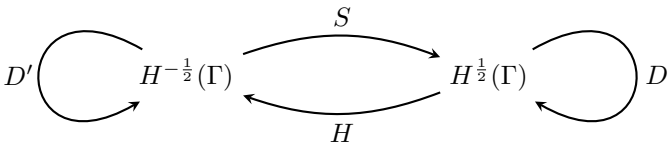
$$\boxed{\begin{aligned} &\gamma^\pm \mathcal{S} = S, && \gamma^\pm \mathcal{D} = D \pm \frac{1}{2} I, \\ &\partial_{\mathbf{n}}^\pm \mathcal{S} = D' \mp \frac{1}{2} I, && \partial_{\mathbf{n}}^\pm \mathcal{D} = H. \end{aligned}} \tag{60}$$

Here I is the identity operator. We have already derived the formulas of the Dirichlet traces, those for the Neumann ones are proved in a similar way, [CK1, §2.4–2.5]. Taking the difference between outer and inner traces we find the **jump relations** (we have already encountered those for the Dirichlet traces)

$$\boxed{\begin{aligned} [\![\gamma\mathcal{S}\psi]\!] &= \gamma^+\mathcal{S}\psi - \gamma^-\mathcal{S}\psi = 0, && [\![\gamma\mathcal{D}\psi]\!] = \gamma^+\mathcal{D}\psi - \gamma^-\mathcal{D}\psi = \psi, \\ [\![\partial_{\mathbf{n}}\mathcal{S}\psi]\!] &= \partial_{\mathbf{n}}^+\mathcal{S}\psi - \partial_{\mathbf{n}}^-\mathcal{S}\psi = -\psi, && [\![\partial_{\mathbf{n}}\mathcal{D}\psi]\!] = \partial_{\mathbf{n}}^+\mathcal{D}\psi - \partial_{\mathbf{n}}^-\mathcal{D}\psi = 0. \end{aligned}} \tag{61}$$

From (60), using (58) and the trace theorem 3.7, the mapping (continuity) properties of the BIOs follow:

$$\begin{aligned} &S : H^{-\frac12}(\Gamma) \to H^{\frac12}(\Gamma), \\ &D : H^{\frac12}(\Gamma) \to H^{\frac12}(\Gamma), \\ &D' : H^{-\frac12}(\Gamma) \to H^{-\frac12}(\Gamma), \\ &H : H^{\frac12}(\Gamma) \to H^{-\frac12}(\Gamma). \end{aligned}$$



From (60) we also see that all four operators are averages of traces of the potentials:

$$S = \{\!\{\gamma\mathcal{S}\}\!\} = \frac{\gamma^+\mathcal{S} + \gamma^-\mathcal{S}}{2}, \qquad D = \{\!\{\gamma\mathcal{D}\}\!\} = \frac{\gamma^+\mathcal{D} + \gamma^-\mathcal{D}}{2},$$
$$D' = \{\!\{\partial_{\mathbf{n}}\mathcal{S}\}\!\} = \frac{\partial_{\mathbf{n}}^+\mathcal{S} + \partial_{\mathbf{n}}^-\mathcal{S}}{2}, \qquad H = \{\!\{\partial_{\mathbf{n}}\mathcal{D}\}\!\} = \frac{\partial_{\mathbf{n}}^+\mathcal{D} + \partial_{\mathbf{n}}^-\mathcal{D}}{2}.$$

These formulas can be taken as alternative rigorous definitions of the four BIOs, given those of the two layer potentials (see e.g. [SBH19, §14.2] for a PDE different from Helmholtz).

Remark 5.24: (BIOs notation). There is no universal notation for boundary integral operators and sometimes the same symbol is used by different authors to mean different BIOs: comparing references can be a nightmare. To help navigating the literature, we list here the notation used in several good references on BIEs for Helmholtz (some of these only consider the 3D case).

	Φ_k (38)	$\mathcal{S}$ (39)	$\mathcal{D}$ (53)	S (40)	D (55)	D' (59)	H (59)
[Spence14, p. 36], [CGLS12, pp. 108–113]	Φ_k	$\mathcal{S}_k$	$\mathcal{D}_k$	S_k	D_k	D'_k	H_k
[Sayas06, §3,§11.1]	ϕ	$\mathcal{S}_\Gamma$	$\mathcal{D}_\Gamma$	V_Γ	K_Γ	K_Γ^t	W_Γ
[CK1, §2.7], [CK2, (3.8–11)]	Φ			S	K	K'	T
[Néd01, p. 116]	E			S	D	D^*	N
[SS11, §3.9]	G_k	S_k	D_k	V_k	K_k	K'_k	W_k
[Antoine, notes 2012, §3.3]	G	L	S	$\mathcal{L}$	$\mathcal{N}$	$\mathcal{D}$	$\mathcal{S}$
[Hsiao and Wendland 2008, §2.1]	E_k	V_k	W_k	V_k	K_k	K'_k	D_k
[Martin 2006, §5.1–5.3]	G	S	D	S	K	$\overline{K}^*$	N
[McLean 2000, pp. 217–218]	G	SL	DL	S	T	$\widetilde{T}$	R
[Steinbach 2008, §6.9] (W used for Laplace d.l.p.)	U_k^*	$\widetilde{V}_k$		V_k	K_k	K'_k	D_k

Moreover, in some cases the hypersingular operator H is defined with the opposite sign, e.g. [Sayas06, §11.1]. In other cases all four BIOs include a factor 2 [CK2, eq. (3.8)–(3.11)] (to avoid the factor $\frac12$ in the trace relations). [Martin, (5.1)] defines the fundamental solution as (-2) times our (standard) definition.

Remark 5.25: (Explicit double-layer formulas). Using $\frac{\partial}{\partial z}H_0^{(1)}(z) = -H_1^{(1)}(z)$, we can write more explicit formulas for the double-layer and the adjoint double-layer operators:

$$(D\psi)(\mathbf{x}) = \frac{\mathrm{i}k}{4}\int_\Gamma H_1^{(1)}(k|\mathbf{x}-\mathbf{y}|)\frac{(\mathbf{x}-\mathbf{y})\cdot\mathbf{n}(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\psi(\mathbf{y})\,\mathrm{d}s(\mathbf{y}),$$
$$(D'\psi)(\mathbf{x}) = \frac{\mathrm{i}k}{4}\int_\Gamma H_1^{(1)}(k|\mathbf{x}-\mathbf{y}|)\frac{(\mathbf{y}-\mathbf{x})\cdot\mathbf{n}(\mathbf{x})}{|\mathbf{x}-\mathbf{y}|}\psi(\mathbf{y})\,\mathrm{d}s(\mathbf{y}).$$

From this formulas, we see that if Γ is a polygon and E is one of its edges, the points $\mathbf{y}\in E$ do not contribute to the computation of $(D\psi)(\mathbf{x})$ for $\mathbf{x}\in E$, because $(\mathbf{x}-\mathbf{y})\cdot\mathbf{n}(\mathbf{y}) = 0$. So $H_1^{(1)}$ is evaluated only for $|\mathbf{x}-\mathbf{y}| > \operatorname{dist}(\mathbf{x},\partial E)$, i.e. away from the singularity. The singularity in the integrand has to be treated carefully when $\mathbf{x}$ is very close to a corner, while it is harmless otherwise. If Γ is smooth then $\frac{(\mathbf{x}-\mathbf{y})\cdot\mathbf{n}(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\to 0$ for $\mathbf{y}\to\mathbf{x}$, partially compensating the strong singularity of $H_1^{(1)}$.

5.5.1 Single-layer potential in Ω_-, value of ψ and far-field pattern

Consider the SSSP (33) and the corresponding single-layer BIE $S\psi = g_D$ (42). The representation formula $u^{\mathrm{Scat}} = (\mathcal{S}\psi)|_{\Omega_+}$ (43) gives the value of the scattered field in Ω_+ as a single-layer potential. Denote by u^- the same potential evaluated inside the scatterer Ω_-, i.e. $u^- = (\mathcal{S}\psi)|_{\Omega_-}$. Then u^- is a Helmholtz solution in Ω_- with trace $\gamma^- u^- = \gamma^-\mathcal{S}\psi = S\psi = g_D = -\gamma u^{\mathrm{Inc}}$. We now assume that: (i) k^2 is not a Dirichlet eigenvalue and (ii) u^{Inc} is an incoming wave that is Helmholtz solution also in Ω_-, e.g. a plane wave. Then, by the uniqueness of the solution of the interior Helmholtz Dirichlet problem (Proposition 4.6),

$$(\mathcal{S}\psi)|_{\Omega_-} = u^- = -u^{\mathrm{Inc}}|_{\Omega_-}. \tag{62}$$

This equality has a few useful consequences.

From one of the jump relations (61) we can relate the BIE density ψ to a "physical" quantity, the Neumann trace of the total field:

$$\psi = -[\![\partial_{\mathbf{n}}\mathcal{S}\psi]\!] = \partial_{\mathbf{n}}^-\mathcal{S}\psi - \partial_{\mathbf{n}}^+\mathcal{S}\psi = \partial_{\mathbf{n}}(-u^{\mathrm{Inc}}) - \partial_{\mathbf{n}}^+ u^{\mathrm{Scat}} = -\partial_{\mathbf{n}}^+ u^{\mathrm{Tot}}.$$

This allows to compute the Neumann trace of the scattered field from the data and the BIE solution as

$$\partial_{\mathbf{n}}^+ u^{\mathrm{Scat}} = -\psi - \partial_{\mathbf{n}} u^{\mathrm{Inc}}.$$

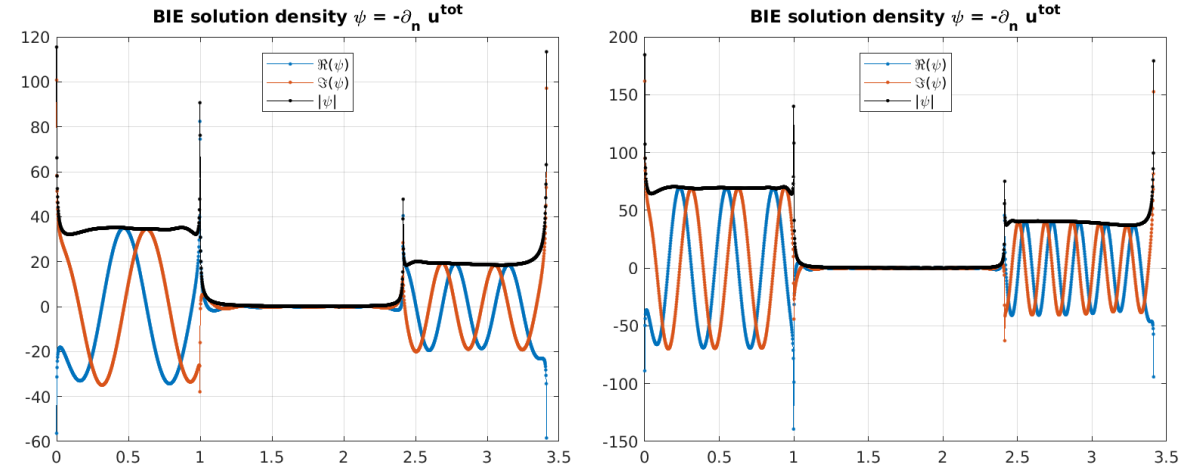


Figure 18: Left: the density $\psi = -\partial_{\mathbf{n}}^+ u^{\mathrm{Tot}}$ for the problem of Figure 12 with $k = 20$, computed with the BEM of §5.2. The x-axis represent the curvilinear abscissa along the boundary of the triangle Γ, starting from the lower-left vertex and proceeding anticlockwise. We observe that ψ oscillates on the two illuminated sides of Γ, is small (but non-zero) on the shadow side, and has singularities at the three vertices. Right: the same for $k = 40$.

We recall from Remark 4.21 that the far-field pattern of the scattered field is the function $u_\infty\in C^\infty(\mathbb{S}^1)$ such that $u^{\mathrm{Scat}}(\mathbf{x}) = \frac{\mathrm{e}^{\mathrm{i}kr}}{\sqrt{r}}(u_\infty(\theta)+\mathcal{O}(r^{-1}))$. Its explicit formula (34) requires the value of $\partial_{\mathbf{n}}^+ u^{\mathrm{Scat}}$ on Γ.

We can use the formula that we have just obtained to compute the far-field pattern of the scattered field (solution of the SSSP (33)) from the BIE (42) solution ψ:

$$\begin{aligned} u_\infty(\theta) &= \frac{\mathrm{e}^{\mathrm{i}\frac{\pi}{4}}}{\sqrt{8\pi k}} \int_\Gamma \left(\gamma^+ u^{\mathrm{Scat}}(\mathbf{y}) \partial_{\mathbf{n}} \mathrm{e}^{-\mathrm{i}k\mathbf{y}\cdot\mathbf{d}} - \partial_{\mathbf{n}}^+ u^{\mathrm{Scat}}(\mathbf{y}) \mathrm{e}^{-\mathrm{i}k\mathbf{y}\cdot\mathbf{d}} \right) \mathrm{d}s(\mathbf{y}) \\ &= \frac{\mathrm{e}^{\mathrm{i}\frac{\pi}{4}}}{\sqrt{8\pi k}} \int_\Gamma \Big(-\gamma u^{\mathrm{Inc}}(\mathbf{y}) \partial_{\mathbf{n}} \mathrm{e}^{-\mathrm{i}k\mathbf{y}\cdot\mathbf{d}} + \big(\psi(\mathbf{y}) + \partial_{\mathbf{n}} u^{\mathrm{Inc}}(\mathbf{y})\big) \mathrm{e}^{-\mathrm{i}k\mathbf{y}\cdot\mathbf{d}} \Big) \mathrm{d}s(\mathbf{y}) \\ &= \frac{\mathrm{e}^{\mathrm{i}\frac{\pi}{4}}}{\sqrt{8\pi k}} \int_\Gamma \psi(\mathbf{y}) \mathrm{e}^{-\mathrm{i}k\mathbf{y}\cdot\mathbf{d}} \, \mathrm{d}s(\mathbf{y}) \qquad\qquad \mathbf{d} = (\cos\theta, \sin\theta). \end{aligned}$$

The last equality comes from Green's second identity (21) in Ω_-, applied to u^{Inc} and the plane wave $\mathbf{y} \mapsto \mathrm{e}^{-\mathrm{i}k\mathbf{y}\cdot\mathbf{d}}$ (recall that here we have assumed that the datum u^{Inc} is a Helmholtz solution in Ω_-).

Exercise 5.26: (Far-field pattern with BEM). Use your BEM code to approximate the far-field pattern of the field scattered by a polygon; see an example in Figure 14.

Remark 5.27: (Checking BEM accuracy). The formula $(\mathcal{S}\psi + u^{\mathrm{Inc}})|_{\Omega_-} = 0$ is useful to check the correctness of a BEM implementation of the BIE (42). The routine used to evaluate the numerical near-field $u_N = \mathcal{S}\psi_N$ in a portion of Ω_+ can be used to approximate $\mathcal{S}\psi$ in Ω_-. The value $|\mathcal{S}\psi_N + u^{\mathrm{Inc}}|$ in Ω_- must be small for an accurate BEM implementation and must decrease to 0 when the BEM mesh is refined.

Choose a simple domain Υ compactly contained in Ω_- and test your BEM code by computing the value $\|\mathcal{S}\psi_N + u^{\mathrm{Inc}}\|_{L^2(\Upsilon)} / \|u^{\mathrm{Inc}}\|_{L^2(\Upsilon)}$ and see how this ratio depends on the problem parameters and on the numerical ones (k, Ω_-, N, quadrature,...). (The reason for taking Υ instead of Ω_- is that the convergence to zero of $(\mathcal{S}\psi_N + u^{\mathrm{Inc}})(\mathbf{x})$ for $N \to \infty$ is slow for $\mathbf{x} \in \Omega_-$ close to Γ.)

Remark 5.28: (Kirchhoff/physical optics approximation). We know from the analysis in §4.1 that, when a plane wave $u^{\mathrm{Inc}}(\mathbf{x}) = \mathrm{e}^{\mathrm{i}k\mathbf{x}\cdot\mathbf{d}}$ is reflected by an infinite flat Dirichlet obstacle, then the corresponding total field satisfies $\partial_{\mathbf{n}} u^{\mathrm{Tot}} = 2\partial_{\mathbf{n}} u^{\mathrm{Inc}}$, where $\mathbf{n}$ is the unit normal to the obstacle. We can imagine that when the same wave hits a flat part of a "very large", convex, bounded, obstacle Ω_- then, at least locally, it behaves as if the obstacle was infinite. For a time-harmonic wave, "very large" means much larger than the problem wavelength $\lambda = \frac{2\pi}{k}$. This suggests to approximate the BIE solution $\psi = -\partial_{\mathbf{n}}^+ u^{\mathrm{Tot}}$ with

$$\psi^{PO} := \begin{cases} -2\partial_{\mathbf{n}} u^{\mathrm{Inc}} = -2\mathrm{i}k\mathbf{d}\cdot\mathbf{n} u^{\mathrm{Inc}} & \text{on the part of } \Gamma \text{ where } \mathbf{d}\cdot\mathbf{n} \le 0 \text{ i.e., in the part illuminated by } u^{\mathrm{Inc}}, \\ 0 & \text{on the part of } \Gamma \text{ where } \mathbf{d}\cdot\mathbf{n} > 0 \text{ i.e., in the shadow part of } \Gamma. \end{cases}$$

(Recall Exercise 4.4.) Then one can compute $u^{PO} = \mathcal{S}(\psi^{PO})$ as an approximation of u^{Scat}. Numerically, this only requires approximating the representation formula, with no need to solve linear systems or to approximate BIEs. This is called *Kirchhoff* or *physical optics* approximation (sometimes *geometrical optics* approximation), see [CK2, pp. 57] and [CGLS12, §3]. Physically, it corresponds to taking into account only the reflection of the incoming wave disregarding more subtle phenomena such as diffraction by the corners and/or the curved parts of Γ. This can be accurate only for large values of k.

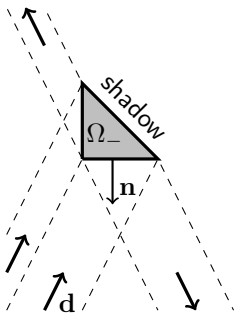


Figure 18 suggests that indeed, for the example with triangular Ω_- of Figure 12, ψ is close to 0 on the hypotenuse (in the shadow) and close to $-2\partial_{\mathbf{n}}\mathrm{e}^{\mathrm{i}k\mathbf{x}\cdot\mathbf{d}}$ on the other sides. The physical optics approximation completely misses the singularities of ψ at the corners of the polygon.

Exercise 5.29: (Physical optics approximation vs BEM). Compute the physical optics approximation ψ^{PO} and u^{PO} and plot them against ψ_N and u_N for a given convex polygon and u^{Inc}. Plot some norm of their difference as function of the wavenumber.

Remark 5.30: (Geometrical theory of diffraction). The Kirchhoff approximation described in Remark 5.28 does not include diffraction phenomena, e.g. it incorrectly predicts $\partial_{\mathbf{n}} u^{\mathrm{Tot}} = 0$ in the shadow region $\{\mathbf{x} \in \Gamma, \mathbf{d}\cdot\mathbf{n}(\mathbf{x}) > 0\}$ of the boundary. An improvement is the *Geometrical Theory of Diffraction* (GTD) developed by J.B. Keller. Assume that Ω_- is a convex polygon and that the incoming field is a plane wave. Let $\mathbf{x} : [0, L_j] \to \Gamma_j$ be the arclength parametrisation of the side Γ_j of Ω_-. The GTD predicts that

$$\psi\big(\mathbf{x}(s)\big) = -\partial_{\mathbf{n}} u^{\mathrm{Tot}}\big(\mathbf{x}(s)\big) = \psi^{PO}\big(\mathbf{x}(s)\big) + v_j^+(s)\mathrm{e}^{\mathrm{i}ks} + v_j^-(L_j - s)\mathrm{e}^{-\mathrm{i}ks},$$

where $v_j^\pm$ are non-oscillatory functions that are singular at $s = 0$ and smooth for $s > 0$. This means that $\psi|_{\Gamma_j}$ is sum of the physical-optics approximation and of two waves creeping along Γ in opposite directions

and that singularities may appear only at corners. This is consistent with what we see in Figure 18 and is the starting point for the construction of the very efficient "hybrid numerical-asymptotic" (HNA) BEM schemes, see [CGLS12, §3].

6 Well-posedness and error analysis

6.1 Well-posedness of the single-layer BIE

We want to study the well-posedness of the single-layer BIE $S\psi = g_D$ (42). To this purpose, we want to verify that the single-layer operator $S : H^{-\frac{1}{2}}(\Gamma) \to H^{\frac{1}{2}}(\Gamma)$ is (i) injective and (ii) Fredholm. When both conditions are satisfied, then Fredholm alternative (Theorem 3.12) implies that S is invertible and the BIE is well-posed. However, injectivity is not always true: the EDP (32) is always well-posed (§4.4) but its BIE (42) might fail.

6.1.1 Injectivity of the single-layer operator

As in §4.2, two cases may happen:

- If $\Lambda = k^2$ is a Dirichlet eigenvalue in Ω_-, then there exists an eigenfunction $w \neq 0$ such that $\Delta w + k^2 w = 0$ and $\gamma^- w = 0$. Define $\psi = \partial_{\mathbf{n}}^- w$. By Green's representation (47), $w = \mathcal{S}\partial_{\mathbf{n}}^- w - \mathcal{D}\gamma^- w = \mathcal{S}\partial_{\mathbf{n}}^- w = \mathcal{S}\psi$. Then $\psi \neq 0$ because a non-trivial Dirichlet eigenfunction has non-trivial Neumann trace (Corollary 5.12) and $S\psi = \gamma^-\mathcal{S}\psi = \gamma^- w = 0$. In this case the single-layer operator is not injective: $0 \neq \psi \in \ker S$. The BIE (42) is not well-posed.
- If $\Lambda = k^2$ is not a Dirichlet eigenvalue in Ω_-, then assume that $S\psi = 0$ and define $u = \mathcal{S}\psi$. We have that $u^- = u|_{\Omega_-}$ is Helmholtz solution in Ω_- and $\gamma^- u^- = \gamma^-\mathcal{S}\psi = S\psi = 0$. But the interior homogeneous Helmholtz Dirichlet BVP is well-posed by Proposition 4.6 and admits only the solution $u^- = 0$. Similarly $u^+ = u|_{\Omega_+}$ is a radiating Helmholtz solution in Ω_+ with $\gamma^+ u^+ = \gamma^+\mathcal{S}\psi = S\psi = 0$. By the well-posedness of the EDP of §4.4 also $u^+ = 0$. The jump relation (61) gives $\psi = -[\![\partial_{\mathbf{n}}\mathcal{S}\psi]\!] = -\partial_{\mathbf{n}}^+ u^+ + \partial_{\mathbf{n}}^- u^- = 0$, so the single-layer operator is injective.

Combining with Proposition 4.6 we obtain the following fact.

Lemma 6.1: (Injectivity of S). For each bounded, Lipschitz $\Omega_- \subset \mathbb{R}^2$, there exist a sequence of positive number $k_1 < k_2 < \ldots$, $\lim_{j\to\infty} k_j = \infty$, such that S is injective if and only if $k \neq k_j$ for all $j \in \mathbb{N}$.

These values are called **spurious resonances** or **spurious frequencies**. Even if the BIE (42) is not solvable in this case, the EDP (32) is well-posed: the interior eigenvalues affect and spoil the BIE formulation of the exterior problem. This can be understood as follows: the same BIE solves both an exterior and an interior Helmholtz Dirichlet problem (with solutions $(\mathcal{S}\psi)|_{\Omega_+}$ and $(\mathcal{S}\psi)|_{\Omega_-}$), when the latter is not well-posed then the BIE cannot be well-posed either. In §6.2 we will see other (slightly more complicated) BIEs that always admit a solution.

Exercise 6.2: (Spurious frequencies and BEM). Spurious resonances affect numerical computations. Plot the condition number of the BEM matrix $\underline{\underline{\mathbf{A}}}^{C/G}$ and the accuracy test of Remark 5.27 (e.g. $\left\|\mathcal{S}\psi_N + u^{\mathrm{Inc}}\right\|_{L^2(\Upsilon)}$) for several values of k close to a resonance to see how they blow up.

Hint: choose Ω_- as a square, so that the values of k_j are easily computed by hand as in §4.2.

6.1.2 The single-layer operator is Fredholm

We now want to show that the single-layer operator $S : H^{-\frac{1}{2}}(\Gamma) \to H^{\frac{1}{2}}(\Gamma)$ is Fredholm, i.e. it is sum of an invertible and a compact operator.

We define the single-layer operator for Laplace equation $-\Delta u = 0$ as

$$(S_0\psi)(\mathbf{x}) := -\frac{1}{2\pi}\int_\Gamma \log\frac{|\mathbf{x}-\mathbf{y}|}{d}\psi(\mathbf{y})\,\mathrm{d}s(\mathbf{y}), \qquad \mathbf{x}\in\Gamma, \tag{63}$$

where d is a positive parameter[21] satisfying $d > \mathrm{diam}(\Gamma) = \sup_{\mathbf{x},\mathbf{y}\in\Gamma}|\mathbf{x}-\mathbf{y}|$.

[21] Different values of d give different "versions" of the single-layer. This corresponds to adding a constant to S_0; recall the following difference between Laplace and Helmholtz solutions: $\Delta u = 0 \Rightarrow \Delta(u+C) = 0$ but $(\Delta + k^2)u = 0 \Rightarrow (\Delta+k^2)(u+C) \neq 0$ for all constants $C \neq 0$. This is related to the fact that the 2D Laplace fundamental solution does not decay to 0 at infinity. We will see in Remark 6.8 that the precise value of d only matters to ensure the coercivity of S_0. This is not true in 3D, where there is no need for the parameter d.

Given a positive number c, we also define the single-layer operator for the homogeneous reaction–diffusion equation (in this setting sometimes called "Yukawa equation" [SBH19], or "modified Helmholtz equation" [Spence14]) $-\Delta u + c^2 u = 0$ as

$$(S_c\psi)(\mathbf{x}) := \frac{1}{2\pi}\int_\Gamma K_0(c|\mathbf{x}-\mathbf{y}|)\psi(\mathbf{y})\,\mathrm{d}s(\mathbf{y}), \qquad \mathbf{x}\in\Gamma, \tag{64}$$

where $K_0(t) := \frac{\mathrm{i}\pi}{2}H_0^{(1)}(\mathrm{i}t)$ for $t \geq 0$ is the "modified Bessel function of the second kind". [22] Formally, S_c equals the single-layer operator for the Helmholtz equation with purely imaginary wavenumber (i.e. with $\mathrm{i}c$ in place of k).

The following two facts hold:

Lemma 6.3: (Single-layer: compactness of $S - S_c$ and coercivity of S_c). For all $c \geq 0$,

- $S - S_c : H^{-\frac{1}{2}}(\Gamma) \to H^{\frac{1}{2}}(\Gamma)$ is compact;
- $S_c : H^{-\frac{1}{2}}(\Gamma) \to H^{\frac{1}{2}}(\Gamma)$ is coercive, i.e. $\exists \alpha_c > 0$ such that $\langle S_c\psi, \psi\rangle_\Gamma \geq \alpha_c \|\psi\|^2_{H^{-\frac{1}{2}}(\Gamma)} \quad \forall \psi \in H^{-\frac{1}{2}}(\Gamma)$.

In the next pages we explain where Lemma 6.3 comes from, but we don't give a complete proof.

Note that we allow also the case $c = 0$: S_c can stand either for the Laplace or for the reaction–diffusion operator.

The second item of Lemma 6.3 and Lax–Milgram theorem imply that $S_c : H^{-\frac{1}{2}}(\Gamma) \to H^{\frac{1}{2}}(\Gamma)$ is invertible and that the corresponding BIE (for Laplace or reaction–diffusion) $S_c\psi = g$ is well-posed for all $g \in H^{\frac{1}{2}}(\Gamma)$. From Definition 3.11, $S = S_c + (S - S_c)$ is a Fredholm operator. Moreover, the sesquilinear form $\mathcal{A}(\cdot,\cdot)$ defined in (44) and used to define and implement the Galerkin-BEM method is sum of a coercive and a compact form: $\mathcal{A}(\psi,\xi) = \langle S\psi,\xi\rangle_\Gamma = \langle S_c\psi,\xi\rangle_\Gamma + \langle (S-S_c)\psi,\xi\rangle_\Gamma$.

By Fredholm alternative 3.12, S is invertible if and only if it is injective. By §6.1.1, the single-layer BIE $S\psi = g_D$ is well-posed (i.e. S is invertible) if and only if $-k^2$ is not a Dirichlet eigenvalue for Ω_-. Combining with what we already know about the eigenvalues, we obtain the following fact.

Theorem 6.4: (Single-layer BIE well-posedness). For each Ω_- there exist a sequence of positive number $k_1 < k_2 < \ldots$, $\lim_{j\to\infty} k_j = \infty$, such that the BIE $S\psi = g_D$ is well-posed for all $g_D \in H^{\frac{1}{2}}(\Gamma)$ if and only if $k \neq k_j$ for all $j \in \mathbb{N}$.

We repeat: Theorem 6.4 follows from Lemma 6.3, Lax–Milgram theorem, Fredholm alternative 3.12, and Lemma 6.1.

Why do we consider both the Laplace (S_0) and the reaction–diffusion (S_c) cases? To prove the Fredholm property of S any of them would be enough. On one hand, the Laplace single-layer S_0 is easier to study because it does not require special functions, while, on the other hand, the coercivity is easier to prove for S_c than for S_0 (in 2D).[23]

We study the properties of compactness and coercivity in the following, starting from the case of a circular scatterer.

Remark 6.5: (Helmholtz=Laplace+low-order, once again). This decomposition of the Helmholtz operator in a "Laplace part" and "whatever is left" should remind you the technique used in §4.2 for Helmholtz problems on bounded domains. In that case, using the Gårding inequality, we have decomposed the Helmholtz sesquilinear form (either $\mathcal{A}$ of (26) or $\mathcal{A}_I$ of (28)) in a coercive part corresponding to an elliptic equation and a compact perturbation term multiple of $k^2\int_\Omega u\overline{w}\,\mathrm{d}\mathbf{x}$.

Remark 6.6: (Helmholtz vs coercive cases: BIE well-posedness and Galerkin method). The coercivity of S_0 implies that the single-layer BIE for the Laplace equation $S_0\psi = g_D$ is always well-posed. Moreover, as we are in a Lax–Milgram setting, every Galerkin-BEM discretisation of the variational problem $\langle S_0\psi,\xi\rangle_\Gamma = \langle g_D,\xi\rangle_\Gamma$ is well-posed, quasi-optimal and gives symmetric positive-definite matrices.

[22] This is defined from the complex-analytic extension of $H_0^{(1)}$ to complex argument, in such a way that K_0 is a real-valued, positive, monotonically decreasing function; see [DLMF, eq. 10.27.E8]. It satisfies $K_0(t) \sim -\log t$ for $t \to 0$ and $K_0(t) \sim \sqrt{\pi/2t}\,\mathrm{e}^{-t}$ for $t \to \infty$, [DLMF, eq. 10.25.3 and 10.30.3]. It is easy to see that it satisfies the "modified Bessel equation" $t^2K_0''(t) + tK_0'(t) - t^2K_0(t) = 0$ and from this that $-\Delta_{\mathbf{x}}\Phi_c(\mathbf{x},\mathbf{y}) + c^2\Phi_c(\mathbf{x},\mathbf{y}) = \delta_{\mathbf{y}}(\mathbf{x})$ where $\Phi_c(\mathbf{x},\mathbf{y}) := \frac{1}{2\pi}K_0(c|\mathbf{x}-\mathbf{y}|) = \frac{\mathrm{i}}{4}H_0^{(1)}(\mathrm{i}c|\mathbf{x}-\mathbf{y}|)$.

[23] We write also the fundamental solutions of the Helmholtz, Laplace and reaction–diffusion PDEs in $\mathbb{R}^3$:

$$\Delta u + k^2u = 0 \;\to\; \Phi_k(\mathbf{x},\mathbf{y}) = \frac{\mathrm{e}^{\mathrm{i}k|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}, \quad \Delta u = 0 \;\to\; \Phi_0(\mathbf{x},\mathbf{y}) = \frac{1}{4\pi|\mathbf{x}-\mathbf{y}|}, \quad -\Delta u + c^2u = 0 \;\to\; \Phi_c(\mathbf{x},\mathbf{y}) = \frac{\mathrm{e}^{-c|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}$$

([Spence14, §9.1]). Differently from the 2D case, all three fundamental solutions can be written as special instances of Φ_c with complex c (or Φ_k with complex k). In particular, in 3D the Laplace case is the limit for $k \searrow 0$ or $c \searrow 0$ of the other two.

This is not true in the Helmholtz case. To ensure well-posedness of the Galerkin method, some conditions on the discrete space are needed (roughly speaking, it must have sufficiently good approximation properties). Under these conditions, we have a quasi-optimality estimate as in Céa lemma. We sketch this theory in §6.3.

6.1.3 Continuity, compactness and coercivity of single-layer BIOs on a circle

Let K be a BIO on the circle ∂B_R with kernel $\kappa : (0,\infty) \to \mathbb{C}$, i.e.

$$(Kv)(\mathbf{x}) = \int_{\partial B_R} \kappa(|\mathbf{x}-\mathbf{y}|) v(\mathbf{y}) \,\mathrm{d}s(\mathbf{y}), \tag{65}$$

for v defined on Γ. The distance between two points on the circle can be computed using polar coordinates:

$$|\mathbf{x}-\mathbf{y}| = |R\mathrm{e}^{\mathrm{i}\theta_{\mathbf{x}}} - R\mathrm{e}^{\mathrm{i}\theta_{\mathbf{y}}}| = R|1-\mathrm{e}^{\mathrm{i}(\theta_{\mathbf{x}}-\theta_{\mathbf{y}})}| = R\sqrt{2-2\cos(\theta_{\mathbf{x}}-\theta_{\mathbf{y}})}.$$

The action of the operator K on a function v can be written as the multiplication of the Fourier coefficients of the argument $v(\mathbf{x}) = \sum_{\ell\in\mathbb{Z}} \hat{v}_\ell \mathrm{e}^{\mathrm{i}\ell\theta}$ by some coefficients K_ℓ:

$$\begin{aligned}(Kv)(\mathbf{x}) = \int_{\partial B_R} \kappa(|\mathbf{x}-\mathbf{y}|)v(\mathbf{y})\,\mathrm{d}s(\mathbf{y}) &= R\int_0^{2\pi} \kappa(R\sqrt{2-2\cos(\theta_{\mathbf{x}}-\theta)}) \sum_{\ell\in\mathbb{Z}} \hat{v}_\ell \mathrm{e}^{\mathrm{i}\ell\theta}\,\mathrm{d}\theta \qquad (\alpha = \theta - \theta_{\mathbf{x}}) \\ &= \sum_{\ell\in\mathbb{Z}} \hat{v}_\ell \mathrm{e}^{\mathrm{i}\ell\theta_{\mathbf{x}}} \underbrace{R\int_0^{2\pi} \kappa(R\sqrt{2-2\cos\alpha})\mathrm{e}^{\mathrm{i}\ell\alpha}\,\mathrm{d}\alpha}_{=:K_\ell} = \sum_{\ell\in\mathbb{Z}} \hat{v}_\ell K_\ell\, \mathrm{e}^{\mathrm{i}\ell\theta_{\mathbf{x}}}.\end{aligned}$$

If $K_\ell = \mathcal{O}(\ell^a)$ for some $a\in\mathbb{R}$ then, from the definition (19) of the Sobolev spaces on the circular boundary, $K : H^s(\partial B_R) \to H^{s-a}(\partial B_R)$ as a bounded operator. But, how to estimate the coefficients K_ℓ?

The values K_ℓ are the Fourier coefficients of the function $\alpha \mapsto R\kappa(R\sqrt{2-2\cos\alpha})$ on $(0,2\pi)$. Parseval's theorem ($\int_0^{2\pi} |f(\theta)|^2\,\mathrm{d}\theta = 2\pi\sum_{\ell\in\mathbb{Z}} |\hat{f}_\ell|^2$) implies that the Fourier coefficients of an $L^2(0,2\pi)$ function decay as $o(\ell^{-1/2})$. As we have seen in §3.3 using that $(\mathrm{e}^{\mathrm{i}\ell\theta})' = \mathrm{i}\ell\mathrm{e}^{\mathrm{i}\ell\theta}$, if $f'\in L^2(0,2\pi)$ then $\hat{f}_\ell = o(\ell^{-3/2})$.

The function $\alpha\mapsto R\sqrt{2-2\cos\alpha}$ is Lipschitz (you can verify that its derivative is $\pm\sqrt{\frac{1+\cos\alpha}{2}}$). Thus, if $\kappa\in L^2(0,2R)$ then $K_\ell = o(\ell^{-1/2})$, if moreover $\kappa'\in L^2(0,2R)$ then $K_\ell = o(\ell^{-3/2})$. So, useful relations between the properties of the kernel κ and the continuity of the operator K in the form (65) are

$$\kappa\in L^2(0,2R) \;\Rightarrow\; K : H^s(\partial B_R) \to H^{s+\frac12}(\partial B_R), \qquad \kappa\in H^1(0,2R) \;\Rightarrow\; K : H^s(\partial B_R) \to H^{s+\frac32}(\partial B_R).$$

(From the formulas above, the properties of $\kappa(t)$ for $t > 2R$, the diameter of the circle, are irrelevant.)

What are the kernels of the Helmholtz and Laplace single-layer operators? We have

$$\begin{aligned} K = S \quad &\Rightarrow \quad \kappa(t) = \frac{\mathrm{i}}{4} H_0^{(1)}(kt) && \in L^2(0,2R), \\ K = S_0 \quad &\Rightarrow \quad \kappa(t) = -\frac{1}{2\pi}\log\frac{|t|}{d} && \in L^2(0,2R), \\ K = S - S_0 \quad &\Rightarrow \quad \kappa(t) = \frac{\mathrm{i}}{4} H_0^{(1)}(kt) + \frac{1}{2\pi}\log\frac{|t|}{d} && \in H^1(0,2R). \end{aligned}$$

The first line gives $S : H^s(\partial B_R) \to H^{s+\frac12}(\partial B_R)$. This is not new: we already mentioned that $S : H^{-\frac12}(\Gamma) \to H^{\frac12}(\Gamma)$ for all Lipschitz boundaries, which, for $s = -\frac12$, is a stronger result. We have also seen from the numerical computations in Figure 15 and Remark 5.1 that $K_\ell \sim \ell^{-1}$, so we cannot expect any stronger continuity property than this.

We now look at the difference between Helmholtz and Laplace single-layer operators $S - S_0$ (on ∂B_R). From the asymptotic formula $H_0^{(1)}(z) = \mathrm{i}\frac{2}{\pi}\log z + 1 + \mathrm{i}\frac{2}{\pi}(\gamma-2) + \mathcal{O}(z^2)$ by, e.g., [DLMF, eq. 10.8.2] with the Euler's constant $\gamma\approx 0.57721$ we have that κ is bounded and κ' is bounded (with a jump at 0 because of the $\sqrt{2-2\cos\alpha}$ term). So $S - S_0 : H^{-\frac12}(\partial B_R) \to H^1(\partial B_R)$. (In Figure 19 we demonstrate numerically a stronger continuity property, i.e. that $S - S_0 : H^{-\frac12}(\partial B_R) \to H^{\frac52}(\partial B_R)$.) Since the inclusion $\iota : H^1(\partial B_R) \to H^{\frac12}(\partial B_R)$ is compact, then **the difference between the two single-layers $S - S_0 : H^{-\frac12}(\partial B_R) \to H^{\frac12}(\partial B_R)$ is compact**.[24]

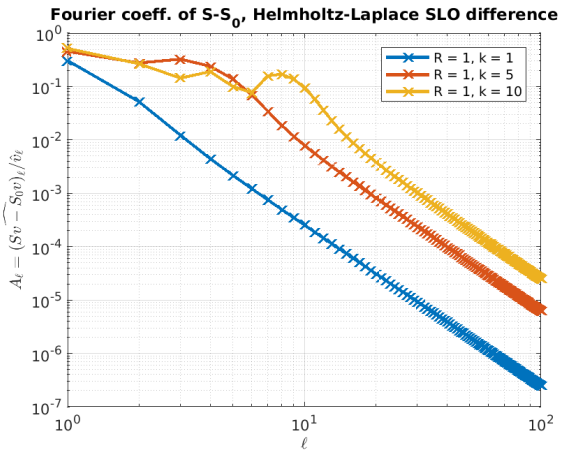


Figure 19: A log–log plot of the Fourier coefficients K_ℓ for the operator $S - S_0$, difference between Helmholtz and Laplace single-layer operators, on a circle. The coefficients decay as $K_\ell \sim \ell^{-3}$, so the operator appears to be continuous $H^s(\partial B_R) \to H^{s+3}(\partial B_R)$. In particular it is compact $H^{-\frac{1}{2}}(\partial B_R) \to H^{\frac{1}{2}}(\partial B_R)$. Compare with the coefficients of S in Figure 15.

Precisely the same holds with S_c in place of S_0: this is because the corresponding kernels $\kappa(t) = -\frac{1}{2\pi}\log\frac{|t|}{d}$ and $\kappa(t) = \frac{1}{2\pi}K_0(c|t|)$ have the same behaviour at $t = 0$.

We can use the expansion in circular harmonics also to check the coercivity of an integral operator. The sesquilinear form associated to K diagonalises in the Fourier basis (recall §3.3):

$$\langle Kv, w\rangle_{\partial B_R} = \left\langle \sum_{\ell\in\mathbb{Z}} \hat{v}_\ell K_\ell \mathrm{e}^{\mathrm{i}\ell\theta_{\mathbf{x}}}, \sum_{m\in\mathbb{Z}} \hat{w}_\ell \mathrm{e}^{\mathrm{i}m\theta_{\mathbf{x}}} \right\rangle_{\partial B_R} = 2\pi R \sum_{\ell\in\mathbb{Z}} \hat{v}_\ell K_\ell \overline{\hat{w}_\ell}.$$

If $K_\ell \in \mathbb{R}$ and $K_\ell \geq c(1+\ell^2)^s$ for all ℓ and some $c > 0$, $s \in \mathbb{R}$ [25], then $|\langle Kv, v\rangle_{\partial B_R}| \geq 2\pi Rc \sum_{\ell\in\mathbb{Z}} |\hat{v}_\ell|^2 (1+\ell^2)^s = c\,\|v\|^2_{H^s(\partial B_R)}$, i.e. K is coercive in $H^s(\partial B_R)$.

Now look at the Laplace single-layer, i.e. K with $\kappa(t) = -\frac{1}{2\pi}\log\frac{|t|}{d}$. Since $\kappa(R\sqrt{2-2\cos(t)})$ is real and even-symmetric, its Fourier coefficients K_ℓ are real.

Exercise 6.7: (Laplace single-layer Fourier coefficients). Compute numerically the coefficients K_ℓ for the Laplace single-layer S_0 and verify that they satisfy $K_\ell(1+\ell^2)^{\frac{1}{2}} > c$. (For $R = 1$, $c \approx \frac{1}{2}$.)

From this exercise it follows that the **Laplace single-layer operator S_0 is coercive in $H^{-\frac{1}{2}}(\partial B_R)$**. The same argument can be repeated for S_c.

We have verified, partly with numerical experiments, both items in Lemma 6.3 for the case $\Gamma = \mathbb{S}^1$.

6.1.4 Compactness of $S - S_0$

We sketch the main ideas used to prove the compactness of $(S - S_0) : H^{-\frac{1}{2}}(\Gamma) \to H^{\frac{1}{2}}(\Gamma)$. Making them precise and rigorous is not trivial.

The key result to prove compactness of BIOs is the following: an operator $K : L^2(\Gamma) \to L^2(\Gamma)$ in the form $(Kv)(\mathbf{x}) = \int_\Gamma \kappa(|\mathbf{x}-\mathbf{y}|)v(\mathbf{y})\,\mathrm{d}s(\mathbf{y})$ is compact if the kernel $\kappa : [0, \operatorname{diam}\Gamma] \to \mathbb{C}$ is a *bounded* function (L^∞).

From the asymptotic expansion of the Hankel function at the origin, as in §6.1.3, we see that the kernel $\kappa(t) = \frac{\mathrm{i}}{4}H_0^{(1)}(kt) + \frac{1}{2\pi}\log\frac{t}{d}$ of $S - S_0$ is bounded (and continuous). The operator T defined by

[24] Recall that the property of compactness of an operator depends heavily on the norms of the function spaces chosen as domain and codomain. E.g. the identity operator $I : H^1(\Omega) \to H^1(\Omega)$ is not compact for a bounded Lipschitz Ω, but when we view it as $I : H^1(\Omega) \to L^2(\Omega)$ (and we call it embedding) then it is compact (Rellich theorem). The technique used in this section is very standard: we show that an operator K maps in a space that is slightly smoother (here $H^1(\partial B_R)$) than the desired domain (here $H^{\frac{1}{2}}(\partial B_R)$), then we compose the operator with the embedding (here $\iota : H^1(\partial B_R) \to H^{\frac{1}{2}}(\partial B_R)$) and if this embedding is compact the same holds for $\iota \circ K$. Recall that the composition of a continuous operator and a compact one is compact; you can prove this from the definition.

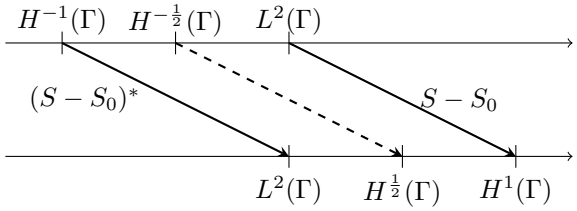
[25] More generally, if the K_ℓ are complex we can also take $\Re\{\sigma K_\ell\} > c(1+\ell^2)^s$ for some $\sigma \in \mathbb{C}$, $|\sigma| = 1$. E.g., K is coercive in $L^2(\partial B_R)$ ($s = 0$) if there exists a straight line in the complex plane $\mathbb{C}$ that separates the origin from all K_ℓ.

$Tv = (Sv - S_0 v)'$, where the derivative is the tangential derivative along Γ, also has a bounded (but discontinuous) kernel, for the same reason. From this it follows that $S - S_0 : L^2(\Gamma) \to H^1(\Gamma)$ is a compact operator. However we want to lower the Sobolev exponents of both spaces by $\frac{1}{2}$.

From functional analysis ([Brezis11, Thm. 6.4]) we know that if an operator $K : H_1 \to H_2$ is compact, then its adjoint $K^* : H_2^* \to H_1^*$ (defined by $(K^*\varphi)(\psi) = \varphi(K\psi)$ for $\varphi \in H_2^*$ and $\psi \in H_1$) is also compact.

Fubini's theorem implies that the single-layer is self-adjoint (up to conjugation) in $L^2(\Gamma)$: $\int_\Gamma (S\varphi)\psi \,\mathrm{d}s = \int_\Gamma \varphi(S\psi)\,\mathrm{d}s$ for all $\varphi, \psi \in L^2(\Gamma)$ (you can prove this). The same holds for S_0. Thus the adjoint of $S - S_0 : L^2(\Gamma) \to H^1(\Gamma)$ is the extension of $S - S_0$ itself to $(S - S_0)^* = S - S_0 : H^{-1}(\Gamma) \to L^2(\Gamma)$, where $H^{-1}(\Gamma)$ is the dual of $H^1(\Gamma)$ as in §3.3.3, and is compact.

A technique called "operator interpolation" (see, e.g., [SBH19, §9.8]) allows to deduce from the compactness of $S - S_0$ in $L^2(\Gamma) \to H^1(\Gamma)$ and in $H^{-1}(\Gamma) \to L^2(\Gamma)$ the boundedness and the compactness in all intermediate spaces $S - S_0 : H^{s-1}(\Gamma) \to H^s(\Gamma)$ for $0 < s < 1$, in particular the compactness of $S - S_0 : H^{-\frac{1}{2}}(\Gamma) \to H^{\frac{1}{2}}(\Gamma)$.



The same reasoning follows verbatim if we take S_c for $c > 0$ (64) in place of S_0 (63): the key point is the boundedness of the kernel of $S - S_c$ and of $v \mapsto (Sv - S_c v)'$. This follows from the fact that all kernels considered have the same singularity at 0. A different proof of the compactness of $S - S_c$ can be obtained by adapting that in [SBH19, Thm. 14.12].

6.1.5 Coercivity of S_c

We sketch the proof of the coercivity of the reaction–diffusion single-layer operator S_c. The coercivity of the Laplace single-layer operator S_0 is slightly more complicated; we discuss it in Remark 6.8.

Most of the results derived and stated in the previous sections (traces, jumps, ...) for the Helmholtz equation hold also for the reaction–diffusion equation, using $\Phi_c(\mathbf{x}, \mathbf{y}) = \frac{1}{2\pi} K_0(c|\mathbf{x} - \mathbf{y}|)$ as fundamental solution.

Fix $\psi \in H^{-\frac{1}{2}}(\Gamma)$ and denote $u = \mathcal{S}_c\psi \in H^1(\Omega_-, \Delta) \times H^1_{\mathrm{loc}}(\overline{\Omega_+}, \Delta)$, $\mathcal{S}_c$ being the reaction–diffusion single-layer potential. Then $-\Delta u + c^2 u = 0$ in $\Omega_- \cup \Omega_+$, $[\![\gamma u]\!] = 0$ and $[\![\partial_{\mathbf{n}} u]\!] = -\psi$, in analogy to (61). From the properties of K_0, both u and its radial derivative decay exponentially to 0 for $|\mathbf{x}| \to \infty$.

Integration by parts (Green's first identity (20) with $k = \mathrm{i}c$) on Ω_- gives the identity

$$\|\nabla u\|^2_{L^2(\Omega_-)} + c^2 \|u\|^2_{L^2(\Omega_-)} = \int_{\Omega_-} (\nabla u \cdot \nabla \overline{u} + c^2 u\overline{u}) \,\mathrm{d}\mathbf{x} = \int_\Gamma \partial_{\mathbf{n}}^- u \, \gamma^- \overline{u} \,\mathrm{d}s.$$

Proceeding similarly to §5.3 and exploiting the decay of u and $\frac{\partial u}{\partial r}$, the same holds in Ω_+:

$$\begin{aligned} \|\nabla u\|^2_{L^2(\Omega_+)} + c^2 \|u\|^2_{L^2(\Omega_+)} &= \lim_{R\to\infty} \Big(\|\nabla u\|^2_{L^2(\Omega_+ \cap B_R)} + c^2 \|u\|^2_{L^2(\Omega_+ \cap B_R)} \Big) \\ &= \lim_{R\to\infty} \int_{\Omega_+ \cap B_R} (\nabla u \cdot \nabla \overline{u} + c^2 u\overline{u}) \,\mathrm{d}\mathbf{x} \\ &= \lim_{R\to\infty} \Big(- \int_\Gamma \partial_{\mathbf{n}}^+ u \, \gamma^+ \overline{u} \,\mathrm{d}s + \underbrace{\int_{\partial B_R} \partial_{\mathbf{n}} u \, \overline{u} \,\mathrm{d}s}_{\xrightarrow{R\to\infty} 0} + \int_{\Omega_+ \cap B_R} \underbrace{(-\Delta u + c^2 u)}_{=0} \overline{u} \,\mathrm{d}\mathbf{x} \Big) = - \int_\Gamma \partial_{\mathbf{n}}^+ u \, \gamma^+ \overline{u} \,\mathrm{d}s. \end{aligned} \tag{66}$$

(Note that the $H^1(\Omega_+)$ norm would not be bounded if $u = \mathcal{S}\psi$, with the Helmholtz single-layer potential, because in this case u would have a slower decay at infinity.)

Taking the normal trace of vector fields, $\mathbf{v} \mapsto \mathbf{v}|_\Gamma \cdot \mathbf{n}$, is continuous as an operator from $H(\mathrm{div}; \Omega_\pm)$ to $H^{-\frac{1}{2}}(\Gamma)$ ([SBH19, Thm. 6.1]):

$$\|\mathbf{v}|_\Gamma \cdot \mathbf{n}\|^2_{H^{-\frac{1}{2}}(\Gamma)} \le C_{\mathrm{tr}} \big(\|\mathbf{v}\|^2_{L^2(\Omega_\pm)^2} + \|\mathrm{div}\, \mathbf{v}\|^2_{L^2(\Omega_\pm)} \big) \qquad \forall \mathbf{v} \in H(\mathrm{div}; \Omega_\pm) := \big\{ \mathbf{v} \in L^2(\Omega_\pm)^2 ; \mathrm{div}\, \mathbf{v} \in L^2(\Omega_\pm) \big\}. \tag{67}$$

Here $C_{\mathrm{tr}} > 0$ only depends on Γ. Applying this to ∇u, whose divergence is in L^2 because $\Delta u = c^2 u$, allows to control the Neumann traces with the L^2 norms of the gradient and the Laplacian of u. Combining all these results we have that, for all $\psi \in H^{-\frac{1}{2}}(\Gamma)$,

$$\|\psi\|^2_{H^{-\frac{1}{2}}(\Gamma)} = \|[\![\partial_{\mathbf{n}} u]\!]\|^2_{H^{-\frac{1}{2}}(\Gamma)} \qquad \psi = -[\![\partial_{\mathbf{n}} u]\!]$$

$$
\begin{aligned}
&\le 2\big(\|\partial_{\mathbf{n}}^- u\|^2_{H^{-\frac12}(\Gamma)} + \|\partial_{\mathbf{n}}^+ u\|^2_{H^{-\frac12}(\Gamma)} \big) && \text{triangle inequality}\\
&\le 2C_{\mathrm{tr}}\big(\|\nabla u\|^2_{L^2(\Omega_-)^2} + \|\Delta u\|^2_{L^2(\Omega_-)} + \|\nabla u\|^2_{L^2(\Omega_+)^2} + \|\Delta u\|^2_{L^2(\Omega_+)} \big) && \text{normal trace continuity in } H(\mathrm{div};\Omega_\pm)\ (67)\\
&\le 2C_{\mathrm{tr}} \max\{1,c^2\} \int_{\Omega_-\cup\Omega_+} (\nabla u\cdot\nabla\overline{u} + c^2 u\overline{u})\,\mathrm{d}\mathbf{x} && \Delta u = c^2 u\\
&= 2C_{\mathrm{tr}} \max\{1,c^2\} \int_\Gamma (\partial_{\mathbf{n}}^- u\gamma^-\overline{u} - \partial_{\mathbf{n}}^+ u\gamma^+\overline{u})\,\mathrm{d}s && \text{Green's 1st identity } (66)\\
&= 2C_{\mathrm{tr}} \max\{1,c^2\} \int_\Gamma -[\![\partial_{\mathbf{n}} u]\!]\gamma\overline{u}\,\mathrm{d}s && \gamma^+ u = \gamma^- u\\
&= 2C_{\mathrm{tr}} \max\{1,c^2\} \int_\Gamma \psi\overline{S_c\psi}\,\mathrm{d}s && [\![\partial_{\mathbf{n}} u]\!] = -\psi,\quad \gamma u = \gamma\mathcal{S}_c\psi = S_c\psi\\
&= 2C_{\mathrm{tr}} \max\{1,c^2\} \int_\Gamma (S_c\psi)\overline{\psi}\,\mathrm{d}s && \overline{S_c\psi} = S_c\overline{\psi},\quad \int_\Gamma \psi S_c\phi\,\mathrm{d}s = \int_\Gamma (S_c\psi)\phi\,\mathrm{d}s.
\end{aligned}
$$

This is precisely the coercivity of S_c in $H^{-\frac12}(\Gamma)$.

Remark 6.8: (Coercivity of S_0 in 3D and 2D). The proof of the coercivity of the Laplace single-layer operator S_0 in $\mathbb{R}^3$ works exactly in the same way setting $c = 0$, noting that the algebraic decay of u for $|\mathbf{x}| \to \infty$ is enough to ensure $\nabla u \in H^1(\Omega_+)$ (see [SBH19, §14.6–7])

In $\mathbb{R}^2$ the proof of the coercivity of S_0 is more complicated; here we follow [Steinbach 2008, Thm. 6.22–23] and [McLean 2000, Thm. 8.12–16]. The key difference is that the relation $\|\nabla u\|^2_{L^2(\Omega_+)} = -\int_\Gamma \partial_{\mathbf{n}}^+ u\,\gamma^+\overline{u}\,\mathrm{d}s$ for $u = S_0\psi$ only holds if ψ satisfies the 0-average condition $\langle\psi, 1\rangle_\Gamma = 0$. This is due to the absence of decay to zero of the fundamental solution Φ_0 at infinity. Thus the reasoning above gives the coercivity of S_0 in $H^{-\frac12}_*(\Gamma) := \{\psi \in H^{-\frac12}(\Gamma) : \langle\psi, 1\rangle_\Gamma = 0\}$.

To deal with the general case ($\psi \in H^{-\frac12}(\Gamma)$ instead of $\psi \in H^{-\frac12}_*(\Gamma)$) we have to work a bit more. Lax–Milgram and the coercivity for 0-average densities ensure that there exists a unique (non zero)

$$
\beta_* \in H^{-\frac12}_*(\Gamma) \qquad \langle S_0\beta_*, \xi_*\rangle_\Gamma = \langle S_0 1, \xi_*\rangle_\Gamma \qquad \forall \xi_* \in H^{-\frac12}_*(\Gamma).
$$

Define $\beta_{\mathrm{eq}} := \frac{1}{|\Gamma|}(1 - \beta_*) \in H^{-\frac12}(\Gamma)$. The two densities $\beta_*, \beta_{\mathrm{eq}}$ only depend on Γ. Then β_{eq} is real-valued, $\langle\beta_{\mathrm{eq}}, 1\rangle_\Gamma = 1$, and $\langle S_0\beta_{\mathrm{eq}}, \xi_*\rangle_\Gamma = 0$ for all $\xi_* \in H^{-\frac12}_*(\Gamma)$, so $S_0\beta_{\mathrm{eq}}$ is constant on Γ. Recalling the definition (63) of S_0,

$$
\begin{aligned}
(S_0\beta_{\mathrm{eq}})(\mathbf{x}) &= \frac{1}{2\pi}\int_\Gamma \beta_{\mathrm{eq}}(\mathbf{y})\big(\log d - \log|\mathbf{x}-\mathbf{y}|\big)\,\mathrm{d}s(\mathbf{y})\\
&= \frac{1}{2\pi}\log d \int_\Gamma \beta_{\mathrm{eq}}(\mathbf{y})\,\mathrm{d}s(\mathbf{y}) - \frac{1}{2\pi}\int_\Gamma \beta_{\mathrm{eq}}(\mathbf{y})\log|\mathbf{x}-\mathbf{y}|\,\mathrm{d}s(\mathbf{y})\\
&= \frac{1}{2\pi}\log d - \frac{1}{2\pi}\int_\Gamma \beta_{\mathrm{eq}}(\mathbf{y})\log|\mathbf{x}-\mathbf{y}|\,\mathrm{d}s(\mathbf{y}) \qquad \forall \mathbf{x}\in\Gamma.
\end{aligned}
$$

This is where we need the (so far unused) parameter $d > 0$: if d is sufficiently large then $S_0\beta_{\mathrm{eq}} > 0$. It is possible to show that $d > \operatorname{diam}\Gamma$ is enough to guarantee that $S_0\beta_{\mathrm{eq}} > 0$.[26] Then also

$$
\langle S_0\beta_{\mathrm{eq}}, \beta_{\mathrm{eq}}\rangle_\Gamma = S_0\beta_{\mathrm{eq}}\langle 1, \beta_{\mathrm{eq}}\rangle_\Gamma = S_0\beta_{\mathrm{eq}} > 0.
$$

We want to decompose a general $\psi \in H^{-\frac12}(\Gamma)$ in a $H^{-\frac12}_*(\Gamma)$ component and a remainder: instead of taking a constant remainder as one might expect, we take a remainder whose image under S_0 is constant, i.e. a multiple of β_{eq}. For all $\psi \in H^{-\frac12}(\Gamma)$ define

$$
\psi_* := \psi - \langle\psi, 1\rangle_\Gamma\beta_{\mathrm{eq}} \quad\Rightarrow\quad \langle\psi_*, 1\rangle_\Gamma = \langle\psi, 1\rangle_\Gamma\Big(1 - \langle\beta_{\mathrm{eq}}, 1\rangle_\Gamma\Big) = 0.
$$

Now we use the decomposition $\psi = \psi_* + \langle\psi, 1\rangle_\Gamma\beta_{\mathrm{eq}}$ to show the positivity of the single-layer potential:

$$
\langle S_0\psi, \psi\rangle_\Gamma = \langle S_0(\psi_* + \langle\psi, 1\rangle_\Gamma\beta_{\mathrm{eq}}), \psi_* + \langle\psi, 1\rangle_\Gamma\beta_{\mathrm{eq}}\rangle_\Gamma
$$

[26] The value $\mathrm{e}^{\int_\Gamma \beta_{\mathrm{eq}}(\mathbf{y})\log|\mathbf{x}-\mathbf{y}|\,\mathrm{d}s(\mathbf{y})}$, which is independent of $\mathbf{x} \in \Gamma$, is called "logarithmic capacity of Γ", while β_{eq} is the "equilibrium density". In 2D electrostatic, the electric charge on an insulated conductor Ω_- distributes on the boundary Γ proportionally to β_{eq}, in such a way that the electrostatic potential $S_0\beta_{\mathrm{eq}}$ is constant on Γ and takes value $\mathcal{S}_0\beta_{\mathrm{eq}}$ in Ω_+.

$$= \langle S_0\psi_*, \psi_*\rangle_\Gamma + 2\langle\psi,1\rangle_\Gamma \underbrace{\langle S_0\beta_{\rm eq}, \psi_*\rangle_\Gamma}_{=0} + |\langle\psi,1\rangle_\Gamma|^2 \underbrace{\langle S_0\beta_{\rm eq}, \beta_{\rm eq}\rangle_\Gamma}_{>0} \geq C\Big(\|\psi_*\|^2_{H^{-\frac12}(\Gamma)} + |\langle\psi,1\rangle_\Gamma|^2\Big),$$

which gives coercivity when combined with the triangle inequality

$$\|\psi\|_{H^{-\frac12}(\Gamma)} \leq \|\psi_*\|_{H^{-\frac12}(\Gamma)} + |\langle\psi,1\rangle_\Gamma|\,\|\beta_{\rm eq}\|_{H^{-\frac12}(\Gamma)} \leq C\Big(\|\psi_*\|_{H^{-\frac12}(\Gamma)} + |\langle\psi,1\rangle_\Gamma|\Big).$$

Exercise 6.9: (Laplace single-layer parameter d). Using the BIO expansion in §6.1.3 show that for a circle $\Gamma = \partial B_R$, we have $\beta_* = 0$, $\beta_{\rm eq} = \frac{1}{2\pi R}$, $S_0\beta_{\rm eq} = \frac{1}{2\pi R}S_0 1 = \frac{1}{2\pi}\log\frac{d}{R}$. Thus $d > R$ is enough to prove the coercivity of S_0 on $H^{-\frac12}(\partial B_R)$.

Hint: use that $\int_0^{2\pi}\log\sqrt{2-2\cos\alpha}\,\mathrm{d}\alpha = 0$ and the properties of the logarithm.

6.2 The BIE zoo

We have seen that the single-layer BIE fails for some values of k. We want to derive some other BIEs that allow to compute the solution of the EDP/SSSP also for these values of k. We write a total of six BIEs; their properties are summarised in Table 1.

First of all, it is instructive to recall how we found the BIE (42). We wrote the solution u of the EDP (32) as a single-layer $u = \mathcal{S}\psi$, then we took the Dirichlet trace γ^+ of this representation, and from one of the trace formulas (60) obtained the BIE $S\psi = g_D$ (recall that we need to impose the boundary condition $\gamma^+ u = g_D$). Also for the other BIEs the key steps are the same:

(i) choose a potential representation,

(ii) take a trace using (60).

6.2.1 Indirect double-layer BIE

If instead of a single-layer we assume that the EDP solution is a double-layer potential

$$u = \mathcal{D}\psi, \qquad \psi \in H^{\frac12}(\Gamma),$$

taking the Dirichlet trace γ^+ (60) we obtain

$$\boxed{\Big(\frac12 + D\Big)\psi = g_D} \quad \text{in } H^{\frac12}(\Gamma), \qquad \psi \in H^{\frac12}(\Gamma). \tag{68}$$

This is another BIE for the same BVP. Here and in the following, $\frac12$ stands for the identity operator multiplied by the number $\frac12$, i.e. the equation is to be read $\frac12\psi + D\psi = g_D$.

This can be discretised with collocation-BEM or Galerkin-BEM in the same way as §5.2. We encounter a couple of difficulties. A first difference is that the singularity of D is stronger than that of S, so the quadrature requires more care.

A second difference is that this BIE is posed in $H^{\frac12}(\Gamma)$ instead of $H^{-\frac12}(\Gamma)$. The functions of $H^{\frac12}(\Gamma)$ are in general not necessarily continuous, but if they are piecewise-polynomial then they must also be continuous. This implies that the BEM discrete space V_N cannot be made of piecewise-constant functions (recall Exercise 3.4). The simplest choice is to take V_N as the space of **continuous piecewise-linear functions** on a mesh.

Is the BIE (68) well-posed? I.e. is the operator $(\frac12 + D) : H^{\frac12}(\Gamma) \to H^{\frac12}(\Gamma)$ invertible? As proved in [CGLS12, Thm. 2.25], this operator is Fredholm. To study the injectivity, once again we have to look at some interior problem.

Exercise 6.10: (Injectivity of $\frac12 + D$). Show the following.

- If w is a Neumann eigenfunction in Ω_- for $\Lambda = k^2$, then its trace $\psi = \gamma^- w$ satisfies $\frac12\psi + D\psi = 0$.

  Hint: use Green's representation.
- If $\Lambda = k^2$ is not a Neumann eigenvalue in Ω_-, then $\frac12 + D$ is injective.

  Hint: take $u = \mathcal{D}\psi$ for $\psi \in \ker(\frac12 + D)$. Use the well-posedness of the EDP and both jump relations.
- Deduce that the BIE (68) is injective if and only if k^2 is not a Neumann eigenfunction.

From this exercise it follows that the BIE (68) is well-posed except for a discrete set of frequencies.

Remark 6.11: (What is ψ?). If ψ is solution of (68), then $u^- = (\mathcal{D}\psi)|_{\Omega_-}$ is Helmholtz solution in Ω_- with Neumann trace $\partial_{\mathbf{n}}^- u^- = \partial_{\mathbf{n}}^+ u$. Differently from §5.5.1, this is not immediately related to the incoming field u^{Inc}. However, if k^2 is not a Neumann eigenvalue, u^- is well-defined as a solution of an interior Neumann problem. From the jump relation (61), $\psi = [\![\gamma\mathcal{D}\psi]\!] = \gamma^+ u - \gamma^- u^- = g_D - \gamma^- u^-$. The solution of the BIE (68) is the difference between the datum g_D and the Dirichlet trace of the solution of an auxiliary interior Neumann problem, whose boundary datum is the Neumann trace of u itself.

6.2.2 Direct BIE

We have constructed two BIEs (42) and (68) by searching for EDP solutions in the form $u = \mathcal{S}\psi$ and $u = \mathcal{D}\psi$, respectively. Green's representation formula (54) allows to represent any radiating solution u in Ω_+ as linear combination of both potentials applied to the traces of u itself:

$$u = -\mathcal{S}\partial_{\mathbf{n}}^+ u + \mathcal{D}\gamma^+ u.$$

When u is solution of the EDP, one of the traces is given: $\gamma^+ u = g_D$. So we can choose as unknown the other one: $\psi = \partial_{\mathbf{n}}^+ u$. How to obtain a BIE from this?

Taking the Dirichlet and the Neumann traces γ^+ and $\partial_{\mathbf{n}}^+$ of Green's representation and using the trace formulas (60), we obtain

$$\boxed{S\psi = \left(D - \frac{1}{2}\right) g_D} \qquad \text{in } H^{\frac{1}{2}}(\Gamma), \tag{69}$$

$$\boxed{\left(\frac{1}{2} + D'\right)\psi = H g_D} \qquad \text{in } H^{-\frac{1}{2}}(\Gamma). \tag{70}$$

We reiterate that here the unknown stands for the Neumann datum and the underlying representation formula is that coming from Green's formula:

$$\psi = \partial_{\mathbf{n}}^+ u \quad \in H^{-\frac{1}{2}}(\Gamma), \qquad u = -\mathcal{S}\psi + \mathcal{D} g_D \quad \text{in } \Omega_+.$$

Some terminology. BIEs such as (69)–(70) whose unknown is the missing Cauchy datum are called **direct** BIEs; BIEs such as (42) and (68) where the unknown is not directly linked to the EDP are called **indirect**. In general, a linear BIE with BIO T and data f can be written as $\alpha\psi + T\psi = f$: if $\alpha = 0$ then the BIE is called "of the first kind", if α is a non-zero coefficient then it is called "of the second kind". Thus BIEs (42) and (69) are called BIEs of the **first kind**, as the unknown ψ only appears as argument of a BIO, while (68) and (70) are called of the **second kind** as the unknown ψ also appears outside the integral operator ($\alpha = \frac{1}{2}$ in both cases).

The first-kind direct BIE (69) has at the left-hand side the same operator S as the indirect BIE (42) we know well. The right-hand side instead is slightly more complicated, as it involves the double-layer operator. So (69) is well-posed exactly when (42) is, i.e. away from Dirichlet eigenvalues. The matrix $\underline{\underline{\mathbf{A}}}^{C/G}$ of a BEM implementation for this formulation is identical to the matrix for the same method applied to (42). The right-hand side vector $\mathbf{F}^{C/G}$ is slightly more complicated to code and more expensive to compute as it requires the implementation of the double-layer operator.

The second-kind direct BIE (70) has at the left-hand side the *adjoint* of the operator present in the indirect BIE (68). Theorem 1.28 of [CK1] implies that $(\frac{1}{2} + D')$ is injective if and only if $(\frac{1}{2} + D)$ is injective (this requires the Fredholm property of the operators). So (70) is well-posed away from Neumann eigenvalues, exactly as (68). The implementation of a BEM discretisation of (70) also requires an approximation of the hypersingular operator H (59) for the right-hand side.

Finally, the evaluation of u_N in the volume through the Green's representation formula is slightly more complicated for the direct BIEs (69)–(70) than for the indirect ones ((42) and (68)), because it involves both the single- and the double-layer potential (applied to the unknown and the datum, respectively).

6.2.3 Indirect combined-field integral equation: Brakhage–Werner equation

We have seen four different BIEs for the same EDP, and none of them is invertible for *all* positive values of k, which is quite disappointing. However all is not lost: the formulations considered were deduced from some special representations of u in terms of layer potentials, we need to choose some other such representation.

We now choose, arbitrarily, to search for some u in the form

$$u = (\mathcal{D} - \mathrm{i}\eta\mathcal{S})\psi, \qquad \psi \in H^{\frac{1}{2}}(\Gamma) \tag{71}$$

where $\eta > 0$ is a parameter. Taking the Dirichlet trace, this is solution of the EDP if ψ is solution of

$$\boxed{\Big(\frac{1}{2} + D - \mathrm{i}\eta S\Big)\psi = g_D} \quad \text{in } H^{\frac{1}{2}}(\Gamma). \tag{72}$$

The operator $A := (\frac{1}{2} + D - \mathrm{i}\eta S) : H^{\frac{1}{2}}(\Gamma) \to H^{\frac{1}{2}}(\Gamma)$ is Fredholm.[27] Is it injective?

Let $A\psi = 0$ for some $\psi \in H^{\frac{1}{2}}(\Gamma)$. Define u as in (71). Then $u|_{\Omega_+}$ is solution of the EDP with $g_D = 0$, so $u = 0$ in Ω_+ and $\gamma^+ u = \partial_{\mathbf{n}}^+ u = 0$. The jump relations (61) give

$$\begin{aligned} -\gamma^- u &= [\![\gamma u]\!] = [\![\gamma(\mathcal{D}\psi - \mathrm{i}\eta\mathcal{S}\psi)]\!] = [\![\gamma\mathcal{D}\psi]\!] = \psi \\ -\partial_{\mathbf{n}}^- u &= [\![\partial_{\mathbf{n}} u]\!] = [\![\partial_{\mathbf{n}}(\mathcal{D}\psi - \mathrm{i}\eta\mathcal{S}\psi)]\!] = [\![-\mathrm{i}\eta\partial_{\mathbf{n}}\mathcal{S}\psi]\!] = \mathrm{i}\eta\psi \end{aligned} \qquad \Rightarrow \qquad \partial_{\mathbf{n}}^- u - \mathrm{i}\eta\gamma^- u = 0.$$

So $u|_{\Omega_-}$ is solution of a homogeneous impedance BVP (27) in Ω_- with $\vartheta = \frac{\eta}{k}$. From the well-posedness of the impedance BVP (§4.2 and Corollary 5.13) $u = 0$ and from the jump relation again $\psi = -\gamma^- u = 0$. We conclude that the operator A is injective.

The BIE (72) is well-posed for all Γ, $k > 0$, $\eta > 0$ and $g_D \in H^{\frac{1}{2}}(\Gamma)$.

We have finally found a BIE that is invertible for all wavenumbers! The BIE (72) is often called **Brakhage–Werner equation** (even if it was introduce independently in three papers by Brakhage and Werner, by Leis and by Panič, all in 1965).

Exercise 6.12: (Brakhage–Werner density). Let ψ be the solution of the Brakhage–Werner BIE (72) and $u = (\mathcal{D} - \mathrm{i}\eta\mathcal{S})\psi$ in $\Omega_+ \cup \Omega_-$. Use the trace formulas (60) to show that $\partial_{\mathbf{n}}^- u - \mathrm{i}\eta\gamma^- u = \partial_{\mathbf{n}}^+ u - \mathrm{i}\eta\gamma^+ u$ (careful with the signs!).

Deduce, using the jump relations, that the solution ψ of the BIE is the jump between the Dirichlet traces of the EDP solution $u|_{\Omega_+}$ and the solution $u|_{\Omega_-}$ of an impedance BVP in Ω_- with data $\partial_{\mathbf{n}}^+ u - \mathrm{i}\eta\gamma^+ u$ and impedance parameter $\vartheta = \frac{\eta}{k}$. (The relation with an impedance BVP further confirms the well-posedness of the BIE.)

6.2.4 Direct combined-field integral equation: Burton–Miller equation

Can we find a *direct* method that is well-posed for all values of k? We know that $\psi = \partial_{\mathbf{n}}^+ u$ solves both direct equations (69)–(70). We take a linear combination of the two equations:

$$\boxed{\Big(\frac{1}{2} + D' - \mathrm{i}\eta S\Big)\psi = \Big[H - \mathrm{i}\eta\Big(D - \frac{1}{2}\Big)\Big] g_D} \quad \text{in } H^{-\frac{1}{2}}(\Gamma). \tag{73}$$

Again, here $\eta > 0$ is a parameter. This is called **Burton–Miller** or (direct) **combined-field** integral equation (CFIE). This is a second-kind direct equation, so, as in §6.2.2, the density and the representation formula are

$$\psi = \partial_{\mathbf{n}}^+ u \quad \in H^{-\frac{1}{2}}(\Gamma) \qquad \text{and} \qquad u = -\mathcal{S}\partial_{\mathbf{n}}^+ u + \mathcal{D}\gamma^+ u = -\mathcal{S}\psi + \mathcal{D}g_D \quad \text{in } \Omega_+.$$

The operator $A' := (\frac{1}{2} + D' - \mathrm{i}\eta S) : H^{-\frac{1}{2}}(\Gamma) \to H^{-\frac{1}{2}}(\Gamma)$ at the left-hand side differs from the operator A of the Brakhage–Werner equation only in that D is substituted by D'.

To study the injectivity of A', let $A'\psi = 0$ and $u = -\mathcal{S}\psi$. Then $\partial_{\mathbf{n}}^- u - \mathrm{i}\eta\gamma^- u = -A'\psi = 0$, so $u = 0$ in Ω_- by the well-posedness of the homogeneous interior impedance BVP. By the jump formula $\gamma^+ u = \gamma^- u$, $u|_{\Omega_+}$ is solution of the homogeneous EDP, so it also vanish, and $\psi = -[\![\partial_{\mathbf{n}}\mathcal{S}\psi]\!] = [\![u]\!] = 0$.

Similarly to the previous section, A' is also Fredholm, thus the BIE (73) is well-posed.

Under some strong regularity and convexity assumptions on the scatterer Ω_-, the operators A and A' are coercive in $L^2(\Gamma)$, [CGLS12, §5.7, Thm. 5.25]: in these cases all Galerkin BEM discretisations of the formulations (72) and (73) are well-posed and quasi-optimal.

Remark 6.13: (Combined-field integral equation for scattering problems [CGLS12, §2.10]). Imagine that we want to solve a scattering problem (33) with an incoming wave u^{Inc} defined over a neighbourhood of $\overline{\Omega_-}$ (for example a plane wave).

[27] We have mentioned in §6.2.1, without proof, that $\frac{1}{2} + D$ is Fredholm. The single-layer operator is continuous as a map $S : H^{\frac{1}{2}}(\Gamma) \to H^1(\Gamma)$ ([CGLS12, Thm. 2.25], recall Remark 5.1 for the circle), thus, by composition with the embedding of Sobolev spaces, it is compact as a map $S : H^{\frac{1}{2}}(\Gamma) \to H^{\frac{1}{2}}(\Gamma)$. Finally, A is a compact perturbation of a Fredholm operator, so it is Fredholm as well.

Green's representation (47)–(50) applied to u^{Scat} in Ω_+ and to u^{Inc} in Ω_- gives

$$\mathcal{D}\gamma^+u^{\mathrm{Scat}} - \mathcal{S}\partial_{\mathbf{n}}^+u^{\mathrm{Scat}} = u^{\mathrm{Scat}}, \qquad \mathcal{D}\gamma^+u^{\mathrm{Inc}} - \mathcal{S}\partial_{\mathbf{n}}^+u^{\mathrm{Inc}} = \mathcal{D}\gamma^-u^{\mathrm{Inc}} - \mathcal{S}\partial_{\mathbf{n}}^-u^{\mathrm{Inc}} = 0 \qquad \text{in } \Omega_+.$$

Summing these two equalities and using that $\gamma^+u^{\mathrm{Tot}} = 0$, we have

$$\mathcal{S}(-\partial_{\mathbf{n}}^+u^{\mathrm{Tot}}) = \mathcal{D}\gamma^+u^{\mathrm{Tot}} - \mathcal{S}\partial_{\mathbf{n}}^+u^{\mathrm{Tot}} = u^{\mathrm{Scat}} = u^{\mathrm{Tot}} - u^{\mathrm{Inc}} \qquad \text{in } \Omega_+.$$

(We already knew this from §5.5.1 in the case where k^2 is not a Dirichlet eigenvalue, now it is clear that this assumption is not needed.) The exterior Dirichlet and Neumann traces of this identity give

$$-S\partial_{\mathbf{n}}^+u^{\mathrm{Tot}} \overset{(60)}{=} \gamma^+\mathcal{S}(-\partial_{\mathbf{n}}^+u^{\mathrm{Tot}}) = \gamma^+u^{\mathrm{Tot}} - \gamma^+u^{\mathrm{Inc}} \overset{(33)}{=} -\gamma^+u^{\mathrm{Inc}},$$
$$\Big(\frac{1}{2} - D'\Big)\partial_{\mathbf{n}}^+u^{\mathrm{Tot}} \overset{(60)}{=} \partial_{\mathbf{n}}^+\mathcal{S}(-\partial_{\mathbf{n}}^+u^{\mathrm{Tot}}) = \partial_{\mathbf{n}}^+u^{\mathrm{Tot}} - \partial_{\mathbf{n}}^+u^{\mathrm{Inc}} \qquad \Rightarrow \qquad \Big(\frac{1}{2} + D'\Big)\partial_{\mathbf{n}}^+u^{\mathrm{Tot}} = \partial_{\mathbf{n}}^+u^{\mathrm{Inc}}.$$

As before, we linearly combine these two identities using a coupling parameter $\eta > 0$:

$$\boxed{\Big(\frac{1}{2} + D' - \mathrm{i}\eta S\Big)\psi = \partial_{\mathbf{n}}^+u^{\mathrm{Inc}} - \mathrm{i}\eta\,\gamma^+u^{\mathrm{Inc}}, \qquad \psi = \partial_{\mathbf{n}}^+u^{\mathrm{Tot}}.} \tag{74}$$

This is a variant of the Burton–Miller formulation (73): since the integral operator to be inverted is the same (A'), the BIE is well-posed for all $k > 0$.

Equation (74) is simpler to implement than (73) because the right-hand side does not need any integral operator. On the other hand, the right-hand side involves *both traces* γ^+u^{Inc} and $\partial_{\mathbf{n}}^+u^{\mathrm{Inc}}$ of the incoming field, so this BIE cannot be used if only the Dirichlet trace $g_D = -\gamma^+u^{\mathrm{Inc}}$ is known.

Another important difference is that the unknown in (74) is $\psi = \partial_{\mathbf{n}}^+u^{\mathrm{Tot}}$, as opposed to $\partial_{\mathbf{n}}^+u^{\mathrm{Scat}}$ in (73). Thus the corresponding representation formula is simply $\boxed{u^{\mathrm{Scat}} = -\mathcal{S}\psi}$.

A modification of this BIE that is continuous and coercive in $L^2(\Gamma)$ for any bounded star-shaped Lipschitz domain has been recently discovered, see [CGLS12, §2.9].

Remark 6.14: (Choice of the parameter η). How to choose the coupling parameter $\eta > 0$ in (72) and (73)? From the expression of A and A', we can guess that η has the dimension of the inverse of a length: η multiplies the operator S which acts as the inverse of a derivation ($S : H^{-\frac{1}{2}}(\Gamma) \to H^{\frac{1}{2}}(\Gamma)$) and is added to the identity. So plausible choices are $\eta \sim k$ or $\eta \sim \frac{1}{\operatorname{diam}(\Gamma)}$. It turns out that $\eta = k$ is also a good choice to reduce the condition number of a BEM discretisation of either (72) or (73) for large values of k.

6.2.5 Further remarks on the BIEs and extensions

Remark 6.15: (Variational formulations of II kind BIEs). In the first-kind equations (42) and (69) the operator to be inverted is S, which maps $H^{-\frac{1}{2}}(\Gamma)$ (the space where we look for the unknown) to its dual $H^{\frac{1}{2}}(\Gamma)$. So testing the BIEs against elements of the same space is simple: the sesquilinear form $\mathcal{A}_S(\psi,\xi) = \langle S\psi, \xi\rangle_\Gamma$ is well-defined for $\psi, \xi \in H^{-\frac{1}{2}}(\Gamma)$ and involves the extension $\langle\cdot,\cdot\rangle_\Gamma$ of the $L^2(\Gamma)$ scalar product. This is why in the implementation of the Galerkin-BEM matrix $\underline{\underline{\mathbf{A}}}^{\mathrm{Gal}}$ we are allowed to use integrals over Γ (recall that we chose piecewise-polynomial basis functions, which are in $L^2(\Gamma)$).

For the second-kind integral equations we have to be more careful. E.g., in (73), the operator A' maps $H^{-\frac{1}{2}}(\Gamma)$ to itself. So, we cannot write $\langle A'\psi, \xi\rangle_\Gamma$ for $\psi, \xi \in H^{-\frac{1}{2}}(\Gamma)$ because it is not defined, but should use the sesquilinear form $(A'\psi, \xi)_{H^{-\frac{1}{2}}(\Gamma)}$, where $(\cdot,\cdot)_{H^{-\frac{1}{2}}(\Gamma)}$ is the scalar product in $H^{-\frac{1}{2}}(\Gamma)$. Implementing a BEM discretisation of this variational problem is hard, as it requires to evaluate the non-local $H^{-\frac{1}{2}}(\Gamma)$ scalar product.[28]

On the other hand, if the EDP datum g_D is at least in $H^1(\Gamma)$, which is the case for smooth incoming waves, then the right-hand side $[H - \mathrm{i}\eta(D - \frac{1}{2})]g_D \in L^2(\Gamma)$ so we can use the variational formulation

$$(A'\psi,\xi)_{L^2(\Gamma)} = \int_\Gamma \Big(\frac{1}{2} + D' - \mathrm{i}\eta S\Big)\psi\,\overline{\xi}\,\mathrm{d}s = \int_\Gamma \Big[H - \mathrm{i}\eta\Big(D - \frac{1}{2}\Big)\Big]g_D\overline{\xi}\,\mathrm{d}s \qquad \forall \xi \in L^2(\Gamma),$$

and look for $\psi \in L^2(\Gamma)$. This is well-posed because A' is Fredholm also as a mapping $A' : L^2(\Gamma) \to L^2(\Gamma)$, [CGLS12, Thm. 2.27]. The Galerkin-BEM is then implemented with the same techniques of §5.2.

[28] On a circular boundary $\Gamma = \mathbb{S}^1$, this would be easier as the scalar product is computed from the circular harmonics expansion: $(\sum_{\ell\in\mathbb{Z}} \hat{v}_\ell \mathrm{e}^{\mathrm{i}\ell\theta}, \sum_{\ell\in\mathbb{Z}} \hat{w}_\ell \mathrm{e}^{\mathrm{i}\ell\theta})_{H^{-\frac{1}{2}}(\mathbb{S}^1)} = 2\pi\sum_{\ell\in\mathbb{Z}} \hat{v}_\ell\overline{\hat{w}_\ell}(1+\ell^2)^{-\frac{1}{2}}$ (recall definition (19)). Still, unless the Fourier coefficients of all the functions involved are already known, this is more complicated than the simple integral appearing in the $\langle\cdot,\cdot\rangle_\Gamma$ duality.

An alternative would be to use a Petrov–Galerkin approach: take test functions in a space different from the trial space. For example one can seek $\psi \in H^{-\frac{1}{2}}(\Gamma)$ such that $\langle A'\psi, \xi\rangle_\Gamma = \langle [H - \mathrm{i}\eta(D - \frac{1}{2})]g_D, \xi\rangle_\Gamma$ for all $\xi \in H^{\frac{1}{2}}(\Gamma)$. The Galerkin-BEM discretisation of this problem can use piecewise-constant trial functions ψ_N and continuous piecewise-linear test functions ξ_N. The two discrete spaces have to be defined on different compatible "dual" meshes.

All this applies also to the second-kind direct BIE (70).

On the other hand, the indirect BIEs (68) and (72) are posed in $H^{\frac{1}{2}}(\Gamma)$, so we can use the $L^2(\Gamma)$ scalar product to write the variational form ($\int_\Gamma A\psi\,\overline{\xi}\,\mathrm{d}s = \int_\Gamma g_D\overline{\xi}\,\mathrm{d}s$ for all $\xi \in H^{\frac{1}{2}}(\Gamma)$ for (72)) but we need to use discrete functions that are continuous on Γ, e.g. piecewise-linear polynomials.

Remark 6.16: (Advantages of direct formulations). The BEM approximation of Burton–Miller equation (73) is slightly more complicated and expensive than Brakhage–Werner (72), as the right-hand side involves two BIOs. Similarly, the direct equations (69)&(70) require a more complicated right-hand side and representation formula than (42)&(68). What is the advantage of a direct formulation against an indirect one?

In a direct formulation, if we have some information on the properties of the EDP solution u we can include it in the design of the approximating space V_N to improve its accuracy and efficiency. For instance, in some situations, PDE theory and high-frequency asymptotics (recall the GTD mentioned in Remark 5.30) permit to estimate the location and the strength of the singularities of $\psi = \partial_{\mathbf{n}}^+ u^{\mathrm{Scat}}$, its oscillations, the different behaviour in the shadow and the illuminated parts of Γ; see, e.g., Figure 18[29]. This knowledge allows to construct discrete spaces V_N that ensure high accuracy with small numbers of DOFs. This is the basic idea underlying the "hybrid numerical-asymptotic" (HNA) methods that are the main concern of [CGLS12].

In an indirect method, the BIE solution ψ depends also on the trace of some interior problem (see Rem 6.11 and Ex. 6.12), so its efficient approximation would require also the knowledge of the corresponding eigenfunction, which is not directly related to the physical scattering problem and might contain expensive-to-approximate "unphysical" singularities.

Moreover, often the quantity of interest is not u^{Scat} or u^{Tot} in Ω_+, but something dependent on $\partial_{\mathbf{n}}^+ u^{\mathrm{Scat}}$, such as the far-field pattern u_∞ (34). This is easily and accurately computed with a direct method or with (42), but not with (68) and (72).

	BIE	representation formula	density $\psi =$	unknown ψ in	direct/ indirect	kind	fails for
(42)	$S\psi = g_D$	$u = \mathcal{S}\psi$	$-\partial_{\mathbf{n}}^+ u^{\mathrm{Tot}}$	$H^{-\frac{1}{2}}(\Gamma)$	indirect	I	Dir. eig.
(68)	$(\frac{1}{2} + D)\psi = g_D$	$u = \mathcal{D}\psi$	Rem.6.11	$H^{\frac{1}{2}}(\Gamma)$	indirect	II	Neum. eig.
(69)	$S\psi = (D - \frac{1}{2})g_D$	$u = \mathcal{D}g_D - \mathcal{S}\psi$	$\partial_{\mathbf{n}}^+ u$	$H^{-\frac{1}{2}}(\Gamma)$	direct	I	Dir. eig.
(70)	$(\frac{1}{2} + D')\psi = Hg_D$	$u = \mathcal{D}g_D - \mathcal{S}\psi$	$\partial_{\mathbf{n}}^+ u$	$H^{-\frac{1}{2}}(\Gamma)$	direct	II	Neum. eig.
(72)	$(\frac{1}{2} + D - \mathrm{i}\eta S)\psi = g_D$	$u = (\mathcal{D} - \mathrm{i}\eta\mathcal{S})\psi$	Ex. 6.12	$H^{\frac{1}{2}}(\Gamma)$	indirect	II	never!
(73)	$(\frac{1}{2}+D'-\mathrm{i}\eta S)\psi = [H-\mathrm{i}\eta(D-\frac{1}{2})]g_D$	$u = \mathcal{D}g_D - \mathcal{S}\psi$	$\partial_{\mathbf{n}}^+ u$	$H^{-\frac{1}{2}}(\Gamma)$	direct	II	never!
(74)	$(\frac{1}{2}+D'-\mathrm{i}\eta S)\psi = \partial_{\mathbf{n}}^+ u^{\mathrm{Inc}} - \mathrm{i}\eta\gamma^+ u^{\mathrm{Inc}}$	$u^{\mathrm{Scat}} = -\mathcal{S}\psi$	$\partial_{\mathbf{n}}^+ u^{\mathrm{Tot}}$	$H^{-\frac{1}{2}}(\Gamma)$	direct	II	never!

Table 1: Six (+1) BIEs for the EDP (32).
They are described in [CK1]: (42) (3.44), (68) (3.26), (69) (3.83), (70) (3.81), (72) (3.51), (73) (3.84).
The BIE (74) can be used only for the SSSP (33) as it requires both traces of u^{Inc}.

Remark 6.17: (BIO diagonalisation on the circle and BIE stability). Let $\Gamma = \mathbb{S}^1$, the unit circle. Then, all the integral operators we have encountered diagonalise in the circular harmonic basis. This means that $\mathrm{e}^{\mathrm{i}\ell\theta}$ is eigenfunction of S, K, K' and H for all $\ell \in \mathbb{Z}$. In particular the operators take the values

$$
\begin{aligned}
S\mathrm{e}^{\mathrm{i}\ell\theta} &= \frac{\mathrm{i}\pi}{2} J_\ell(k) H_\ell^{(1)}(k)\mathrm{e}^{\mathrm{i}\ell\theta},\\
D\mathrm{e}^{\mathrm{i}\ell\theta} = D'\mathrm{e}^{\mathrm{i}\ell\theta} &= \Big(\frac{\mathrm{i}\pi k}{2} J_\ell(k) H_\ell^{(1)\prime}(k) + \frac{1}{2}\Big)\mathrm{e}^{\mathrm{i}\ell\theta} = \Big(\frac{\mathrm{i}\pi k}{2} J_\ell(k)' H_\ell^{(1)}(k) - \frac{1}{2}\Big)\mathrm{e}^{\mathrm{i}\ell\theta}, \qquad \ell \in \mathbb{Z}, k > 0\\
H\mathrm{e}^{\mathrm{i}\ell\theta} &= \frac{\mathrm{i}\pi k^2}{2} J_\ell'(k) H_\ell^{(1)\prime}(k)\mathrm{e}^{\mathrm{i}\ell\theta}.
\end{aligned}
$$

[29] Figure 18 shows the density ψ for the indirect BIE (42). However we have seen in §5.5.1 that the solution of this BIE has the physical interpretation $\psi = -\partial_{\mathbf{n}}^+ u^{\mathrm{Tot}}$, while this is not true for the solutions of the other indirect BIEs (68) and (72).

This implies that the (spectral) Galerkin-BEM matrix with basis $\{\mathrm{e}^{\mathrm{i}\ell\theta}\}_{\ell=-L,\dots,L}$, for any BIE on the unit circle, gives a diagonal matrix.

This offers a very simple way to study the spurious resonances of the BIEs. We denote $\lambda_{T,\ell}$ the eigenvalue of an operator T associated to $\mathrm{e}^{\mathrm{i}\ell\theta}$, i.e. $T\mathrm{e}^{\mathrm{i}\ell\theta} = \lambda_{T,\ell}\mathrm{e}^{\mathrm{i}\ell\theta}$. In Figure 20 we plot the values $\|(\underline{\underline{\mathbf{A}}}^{\mathrm{Gal}})^{-1}\| = \inf_{\ell\in\{-L,\dots,L\}} |\lambda_{T,\ell}|^{-1}$ for $L = 30$ as a function of the wavenumber k. Large values of this norm signal that the linear system is close to singular, and that the corresponding BIE is not invertible for nearby values of k. We observe these instabilities near the (square roots of the) Dirichlet eigenvalues for the operator $T = S$ (corresponding to the BIEs (42) and (69)) and near the Neumann eigenvalues for the operator $T = \frac{1}{2} + D$ (corresponding to the BIEs (68) and (70), since 'taking the adjoint' does not change the operator norm). On the other hand, the operator $T = A = \frac{1}{2} + D - \mathrm{i}kS$ of the BIEs (72) and (73) is bounded uniformly. (Here we have chosen $\eta = k$ in accordance to Remark 6.14, while choosing $\eta = \max\{k, 1\}$ would stabilise the operator also for $k \to 0$, which is a Neumann eigenvalue.)

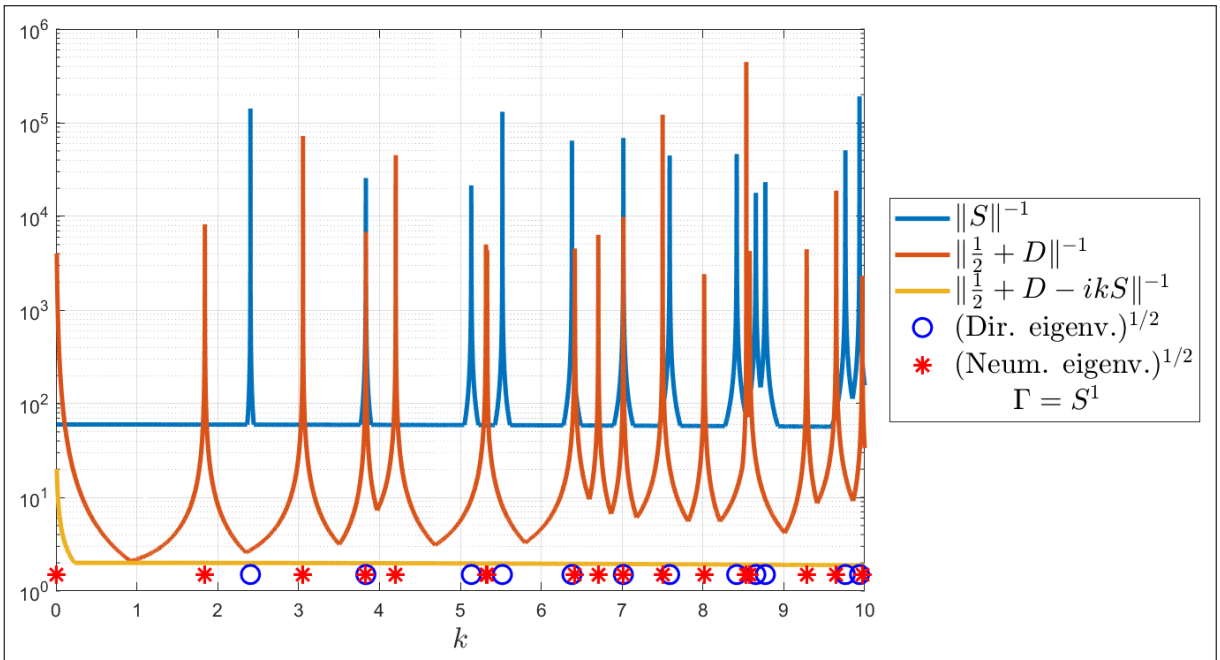


Figure 20: The norm of the inverses of the operators of the BIEs of Table 1 for $\Gamma = \mathbb{S}^1$. The operator norm is approximated by the norm of the inverse of the Galerkin-BEM matrix with basis $(\mathrm{e}^{\mathrm{i}\ell\theta})_{\ell=-30,\dots,30}$, which is a diagonal matrix as described in Remark 6.17. The strong peaks in correspondence of the square roots of the Dirichlet (◯) and Neumann (∗) eigenvalues of the disc are the spurious resonances. Only the combined-field operator $A = \frac{1}{2} + D - ikS$ (yellow line) is stable for all wavenumbers.

Remark 6.18: (The CHIEF method). A different way to deal with spurious resonances and ensure the solvability of the BEM linear system even when k^2 is a Dirichlet eigenvalue of Ω_- is the Combined Helmholtz Integral Equation Formulation (CHIEF) proposed in [Schenck 1968]. We know from (62) that $(S\psi)|_{\Omega_-} = -u^{\mathrm{Inc}}|_{\Omega_-}$, where ψ is the solution of the single-layer BIE (42). The CHIEF augments the BEM linear system $\underline{\underline{\mathbf{A}}}\boldsymbol{\psi} = \mathbf{F}$ with a few more equations $(\mathcal{S}\psi_N)(\mathbf{z}_j) = -u^{\mathrm{Inc}}(\mathbf{z}_j)$ for some points $\mathbf{z}_j \in \Omega_-$, $j = 1, \dots, M$. This gives an overdetermined linear system that can be solved in the least-squares sense. At least some of the points $\mathbf{z}_j$ need to be away from the nodal lines of the eigenfunction corresponding to k^2, which of course is not known, so their choice is delicate (in particular for large frequencies).

Remark 6.19: (Exterior Neumann problem). So far we have only considered exterior Dirichlet problems. All techniques and results can be extended to the **exterior Neumann problem**: find $u \in H^1_{\mathrm{loc}}(\overline{\Omega_+})$ such that

$$\Delta u + k^2 u = 0 \quad \text{in } \Omega_+, \qquad \partial^+_{\mathbf{n}} u = g_N \quad \text{on } \Gamma, \qquad u \text{ radiating}, \tag{75}$$

where $g_N \in H^{-\frac{1}{2}}(\Gamma)$ is a boundary datum. The corresponding **sound-hard scattering problem** is

$$\Delta u^{\mathrm{Scat}} + k^2 u^{\mathrm{Scat}} = 0 \quad \text{in } \Omega_+, \qquad \partial^+_{\mathbf{n}}(u^{\mathrm{Scat}} + u^{\mathrm{Inc}}) = 0 \quad \text{on } \Gamma, \qquad u^{\mathrm{Scat}} \text{ radiating},$$

which is a special case of (75) for $g_N = -\partial^+_{\mathbf{n}} u^{\mathrm{Inc}}$. The tools needed to construct BIEs for this problem are the same already introduced: the four BIOs (S, D, D', H), the two layer potentials $(\mathcal{S}, \mathcal{D})$ and their relations.

We can construct six BIEs, proceeding precisely as we have done for the Dirichlet case in the previous sections: see the summary in Table 2. Coding a BEM for the exterior Neumann problem is slightly harder than for the EDP (32) because there is no BIE involving only the single-layer operator S and many of the BIEs listed require the hypersingular operator H. See also [Sayas06, §11].

Exercise 6.20: (BIEs for the exterior Neumann problem). Derive the six BIEs in Table 2 for the exterior Neumann problem (75). Start from the representation formulas and use the trace relations (60).

[CK1] eq.#	BIE	representation formula	density $\psi =$	direct/ indirect	kind	fails for
(3.29)	$(-\frac{1}{2}+D')\psi = g_N$	$u = \mathcal{S}\psi$		indirect	II	Dir. eig.
(3.47)	$H\psi = g_N$	$u = \mathcal{D}\psi$	$\gamma^+ u^{\mathrm{Tot}}$	indirect	I	Neum. eig.
(3.82)	$(-\frac{1}{2}+D)\psi = Sg_N$	$u = \mathcal{D}\psi - \mathcal{S}g_N$	$\gamma^+ u$	direct	II	Dir. eig.
(3.85)	$H\psi = (\frac{1}{2}+D')g_N$	$u = \mathcal{D}\psi - \mathcal{S}g_N$	$\gamma^+ u$	direct	I	Neum. eig.
(3.53)	$(-\frac{1}{2}+D'+\mathrm{i}\eta H)\psi = g_N$	$u = (\mathcal{S}+\mathrm{i}\eta\mathcal{D})\psi$		indirect	II	never
(3.86)	$(-\frac{1}{2}+D+\mathrm{i}\eta H)\psi$ $=(S+\mathrm{i}\eta(\frac{1}{2}+D'))g_N$	$u = \mathcal{D}\psi - \mathcal{S}g_N$	$\gamma^+ u$	direct	II	never

Table 2: Six BIEs for the exterior Neumann problem (75). The first column indicates the equation number in [CK1]. Note the symmetry with the Dirichlet case in Table 1.

Exercise 6.21: (BIO diagonalisation on the circle). (i) Read carefully Remark 6.17 and reproduce Figure 20.
(ii) Repeat the same for the exterior Neumann problem described in Remark 6.19.

Exercise 6.22: (Exterior impedance problem). Consider the exterior impedance problem

$$\Delta u + k^2 u = 0 \quad \text{in } \Omega_+, \qquad \partial_{\mathbf{n}}^+ u + \mathrm{i}k\vartheta\,\gamma^+ u = g_I \quad \text{on } \Gamma, \qquad u \text{ radiating}, \tag{76}$$

for $g_I \in H^{-\frac{1}{2}}(\Gamma)$ and $\vartheta \in L^\infty(\Gamma)$, $\operatorname{ess\,inf}\vartheta > 0$. (For the interior impedance BVP (27) on Ω_- we used the condition with opposite sign $\partial_{\mathbf{n}}^- u - \mathrm{i}k\vartheta\gamma^- u = g_I$: the two choices are consistent because $\mathbf{n}$, entering the definition of $\partial_{\mathbf{n}}^\pm$, points in Ω_+ and out of Ω_-.)

- Using Lemma 5.17, prove that (76) admits at most one solution.
- Show that $u = \mathcal{S}\psi$, for $\psi \in H^{-\frac{1}{2}}(\Gamma)$ solution of the BIE $(\frac{1}{2} - D' - \mathrm{i}k\vartheta S)\psi = -g_I$, is solution of (76). Show that the operator of this BIE is injective if and only if k^2 is not a Dirichlet eigenvalue in Ω_-.
- Show that $u = (\mathcal{S}+\mathrm{i}\eta\mathcal{D})\psi$ solves (76) if ψ is solution of the combined-field BIE

$$\Big(\frac{1+k\vartheta\eta}{2} - \mathrm{i}\vartheta k S + \eta\vartheta k D - D' - \mathrm{i}\eta H\Big)\psi = -g_I.$$

  Finally, show that the operator of this BIE is injective for all $k > 0$ and $\eta > 0$.

For help and more details, see [CK1, §3.7].

6.2.6 There is much more than this!

In these notes we have considered exterior Dirichlet BVPs for the 2D Helmholtz equation. However BIEs and BEMs have a much broader range of applicability. They may be used to model, analyse and approximate Helmholtz BVPs posed on bounded domains, on domains with unbounded or non-Lipschitz boundaries, transmission problems (i.e. multiple Helmholtz equations with different wavenumbers on different domains, coupled by Dirichlet and Neumann conditions),....

BIEs can be used for much more general linear PDEs (of all kinds: elliptic, parabolic and hyperbolic) in any dimension, including systems of PDEs such as those of elasticity and electromagnetism (recall §1.3–1.4). In computational electromagnetism, the BEM is often called "method of moments" (MoM). The main requirement for implementing a BEM is that the fundamental solution of the problem is known, either in exact or approximate form.

A BEM can be coupled with a FEM (or another volume-based method) for approximating problems with different physical models in different subdomains. Typically, FEMs are used in small regions of high geometric complexity, variable coefficients or non-linearities, and BEMs are used to deal with unbounded regions where coefficients are constant.

The use of BEM in applications requires the solution of large dense linear systems. Often their solution with direct methods (such as Gauss elimination) is too expensive and they are often solved with iterative (Krylov) methods such as GMRES, usually with preconditioning. For large systems, even assembling the matrix may be unfeasible. Thus several techniques to compute matrix–vector multiplications, which are the key steps in Krylov methods, without explicitly assembling the matrix have been developed. These techniques exploit the structure of the BEM matrix and the properties of the fundamental solution. Important realisations of this idea are the fast multipole method (FMM), the panel clustering and the hierarchical matrices ($\mathcal{H}$-matrices).

An open-source, high-performance Galerkin-BEM code with a Python interface is *Bempp*. On its website (`http://bempp.com/`) you can find examples of the use of BEM for high-intensity focused ultrasound medicine (where the underlying PDE is the Helmholtz equation) and electromagnetic wave scattering.

Another important numerical method for the discretisation of BIEs is Nyström method, which can converge extremely fast for smooth scatterers; see [CK2, §3.5] for the application to 2D Brakhage–Werner equation or [Sayas15, pp. 33 and 36] for the Laplace case.

Plenty of information can be found in the references in the bibliography and in the books mentioned in Remark 5.24.

6.3 Error analysis of the Galerkin method applied to Gårding-type problems

6.3.1 Abstract variational framework

We recall the general variational problem (22) and its Galerkin approximation (23):

$$\text{find} \quad u \in H \qquad \text{such that} \quad \mathcal{A}(u,w) = \mathcal{F}(w) \qquad \forall w \in H, \tag{77}$$

$$\text{find} \quad u_N \in V_N \qquad \text{such that} \quad \mathcal{A}(u_N,w_N) = \mathcal{F}(w_N) \qquad \forall w_N \in V_N, \tag{78}$$

where H is a Hilbert space, $V_N \subset H$ is a finite-dimensional subspace, $\mathcal{A}(\cdot,\cdot)$ and $\mathcal{F}(\cdot)$ are a sesquilinear and a continuous antilinear functional on H.

If $\mathcal{A}(\cdot,\cdot)$ is continuous and coercive in H, the stability and error analysis of the Galerkin method is simple, thanks to Lax–Milgram theorem and Céa lemma. Unfortunately, the variational problems studied in this course fall outside of this framework. The key result to extend Céa lemma to problems satisfying a Gårding inequality is the following theorem. This is a modification of the classical "Aubin–Nitsche duality trick" used also in finite element analysis; in the context of Helmholtz problems it is often called "Schatz argument". Here we closely follow [Spence14, Thm. 5.21].

Theorem 6.23: (Galerkin method for Gårding inequality). Let $H \subset V$ be Hilbert spaces and the inclusion be compact. Let $\mathcal{A}(\cdot,\cdot)$ be a continuous sesquilinear form on H that satisfies the Gårding inequality (24):

$$|\mathcal{A}(v,w)| \leq C_{\mathcal{A}} \|v\|_H \|w\|_H, \qquad \Re\{\mathcal{A}(v,v)\} \geq \alpha \|v\|_H^2 - C_V \|v\|_V^2, \qquad \forall v,w \in H \qquad (\alpha, C_V, C_{\mathcal{A}} > 0).$$

Assume that the only $u_0 \in H$ such that $\mathcal{A}(u_0,v) = 0$ for all $v \in H$ is $u_0 = 0$ (so that the variational problem (77) is well-posed for any right-hand side). Let $\mathcal{F}(\cdot)$ be a continuous linear functional on H and u be the solution of the variational problem (77).
Given $f \in V$, let $z_f \in H$ be the solution of the **adjoint problem**

$$\mathcal{A}(v,z_f) = (v,f)_V \qquad \forall v \in H, \tag{79}$$

where $(\cdot,\cdot)_V$ is the scalar product in V. Let V_N be a finite-dimensional subspace of H and define

$$\eta(V_N) := \sup_{f\in V, f\neq 0} \min_{v_N \in V_N} \frac{\|z_f - v_N\|_H}{\|f\|_V}. \tag{80}$$

If $\eta(V_N)$ satisfies the **threshold condition**

$$\eta(V_N) \leq \frac{1}{C_{\mathcal{A}}} \sqrt{\frac{\alpha}{2C_V}}, \tag{81}$$

then the Galerkin method (78) is **well-posed** and its solution u_N satisfies the **quasi-optimality** bound

$$\|u - u_N\|_H \leq \frac{2C_{\mathcal{A}}}{\alpha} \min_{v_N \in V_N} \|u - v_N\|_H. \tag{82}$$

The statement of this theorem is not simple and requires some explanation. Our goal is to prove that the Galerkin method is well posed and the quasi-optimality bound (82) holds. This ensures that the Galerkin error $\|u-u_N\|_H$ is controlled by the best-approximation error, i.e. by the best error achievable if we knew the exact solution u and we tried to approximate it with the discrete space V_N. Quasi-optimality holds for all finite-dimensional subspaces $V_N \subset H$ when we are in a Lax–Milgram setting, i.e. when $\mathcal{A}(\cdot,\cdot)$ is coercive. The bad news is that this is not true for all V_N if $\mathcal{A}(\cdot,\cdot)$ only satisfies a Gårding inequality. The good news is that well-posedness and quasi-optimality hold if V_N "has sufficiently good approximation properties". How do we measure the approximation properties of a discrete space? The "adjoint approximability parameter" $\eta(V_N)$ in (80) precisely quantifies how well V_N approximates the solution of the "adjoint problem" (79), whose datum f is an arbitrary element of the larger space V. The smaller $\eta(V_N)$, the better the approximation offered by V_N. If the "threshold condition" (81) holds, i.e. $\eta(V_N)$ is smaller than a certain quantity that depends on $\mathcal{A}(\cdot,\cdot)$, then we have what we want: the well-posedness of the Galerkin method and its quasi-optimality.[30]

In brief, Theorem 6.23 states, in a precise quantitative way, the following:

> if the discrete space is sufficiently fine, then the Galerkin method (applied to a well-posed Gårding-type problem) is well-posed and quasi-optimal.

If we have a dense sequence of discrete spaces $(V_N)_{N\in\mathbb{N}}$ (i.e. such that $\inf_{v_N\in V_N}\|v-v_N\|_H \to 0$ for $N\to\infty$ and for all $v\in H$) then the theorem states that we eventually achieve convergence. Typically, for constant-degree piecewise-polynomial FEM or BEM, this means that the method is guaranteed to be well-posed and quasi-optimal if the mesh employed is sufficiently fine. In the next sections we analyse a couple of examples coming from discretisations of the Helmholtz equation.

The phenomenon described by Theorem 6.23 is often observed numerically. When we approximate a Laplace-type problem with a sequence of meshes we observe a reduction of the error starting from the first mesh refinements. For a Helmholtz-type problem, mesh refinement doesn't give any improvement in the solution until some threshold h_0 on the mesh size h is reached; for $h<h_0$ we observe convergence of the error to zero.

Proof of Theorem 6.23. We follow [Spence14, Thm. 5.21].[31]

We first assume that the Galerkin method (78) admits a solution u_N. We recall the Galerkin orthogonality $\mathcal{A}(u-u_N,v_N)=0$ which holds for all $v_N\in V_N$. In the following bound we use the adjoint problem (79) with source term $f=u-u_N$ and denote by $w_N\in V_N$ the element minimising the ratio in the definition (80) of the approximability parameter $\eta(V_N)$. We first control the (weaker) V norm of the Galerkin error, exploiting the adjoint problem:

$$\begin{aligned}
\|u-u_N\|_V^2 &= \mathcal{A}(u-u_N, z_{u-u_N}) && \text{adjoint problem (79)}\\
&= \mathcal{A}(u-u_N, z_{u-u_N}-w_N) && \text{Galerkin orthogonality}\\
&\le C_{\mathcal{A}}\|u-u_N\|_H\,\|z_{u-u_N}-w_N\|_H && \text{continuity of } \mathcal{A}(\cdot,\cdot)\\
&\le C_{\mathcal{A}}\eta(V_N)\|u-u_N\|_H\,\|u-u_N\|_V && \text{definition of } \eta(V_N)\ (80)\\
&\le \sqrt{\frac{\alpha}{2C_V}}\|u-u_N\|_H\,\|u-u_N\|_V && \text{threshold condition (81).}
\end{aligned}$$

Then also the (stronger) H norm of the error can be controlled:

$$\begin{aligned}
\alpha\|u-u_N\|_H^2 &\le \Re\{\mathcal{A}(u-u_N,u-u_N)\}+C_V\|u-u_N\|_V^2 && \text{Gårding inequality (24)}\\
&= \Re\{\mathcal{A}(u-u_N,u-v_N)\}+C_V\|u-u_N\|_V^2 && \text{Galerkin orthogonality, } \forall v_N\in V_N\\
&\le C_{\mathcal{A}}\|u-u_N\|_H\,\|u-v_N\|_H+C_V\|u-u_N\|_V^2 && \text{continuity of } \mathcal{A}(\cdot,\cdot)\\
&\le C_{\mathcal{A}}\|u-u_N\|_H\,\|u-v_N\|_H+\frac{\alpha}{2}\|u-u_N\|_H^2
\end{aligned}$$

where in the last step we used the previous bound (after simplifying a term and squaring). Moving the last term to the left-hand side of the equation, and simplifying $\|u-u_N\|_H$, we obtain the desired quasi-optimality bound (82).

[30] A (surprising) theorem by Gohberg and Fel'dman (1971) states that, given a continuous linear operator $A: H\to H$, this is sum of a coercive and a compact operator if and only if the Galerkin method for $(Au,v)_H=(f,v)_H$ converges for all sequences $(H_N)_{N\in\mathbb{N}}$ of nested finite-dimensional subspaces such that $\lim_{N\to\infty}\inf_{u_N\in H_N}\|u-u_N\|_H=0$ for all $u\in H$. In words: the Galerkin method converges for all dense subspace sequences iff the problem is a compact perturbation of a coercive problem. In this section we are studying the implication "compactly-perturbed problem⇒Galerkin convergence", the converse means that we cannot expect this to hold in a much more general setting.

[31] This proof should remind you the derivation of error estimates on the L^2 norm of the error of classical finite elements for elliptic problems.

Taking $v_N = 0$ in (82) and using the triangle inequality, we have the stability estimate $\|u_N\|_H \leq (1 + \frac{2C_{\mathcal{A}}}{\alpha}) \|u\|_H$. By the injectivity assumption stipulated in the theorem, if $\mathcal{F} = 0$ then $u = 0$ and, by this stability bound, also $u_N = 0$. This means that the solution of the Galerkin method (78) is at most unique. Since the method is equivalent to a square $N \times N$ linear system, uniqueness implies well-posedness. We conclude that u_N exists, so the assumption made at the beginning of the proof is justified. □

Remark 6.24: (Galerkin method for compactly perturbed problem). A related error analysis of the Galerkin approximation of variational problems whose sesquilinear form is sum of a coercive and a compact one is presented in [SBH19, §8.9]. This theory is simpler, does not use the adjoint problem, and it is slightly more general: it does not require the Gårding inequality but only the Fredholm property. On the other hand, the threshold condition (81) and the quasi-optimality bound (82) obtained this way are not explicit.

Exercise 6.25: (Parameter tuning). Show that if we assume $\eta(V_N) \leq \frac{1}{C_{\mathcal{A}}} \sqrt{\frac{\alpha}{bC_V}}$ for some $b > 1$ instead of (81), then the quasi-optimality constant $\frac{2C_{\mathcal{A}}}{\alpha}$ in (82) can be substituted by $\frac{b}{b-1} \frac{C_{\mathcal{A}}}{\alpha}$.

This means that, with this proof, we can win at most a factor $\sqrt{2}$ in the threshold on $\eta(V_N)$ (i.e. allow a slightly coarser discrete space), paying with a poorer quasi-optimality constant. Conversely (choosing $b > 2$) one can reduce by a factor at most 2 the quasi-optimality constant, paying with a more restrictive threshold condition. In brief: in Theorem 6.23 we arbitrarily chose the factor 2 that appears in (81) and (82), but other choices don't give substantial improvements.

In the next two sections we sketch how to apply Theorem 6.23 to two exemplary problems: the FEM for the interior Helmholtz problem, and the BEM for the exterior Dirichlet problem. We use several results from the theory of PDEs, BIOs and FEMs, without justifying their use in full details.

6.3.2 FEM error analysis for interior problems

To better understand Theorem 6.23 we apply it to the simplest problem we know in this setting: the variational formulation (26) of the interior Helmholtz Dirichlet problem (25). Here $\Omega \subset \mathbb{R}^2$ is a bounded Lipschitz domain,

$$H = H^1_0(\Omega), \qquad V = L^2(\Omega), \qquad \mathcal{A}(u, w) = \int_\Omega (\nabla u \nabla \overline{w} - k^2 u \overline{w}) \, \mathrm{d}\mathbf{x}.$$

It is convenient to use a dimensionally-homogeneous version of the $H^1(\Omega)$ norm, weighted with the wavenumber k:

$$\|v\|^2_{H^1_k(\Omega)} := |v|^2_{H^1(\Omega)} + k^2 \|v\|^2_{L^2(\Omega)}.$$

With this norm, the constants in the Gårding inequality and in the continuity of $\mathcal{A}(\cdot, \cdot)$ are

$$\alpha = 1, \qquad C_V = 2k^2, \qquad C_{\mathcal{A}} = 1.$$

For each $f \in L^2(\Omega)$, the adjoint problem (79) consists in finding $z_f \in H^1_0(\Omega)$ such that

$$\int_\Omega (\nabla v \nabla \overline{z_f} - k^2 v \overline{z_f}) \, \mathrm{d}\mathbf{x} = \int_\Omega v \overline{f} \, \mathrm{d}\mathbf{x} \qquad \forall v \in H^1_0(\Omega),$$

where at the right-hand side we simply have the $V = L^2(\Omega)$ scalar product. Taking the complex conjugate, this is again (26), the variational formulation of the interior Helmholtz Dirichlet problem:

$$\Delta z_f + k^2 z_f = -f \quad \text{in } \Omega, \qquad \gamma u = 0 \quad \text{on } \partial\Omega.$$

This is because the Helmholtz Dirichlet problem is self-adjoint: it coincides with its adjoint. We assume that k^2 is not a Dirichlet eigenvalue in Ω: this corresponds to the assumption made in the theorem that the homogeneous problems admits only the trivial solution. Well-posedness comes with stability, also for the adjoint problem: $\|z_f\|_{H^1_k(\Omega)} \leq C_{\mathrm{stab}} \|f\|_{L^2(\Omega)}$, for some $C_{\mathrm{stab}} > 0$ depending on k and Ω but independent of f (and which may be difficult to estimate).

Since $z_f \in H^1_0(\Omega)$ and $\Delta z_f = -f - k^2 z_f \in L^2(\Omega)$, by elliptic regularity [Ihl98, Prop. 2.24], if Ω is a smooth or convex domain then $z_f \in H^2(\Omega)$ and $\|z_f\|_{H^2(\Omega)} \leq C_{H^2} \|f + k^2 z_f\|_{L^2(\Omega)} \leq C_{H^2}(1 + kC_{\mathrm{stab}}) \|f\|_{L^2(\Omega)}$ for some $C_{H^2} > 0$ depending on Ω. Assume V_N is the space of piecewise-linear finite elements on a quasi-uniform, shape-regular triangulation of Ω with mesh size h. Classical finite-element approximation estimates (Bramble–Hilbert lemma, e.g. as in [Ihl98, eq. (4.1.10)]) state that

$\inf_{v_N\in V_N}\|v-v_N\|_{H^1_k(\Omega)} \le C_{BH}h\,\|v\|_{H^2(\Omega)}$ for all $v\in H^2(\Omega)$ and for some C_{BH} that depends on Ω, k and the "chunkiness" of the mesh elements[32]. Collecting all these bounds:

$$\inf_{v_N\in V_N}\|z_f-v_N\|_{H^1_k(\Omega)} \le C_{BH}h\,\|z_f\|_{H^2(\Omega)} \le hC_{BH}C_{H^2}(1+kC_{\mathrm{stab}})\,\|f\|_{L^2(\Omega)} \qquad \forall f\in L^2(\Omega).$$

We have obtained a bound on the adjoint approximability parameter defined in (80):

$$\eta(V_N) \le hC_{BH}C_{H^2}(1+kC_{\mathrm{stab}}).$$

Then the threshold condition (81) is satisfied if $h\le h_\star$, where

$$h_\star := \frac{1}{C_{\mathcal{A}}C_{BH}C_{H^2}(1+kC_{\mathrm{stab}})}\sqrt{\frac{\alpha}{2C_V}} \le \frac{1}{2k\,C_{BH}C_{H^2}(1+kC_{\mathrm{stab}})}.$$

If the finite element mesh size h is smaller than this value, then the method is well-posed and the quasi-optimality bound (82)

$$\|u-u_N\|_{H^1_k(\Omega)} \le 2\min_{v_N\in V_N}\|u-v_N\|_{H^1_k(\Omega)}$$

holds. This gives a recipe to choose a suitable mesh and predict what is the computational effort required for the FEM to approximate the solution u of the Dirichlet problem (25) to a desired accuracy.

Exercise 6.26: (Galerkin method for the impedance problem). Repeat the argument of this section for the FEM applied to the interior impedance problem (28). Which of the assumptions in Theorem 6.23 comes for free?

Remark 6.27: (k-dependence in the FEM). We observe that the mesh size threshold $h_\star$ decreases with $k\to\infty$: this means that higher frequencies require finer meshes (and larger linear systems: in 2D the number of DOFs involved in a standard FEM discretisation is proportional to h^{-2}). In Remark 5.7 we explained why $h\lesssim k^{-1}$ is needed to maintain a given level of approximation for large k. Here instead we are not looking at the approximability of the solution, but at the stability of the Galerkin scheme: the bound on $h_\star$ is needed to ensure well-posedness and quasi-optimality.

Moreover, we observe that $h_\star$ grows more than linearly in k^{-1}.[33] So, while $h=\mathcal{O}(k^{-1})$ ($N\approx k^2$ DOFs) is enough to ensure good approximation properties, we need finer meshes with $h=\mathcal{O}(k^{-a})$ ($N\approx k^{2a}$ DOFs) for some $a>1$ to ensure stability.

This is not an artefact of the proof but can be observed numerically and is a major problem in the numerical analysis of high-frequency time-harmonic problems. This notorious phenomenon is called **pollution effect**; it is described e.g. in [Ihl98, §4.6]. High-order methods (i.e. FEM with high-degree piecewise-polynomial discrete spaces) perform better under this respect. A great deal of research is devoted to the design of Galerkin schemes for time-harmonic problems that are robust and accurate for large wavenumbers.

Exercise 6.28: (FEM numerical experiments). Use piecewise-linear finite elements on a quasi-uniform mesh to approximate an interior Helmholtz BVP. Choose either a (well-posed) Dirichlet or an impedance problem, in 1D or in 2D.

Plot the norm of (i) the Galerkin error and (ii) the best-approximation error (e.g. computed with the $L^2(\Omega)$ projection) for decreasing mesh sizes h. Repeat the same plot for different values of k. Observe how the best-approximation error starts to converge only for $h\approx k^{-1}$, while the Galerkin method needs even smaller h to start converging. Study numerically the quasi-optimality constant in dependence of k.

You can try to reproduce the example in Figure 21.

6.3.3 BEM error analysis for the single-layer BIE

We want to apply Theorem 6.23 to the Galerkin-BEM discretisation (§5.2) of the single-layer BIE $S\psi=g_D$, (42). Our goal is to show that the piecewise-constant Galerkin-BEM is well-posed and quasi-optimal if the mesh size h is sufficiently small. We recall the variational formulation (44):

$$\text{find } \psi\in H^{-\frac12}(\Gamma) \quad \text{such that} \quad \mathcal{A}(\psi,\xi) := \langle S\psi,\xi\rangle_\Gamma = \langle g_D,\xi\rangle_\Gamma =: \mathcal{F}(\xi) \qquad \forall \xi\in H^{-\frac12}(\Gamma).$$

[32] More precisely, $\inf_{v_N\in V_N}\|v-v_N\|_{H^1_k(\Omega)} \le C'_{BH}(h+h^2k)\,|v|_{H^2(\Omega)}$ for C'_{BH} independent of k.

[33] The bound suggests a quadratic growth $h_\star\approx k^{-2}$, but $h_\star\approx k^{-3/2}$ is typically enough [Ihl98, eq. (4.7.41)]. Here we are not taking into accounts that C_{stab} depends on k; this dependence is more easily studied for impedance problems and simple geometries (Ω convex or star-shaped).

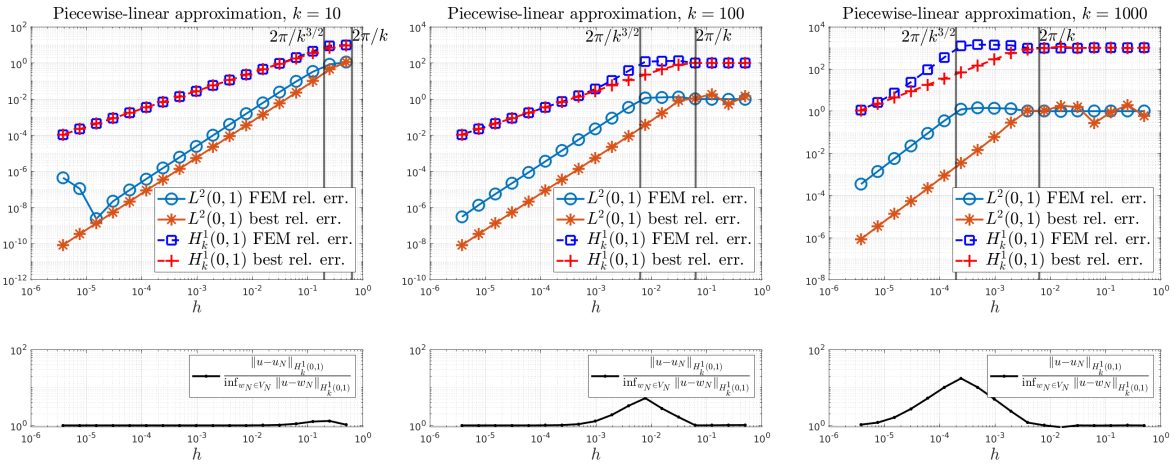


Figure 21: Error analysis of the finite element approximation of an interior impedance Helmholtz BVP. We consider the simple 1D problem $u''+k^2u=0$, $u'(0)+\mathrm{i}ku(0)=2\mathrm{i}k$, $u'(1)-\mathrm{i}ku(1)=0$ on the interval $\Omega=(0,1)$ for $k=10$, 100, 1000. We approximate $u(x)=\mathrm{e}^{\mathrm{i}kx}$ with piecewise-linear finite elements on uniform meshes with element size h and $N=1+1/h$ DOFs.
In blue we see the Galerkin relative error $\frac{\|u-u_N\|_\bullet}{\|u\|_\bullet}$ in $L^2(0,1)$ (◯ markers) and $H^1_k(0,1)$ (□ markers) norms plotted against $h\in\{2^{-1},2^{-2},\dots,2^{-18}\}$. We compare against the $L^2(0,1)$ (∗ markers) and $H^1_k(0,1)$ (+ markers) best approximation errors (computed by the orthogonal projection with the relevant matrices). We observe three regimes:
(i) for $\frac{2\pi}{k}\lesssim h$ (right) u is not approximated at all, the wave is not resolved by the discrete space;
(ii) for $\frac{2\pi}{k^{3/2}}\lesssim h\lesssim\frac{2\pi}{k}$ (centre) the best approximation error decays with optimal order, while the Galerkin error remains roughly constant;
(iii) for $h\lesssim\frac{2\pi}{k^{3/2}}$ (left part of each plot) the Galerkin error decays with optimal rate.
The lower plots show the quasi-optimality ratio.
The presence of the regime (ii) (in which the Galerkin solution performs much worse than the best approximation) and the growth of the quasi-optimality ratio for increasing values of k are signals of the pollution effect. See Remark 6.27 and Exercise 6.28.

The space in which the variational problem is posed is $H=H^{-\frac12}(\Gamma)$. The continuity constant of the sesquilinear form $\mathcal{A}(\cdot,\cdot)$ is $C_{\mathcal{A}}=\|S\|_{H^{-\frac12}(\Gamma)\to H^{\frac12}(\Gamma)}$ (with the $H^s(\Gamma)$ norms defined as in Exercise 3.6).

We have seen in §6.1 that the operator $S:H^{-\frac12}(\Gamma)\to H^{\frac12}(\Gamma)$ is Fredholm but we haven't proved a Gårding inequality yet. The Fredholm property of S was proved from the decomposition $S=S_c+(S-S_c)$ for $c\geq0$, where the first term is coercive and the second compact. Recall from §6.1.2 that S_c is the single-layer operator for the reaction–diffusion or the Laplace equation, for $c>0$ or $c=0$, respectively. In the following we choose $c>0$ for simplicity. To write a Gårding inequality for $\mathcal{A}(\cdot,\cdot)$ we need a space V larger than $H^{-\frac12}(\Gamma)$ (recall that $H^{-\frac12}(\Gamma)$ is larger than $L^2(\Gamma)$). To identify this space V we need to study the mapping properties of $S-S_c$ and write a Gårding inequality.

We have seen numerically in §6.1.3 (in particular in Figure 19) that, on the circular boundary $\Gamma=\partial B_R$, $S-S_c:H^s(\Gamma)\to H^{s+3}(\Gamma)$ for all $s\in\mathbb{R}$. We assume that Γ is sufficiently regular so that $S-S_c:H^{-1}(\Gamma)\to H^1(\Gamma)$ is a continuous operator. We denote its operator norm by $C_1:=\|S-S_c\|_{H^{-1}(\Gamma)\to H^1(\Gamma)}$, depending only on k, c, Γ. Then the bound

$$|\langle(S-S_c)\psi,\xi\rangle_\Gamma|\leq\|(S-S_c)\psi\|_{H^1(\Gamma)}\,\|\xi\|_{H^{-1}(\Gamma)}\leq C_1\,\|\psi\|_{H^{-1}(\Gamma)}\,\|\xi\|_{H^{-1}(\Gamma)}\qquad\forall\psi,\xi\in H^{-1}(\Gamma)$$

and the coercivity of S_c shown in §6.1.5 give the Gårding inequality: $\forall\psi\in H^{-\frac12}(\Gamma)$,

$$\Re\big\{\mathcal{A}(\psi,\psi)\big\}=\Re\big\{\langle S\psi,\psi\rangle_\Gamma\big\}=\Re\big\{\langle S_c\psi,\psi\rangle_\Gamma\big\}+\Re\big\{\langle(S-S_c)\psi,\psi\rangle_\Gamma\big\}\geq\alpha_c\,\|\psi\|^2_{H^{-\frac12}(\Gamma)}-C_1\,\|\psi\|^2_{H^{-1}(\Gamma)}$$

where $\alpha_c:=\frac{1}{2C_{\mathrm{tr}}\max\{1,c^2\}}$ and C_{tr} is the continuity constant of the normal trace operator in (67). This Gårding inequality suggests to choose $V=H^{-1}(\Gamma)$.

The adjoint problem (79) is:

$$\text{given } f\in H^{-1}(\Gamma),\quad\text{find } z_f\in H^{-\frac12}(\Gamma)\quad\text{such that } \mathcal{A}(\xi,z_f)=\langle S\xi,z_f\rangle_\Gamma=(\xi,f)_{H^{-1}(\Gamma)}\quad\forall\xi\in H^{-\frac12}(\Gamma).$$

From $\mathcal{A}(v,w) = \mathcal{A}(\overline{w},\overline{v})$ (which can be verified by the definitions of $\mathcal{A}(\cdot,\cdot)$ and S) and $(v,w)_{H^{-1}(\Gamma)} = (\overline{w},\overline{v})_{H^{-1}(\Gamma)}$, we can rewrite the equation as $\mathcal{A}(\overline{z_f},\overline{\xi}) = (\overline{f},\overline{\xi})_{H^{-1}(\Gamma)}$. Dropping the complex conjugates, we obtain $\mathcal{A}(z_f,\xi) = (f,\xi)_{H^{-1}(\Gamma)}$. Since $H^{-1}(\Gamma) \ni \xi \mapsto (f,\xi)_{H^{-1}(\Gamma)}$ defines a continuous antilinear functional on $H^{-1}(\Gamma)$, it can be represented by an element $F \in H^1(\Gamma)$ with $\|F\|_{H^1(\Gamma)} \leq C_* \|f\|_{H^{-1}(\Gamma)}$ (i.e. $(f,\xi)_{H^{-1}(\Gamma)} = \langle F,\xi\rangle_\Gamma$, recall §3.3.3). Thus the adjoint problem is again the single-layer BIE $Sz_f = F$ for a "smooth" datum $F \in H^1(\Gamma)$.

Assume that k^2 is not a Dirichlet eigenvalue for Ω_- (which we need to assume for the BIE to be well-posed, as required by the theorem). Then $S : H^{-\frac{1}{2}}(\Gamma) \to H^{\frac{1}{2}}(\Gamma)$ is invertible, as explained in §6.1.1. It is possible to show that the single-layer operator is invertible[34] also as $S : L^2(\Gamma) \to H^1(\Gamma)$. This implies a regularity result for the adjoint problem: $z_f \in L^2(\Gamma)$.

The adjoint problem solution z_f is an element of $L^2(\Gamma)$ which can be approximated by the piecewise-constant space V_N defined in §5.2. Indeed, approximation (and operator interpolation) theory gives that there exists a discrete function $v_N \in V_N$ and a $C_{\mathrm{App}} > 0$ independent of h and z_f such that[35]

$$\begin{aligned}\|z_f - v_N\|_{H^{-\frac{1}{2}}(\Gamma)} \leq C_{\mathrm{App}} h^{\frac{1}{2}} \|z_f\|_{L^2(\Gamma)} &\leq C_{\mathrm{App}} h^{\frac{1}{2}} \left\|S^{-1}\right\|_{H^1(\Gamma)\to L^2(\Gamma)} \|F\|_{H^1(\Gamma)} \\ &\leq C_* C_{\mathrm{App}} h^{\frac{1}{2}} \left\|S^{-1}\right\|_{H^1(\Gamma)\to L^2(\Gamma)} \|f\|_{H^{-1}(\Gamma)}\,.\end{aligned}$$

Recall that h is the length of the longest mesh element. This corresponds to an estimate on $\eta(V_N)$:

$$\eta(V_N) = \sup_{f\in H^{-1}(\Gamma)} \min_{v_N\in V_n} \frac{\|z_f - v_N\|_{H^{-\frac{1}{2}}(\Gamma)}}{\|f\|_{H^{-1}(\Gamma)}} \leq C_* C_{\mathrm{App}} h^{\frac{1}{2}} \left\|S^{-1}\right\|_{H^1(\Gamma)\to L^2(\Gamma)}.$$

Similarly to the FEM case in §6.3.2, we have proved that if h is sufficiently small then the threshold condition (81) is satisfied and the Galerkin-BEM is well-posed and quasi-optimal.[36]

[34] We know $S : L^2(\Gamma) \to H^1(\Gamma)$ is injective for this choice of k. We have seen in §6.1.4 that $S - S_0 : L^2(\Gamma) \to H^1(\Gamma)$ is compact, as it corresponds to an integral operator with bounded kernel. It also holds that $S_0 : L^2(\Gamma) \to H^1(\Gamma)$ is invertible. Then Fredholm alternative Theorem 3.12 yields the invertibility of S in the desired spaces.

[35] In order to bound $\eta(V_N)$, both in the FEM and in the BEM cases, we used three main ingredients: (i) the stability of the adjoint problem, (ii) a regularity result for its solution, and (iii) an approximation result. In the FEM case, (i)–(ii)–(iii) correspond to the terms C_{stab}, C_{H^2} and C_{BH}, respectively. In the BEM case, stability and regularity together give $C_* \left\|S^{-1}\right\|_{H^1(\Gamma)\to L^2(\Gamma)}$ and the approximation term gives C_{App}.

[36] More precisely, if $h \leq h_\star$ for

$$h_\star := \left(4\max\{1,c^2\} C_*^2 C_{\mathrm{App}}^2 C_{\mathrm{tr}} \|S\|^2_{H^{-\frac{1}{2}}(\Gamma)\to H^{\frac{1}{2}}(\Gamma)} \|S - S_c\|_{H^{-1}(\Gamma)\to H^1(\Gamma)} \left\|S^{-1}\right\|^2_{H^1(\Gamma)\to L^2(\Gamma)}\right)^{-1},$$

we have the quasi-optimality estimate (82)

$$\|\psi - \psi_N\|_{H^{-\frac{1}{2}}(\Gamma)} \leq C_{qo} \inf_{v_N\in V_N} \|\psi - v_N\|_{H^{-\frac{1}{2}}(\Gamma)}, \qquad C_{qo} := 4C_{\mathrm{tr}} \max\{1,c^2\} \|S\|_{H^{-\frac{1}{2}}(\Gamma)\to H^{\frac{1}{2}}(\Gamma)}.$$

Both $h_\star$ and C_{qo} depends only on Γ, k and c (which can be chosen arbitrarily).

A Useful calculus formulas and notation

$B_R(\mathbf{x}) := \{\mathbf{y}\in\mathbb{R}^n : \ |\mathbf{y}-\mathbf{x}|<R\}, \qquad B_R := B_R(\mathbf{0}),$

polar coordinates: $(x_1,x_2) = (r\cos\theta, r\sin\theta), \quad r\geq 0,\ 0\leq\theta<2\pi, \qquad \mathrm{d}x_1\,\mathrm{d}x_2 = r\,\mathrm{d}r\,\mathrm{d}\theta,$

$\mathrm{e}^{\mathrm{i}z} = \cos z + \mathrm{i}\sin z, \quad \cos z = \dfrac{\mathrm{e}^{\mathrm{i}z}+\mathrm{e}^{-\mathrm{i}z}}{2}, \quad \sin z = \dfrac{\mathrm{e}^{\mathrm{i}z}-\mathrm{e}^{-\mathrm{i}z}}{2\mathrm{i}}, \qquad \Re = \text{real part}, \quad \Im = \text{imaginary part},$

$\mathbf{v}\times\mathbf{w} := (v_2w_3 - v_3w_2,\ v_3w_1 - v_1w_3,\ v_1w_2 - v_2w_1), \qquad \mathbf{u}\times(\mathbf{v}\times\mathbf{w}) = \mathbf{v}(\mathbf{u}\cdot\mathbf{w}) - \mathbf{w}(\mathbf{u}\cdot\mathbf{v}),$

$\nabla u := \Big(\dfrac{\partial u}{\partial x_1}, \ \dots, \ \dfrac{\partial u}{\partial x_n}\Big), \qquad \operatorname{div}\mathbf{v} := \nabla\cdot\mathbf{v} := \dfrac{\partial v_1}{\partial x_1} + \dots + \dfrac{\partial v_n}{\partial x_n},$

$\Delta u := \nabla^2 u := \operatorname{div}(\nabla u) \ = \ \dfrac{\partial^2 u}{\partial x_1^2} + \dots + \dfrac{\partial^2 u}{\partial x_n^2} \ \overset{\text{if } n=2}{=} \ \dfrac{1}{r}\dfrac{\partial}{\partial r}\Big(r\dfrac{\partial u}{\partial r}\Big) + \dfrac{1}{r^2}\dfrac{\partial^2 u}{\partial\theta^2} \ = \ \dfrac{\partial^2 u}{\partial r^2} + \dfrac{1}{r}\dfrac{\partial u}{\partial r} + \dfrac{1}{r^2}\dfrac{\partial^2 u}{\partial\theta^2},$

$\operatorname{curl}\mathbf{v} := \nabla\times\mathbf{v} := \Big(\dfrac{\partial v_3}{\partial x_2} - \dfrac{\partial v_2}{\partial x_3}, \ \dfrac{\partial v_1}{\partial x_3} - \dfrac{\partial v_3}{\partial x_1}, \ \dfrac{\partial v_2}{\partial x_1} - \dfrac{\partial v_1}{\partial x_2}\Big),$

$\operatorname{curl}\nabla u = \mathbf{0}, \qquad \operatorname{div}\operatorname{curl}\mathbf{v} = 0,$

$\operatorname{curl}\operatorname{curl}\mathbf{v} = \nabla(\nabla\cdot\mathbf{v}) - \boldsymbol{\Delta}\mathbf{v}$

$= \Big(\frac{\partial^2 v_2}{\partial x_1\partial x_2} + \frac{\partial^2 v_3}{\partial x_1\partial x_3} - \frac{\partial^2 v_1}{\partial x_2^2} - \frac{\partial^2 v_1}{\partial x_3^2}, \quad \frac{\partial^2 v_1}{\partial x_1\partial x_2} + \frac{\partial^2 v_3}{\partial x_2\partial x_3} - \frac{\partial^2 v_2}{\partial x_1^2} - \frac{\partial^2 v_2}{\partial x_3^2}, \quad \frac{\partial^2 v_1}{\partial x_1\partial x_3} + \frac{\partial^2 v_2}{\partial x_2\partial x_3} - \frac{\partial^2 v_3}{\partial x_1^2} - \frac{\partial^2 v_3}{\partial x_2^2}\Big).$

B Bessel function formulas

All these formulas can be found in [DLMF, §10] and [CK2, §3.4]. Here $r>0$ and $\ell\in\mathbb{Z}$.

$r^2 f''(r) + rf'(r) + (r^2-\ell^2)f(r) = 0, \qquad$ Bessel differential equation, $f\in\{J_\ell, Y_\ell, H_\ell^{(1)}, H_\ell^{(2)}\},$

$J_\ell(r) = \Big(\dfrac{r}{2}\Big)^\ell \sum_{j=0}^{\infty}(-1)^j \dfrac{(\frac{1}{4}r^2)^j}{j!(\ell+j)!} \quad \ell\in\mathbb{N}_0,$

$J_{-\ell} = (-1)^\ell J_\ell, \qquad Y_{-\ell} = (-1)^\ell Y_\ell, \qquad J_\ell' = \dfrac{J_{\ell-1}-J_{\ell+1}}{2}, \qquad Y_\ell' = \dfrac{Y_{\ell-1}-Y_{\ell+1}}{2},$

$H_\ell^{(1)} := J_\ell + \mathrm{i}Y_\ell, \qquad H_\ell^{(2)} := J_\ell - \mathrm{i}Y_\ell = \overline{H_\ell^{(1)}}, \qquad \dfrac{\partial}{\partial r}|H_\ell^{(1)}(r)| < 0,$

$\mathrm{e}^{\mathrm{i}r\cos\alpha} = \sum_{\ell\in\mathbb{Z}} \mathrm{i}^\ell J_\ell(r)\mathrm{e}^{\mathrm{i}\ell\alpha} \qquad$ Jacobi–Anger formula, $\alpha\in\mathbb{R},$

$J_\ell(r) = \sqrt{\dfrac{2}{\pi r}}\cos\Big(r - \dfrac{\ell\pi}{2} - \dfrac{\pi}{4}\Big)\Big(1 + \mathcal{O}_{r\to\infty}\Big(\dfrac{1}{r}\Big)\Big), \quad Y_\ell(r) = \sqrt{\dfrac{2}{\pi r}}\sin\Big(r - \dfrac{\ell\pi}{2} - \dfrac{\pi}{4}\Big)\Big(1 + \mathcal{O}_{r\to\infty}\Big(\dfrac{1}{r}\Big)\Big),$

$J_\ell(r) \sim \dfrac{1}{\sqrt{2\pi}}\Big(\dfrac{\mathrm{e}r}{2}\Big)^\ell \ell^{-\ell-\frac{1}{2}}, \qquad H_\ell^{(1)}(r) \sim -\mathrm{i}\sqrt{\dfrac{2}{\pi}}\Big(\dfrac{2}{\mathrm{e}r}\Big)^\ell \ell^{\ell-\frac{1}{2}} \qquad \text{for } \ell\to\infty,$

$J_\ell(r) \sim \dfrac{r^\ell}{\ell!\,2^\ell} \quad \ell\in\mathbb{N}_0, \qquad H_0^{(1)}(r) \sim \dfrac{2\mathrm{i}}{\pi}\log r, \qquad H_\ell^{(1)}(r) \sim -\dfrac{\mathrm{i}}{\pi}(\ell-1)!\dfrac{2^\ell}{r^\ell} \quad \ell\in\mathbb{N} \qquad \text{for } r\searrow 0.$

Here $a(x)\sim b(x)$ for $x\to X$ means that $\lim_{x\to X}\frac{a(x)}{b(x)} = 1$.

C List of acronyms

- **BCs**: boundary conditions.
- **BEM**: boundary element method.
- **BIE**: boundary integral equation.
- **BIO**: boundary integral operator.
- **BVP**: boundary value problem.
- **DOFs**: degrees of freedom.
- **DtN**: Dirichlet-to-Neumann map.
- **EDP**: exterior Dirichlet problem.
- **FEM**: finite element method.
- **PDE**: partial differential equation.
- **PEC**: perfect electric conductor.
- **SSSP**: sound-soft scattering problem.
- **TE**: transverse-electric.
- **TEM**: transverse-electric and magnetic.
- **TM**: transverse-magnetic.

D Summary

In this course we studied the boundary element method (BEM) for the numerical approximation of sound-soft scattering problems for the homogeneous Helmholtz equation in two dimensions.

The Helmholtz equation $\Delta u + k^2 u = 0$ is relevant because it characterizes the space dependence $u(\mathbf{x})$ of time-harmonic solutions $U(\mathbf{x},t) = \Re\{u(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t}\}$ of the wave equation $\frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} - \Delta U = 0$, where $k = \omega/c > 0$ is the wavenumber, ω the time frequency and c the wave speed (§1.2). The wave equation models the propagation and the scattering of acoustic waves (§1.1). The Helmholtz equation arises also in the modelling of electromagnetic (§1.3) and elastic (§1.4) waves.

The solutions of the Helmholtz equation that are separable in Cartesian coordinates are the plane waves (§2.2), which are either propagative or evanescent (§2.2.1). The solutions that are separable in polar coordinates are the circular waves, which are products of Bessel (J_ℓ, Y_ℓ) and Hankel ($H_\ell^{(1)}$, $H_\ell^{(2)}$) functions (§2.3) in the radial variable r (times k) and circular harmonics $\mathrm{e}^{\mathrm{i}\ell\theta}$ in the angular variable.

We can easily compute by hand the reflection of a plane wave hitting an infinite straight line equipped with Dirichlet, Neumann, or impedance conditions (§4.1). In order to study the wave scattering by bounded obstacles we need to deal with Lipschitz domains (§3.1), function spaces defined on them (§3.2) and on their boundaries (§3.3) and Green's identities (§3.4). With these tools we can formulate the exterior Dirichlet problems, and in particular the sound-soft scattering problems (§4.3.2). The simplest example is given by a circular scatterer (§4.3.1), for which we can write the solution explicitly. The key condition "at infinity", used to select the correct solution, is the Sommerfeld radiation condition $|\partial_r u - \mathrm{i}ku| = o_{r\to\infty}(r^{-1/2})$.

Using the fundamental solution Φ_k of the Helmholtz equation, we define the single-layer potential $\mathcal{S}$ and the single-layer operator S. These allow to write the boundary integral equation (BIE) $S\psi = g_\mathrm{D}$ and the representation formula $u = \mathcal{S}\psi$. Solving the BIE and applying the representation formula we obtain the solution of the exterior Dirichlet problem (§5.1). The BIE can be discretised with a collocation-BEM or a Galerkin-BEM (§5.2). The implementation of the BEM requires a careful use of quadrature formulas (§5.2.1). Several variations and extensions of the BEM are possible (§5.2.2).

The analysis of the Helmholtz equation (in the form of boundary value problem or BIE) involves non-coercive variational problems that admit Gårding inequalities, and relies on Fredholm theory (§3.5). This allows to study BVPs posed both in bounded domains (§4.2), which are closely related to Laplace eigenvalue problems, and in exterior domains (§4.3). An important formula is Green's representation, which allows to write Helmholtz solutions in terms of their traces (§5.3). One can also define the double-layer potential $\mathcal{D}$ and operator D (§5.4), the adjoint double-layer operator D' and the hypersingular operator H (§5.5). The two potentials and the four operators are related to one another by the Dirichlet and Neumann trace operators (60), which determine the jump relations (61).

The well-posedness of the single-layer BIE (§6.1) follows from the injectivity and the Fredholm property of S, The injectivity holds only when k^2 is not a Laplace eigenvalue (§6.1.1). The Fredholm property (§6.1.2) is obtained by decomposing S in the sum of a coercive part related to the reaction-diffusion equation (§6.1.5) and a compact part corresponding to a bounded kernel function (§6.1.4).

Many other integral equations are possible for the same exterior Dirichlet problem (§6.2), some of which are well-posed for all values of k (§6.2.3, §6.2.4).

Galerkin discretisations of non-coercive problems that satisfy a Gårding inequality are well-posed and quasi-optimal if the discrete space is "sufficiently fine" (§6.3.1). This applies to both the finite element method (§6.3.2) and the BEM (§6.3.3) approximations of Helmholtz problems.

Contents

References

References on wave phenomena. [CJ77] is classical book that describes very clearly and succinctly many kinds of wave phenomena, developing both physical intuition and mathematical formalism; [BK00] is similar in spirit but more advanced and up to date. Plenty of animations illustrating concepts from acoustics, mechanical vibrations, and waves can be found on the website [Rus]. [Hel13] is a very clear and instructive (non-mathematical) book to gain an intuitive understanding of acoustic phenomena.

References on the Helmholtz equation. [Ihl98] describes in detail the Helmholtz BVPs and their discretisation with the finite element method. Several Helmholtz (and Maxwell) BVPs and the corresponding boundary integral equations (BIEs) are analysed in mathematically rigorous way in [CK1, §3], [CK2, §1–3], [Néd01, §2–3]. [Spence14] is a survey of several variational formulations for Helmholtz and Laplace BVPs, the corresponding BIEs and the numerical methods for their discretisation; it is a very clear introduction to Helmholtz problems and the related literature. The book [SBH19] is an excellent PDE textbook: it is written with applications to numerical methods in mind and it includes a detailed analysis of the Helmholtz equation.

References on the BEM. The lecture notes [Sayas15] and [Sayas06] are good introductions to BIEs for Laplace and Helmholtz equations, respectively, and their discretisations with the boundary element method (BEM). [Cos87] is a brief and clear introduction to the BEM for several PDEs. [SS11] is a comprehensive and mathematically-oriented textbook on the BEM for elliptic PDEs, including Helmholtz. Section 2 of [CGLS12] is devoted to the BIE formulation of 2D Dirichlet–Helmholtz problems (precisely those we mostly focus on in this notes); the rest of the article analyses in detail a special class of BEM for the same problems.

[BK00] J. Billingham, A.C. King, *Wave motion*, Cambridge University Press, 2000.

[Brezis11] H. Brezis, *Functional analysis, Sobolev spaces and PDEs*, Springer, 2011.

[CGLS12] S.N. Chandler-Wilde, I.G. Graham, S. Langdon, E.A. Spence, *Numerical-asymptotic boundary integral methods in high-frequency acoustic scattering*, Acta Numerica, 21:89–305, 2012.

[CJ77] C.A. Coulson, A. Jeffrey, *Waves: a mathematical approach to the common types of wave motion*, Longman, 1977.

[CK1] D. Colton, R. Kress, *Integral equation methods in scattering theory*, SIAM, 2013, reprint of 1983 original.

[CK2] D. Colton, R. Kress, *Inverse acoustic & electromagnetic scattering theory*, Springer, 2013, 3rd ed.

[Cos87] , M. Costabel, *Principles of boundary element methods*, North-Holland 1987.
Available on https://perso.univ-rennes1.fr/martin.costabel/publis/Co_PrinciplesBEM.pdf

[DLMF] *NIST Digital Library of Mathematical Functions*, https://dlmf.nist.gov/

[Hel13] E.J. Heller, *Why you hear what you hear*, Princeton University Press, 2013.

[Ihl98] F. Ihlenburg, *Finite element analysis of acoustic scattering*, Springer, 1998.

[Néd01] J.-C. Nédélec, *Acoustics and electromagnetic equations. Integral representations for harmonic problems*, Springer, 2001.

[Rus] D.A. Russell, *Acoustics and vibration animations*,
https://www.acs.psu.edu/drussell/demos.html

[Sayas06] F.-J. Sayas, *Introduction to the boundary element method. A case study: the Helmholtz equation*, notes 2006. Available on https://team-pancho.github.io/classnotes.html

[Sayas15] F.-J. Sayas, *First steps in boundary element method*, notes 2015.
Available on https://team-pancho.github.io/classnotes.html

[SBH19] F.-J. Sayas, T.S. Brown, M.E. Hassell, *Variational Techniques for Elliptic Partial Differential Equations*, CRC Press, 2019.

[Spence14] E.A. Spence, *"When all else fails, integrate by parts" – an overview of new and old variational formulations for linear elliptic PDEs*, preprint 2014.
Available on https://people.bath.ac.uk/eas25/ibps.pdf

[SS11] S.A. Sauter, C. Schwab, *Boundary Element Methods*, Springer, 2011.