Lecture Notes for 2011 course

Introduction to Inverse Problems

Guillaume Bal [1]

[1]Departments of Statistics and Mathematics and CCAM, University of Chicago, Chicago, IL 60637; guillaumebal@uchicago.edu

Contents

Chapter 1

Introduction to Inverse Problems

1.1 What constitutes and *Inverse Problem* (IP)

Let us first schematically describe the components of what constitutes an *Inverse Problem* in these notes.

1. The *Measurement Operator (MO)*.

   This the operator that maps objects of interest, which we shall call parameters, to the available information we have about them, which we shall call *data* or *measurements*.

2. *Injectivity*, regularizing properties of MO and related *stability estimates*.

   Injectivity addresses the question of the uniqueness of the reconstruction of the parameters from knowledge of the MO. Regularization theory and related stability estimates describe the functional setting in which the reconstruction of the parameters is well-behaved, in the sense that differences in the reconstruction are *controlled* a priori by a function of the differences in the measurements.

3. "Noise" and Modeling.

   Knowledge of the MO may not be sufficient to solve an IP *satisfactorily* (which is, as I'm sure you noticed, a very subjective notion). Often, various "noisy" contributions need to be added to realistically model the IP of interest. They need to be added because they may have *undesirable* effects on the reconstructions (which makes the reconstruction not *satisfactory*, again, a very subjective notion).

   (a) Errors in Modeling

   The MO may be an approximation of the "real" map between objects of interest and "real" measurements. Sometimes, such approximations constitute the main "noise" contribution.

   (b) Errors in measurements

   This is because detectors are not perfect. It is important to model such "noise" contributions to estimate their effect on reconstructions.

(c) Prior Assumptions on IP Solution and Noise

This item is probably the most important one from a practical standpoint. The very reason "noise" is modeled is because it has an *undesirable* effect. In order to avoid this effect, the whole IP needs to be modified: since available data are not sufficiently rich to allow for *desirable* reconstructions, we need to be more restrictive on what we aim to reconstruct. It is necessary to make *prior* assumptions on the object of interest. There are two main, related, methods to do so.

i. *Penalization theory* is a deterministic methodology that restricts the domain of definition of the parameters. It includes two subcategories, one being *regularization theory*, which assumes that the parameters of interest are "sufficiently smooth", and the other one being *sparsity theory*, which assumes that the parameters are sparsely represented in a given, specific, basis.

ii. The *Bayesian framework* is an alternative, very versatile, methodology to incorporate prior assumptions in a statistical fashion. A *prior* probability density describes the potential values that the parameters can take. A *posterior* probability density describes how these potential values are affected by the measurements.

4. Numerical implementation.

(a) Linear Algebra to solve Linear systems of equations. Cheapest computationally. Should be tried first.

(b) Optimization methods to solve minimization problems such as, e.g., linear and quadratic programming problems. More expensive computationally but still relatively inexpensive and provide extremely accurate results when hypotheses of penalization theory are met.

(c) Monte Carlo methods such as Markov Chain Monte Carlo methods (MCMC) to solve IPs in Bayesian framework. By far the most expensive methodology, which should be used when the other two fail.

(d) Less standard numerical implementations necessary when IP solution parameterized in very specific ways, such as, e.g., small inclusions or point scatterers.

1.2 Some Applications of Inverse Problems

Inverse Problems are ubiquitous in Science and Engineering. This course will focus on some important examples that find applications, e.g., in medical and geophysical imaging. With a few exceptions, we shall be primarily concerned with injective inverse problems, which means inverse problems in which the MO is injective: in the absence of any "noise", a unique set of parameters corresponds to the available measurements.

Such inverse problems may be broadly classified into three different categories. The first category concerns the *well-posed* inverse problems, in which "noise" has a very

limited effect on the reconstruction proportional to its size. The second category is referred to as the *mildly ill-posed* inverse problems. There, the effect of noise is more important so that prior assumptions on the parameters we wish to reconstruct become necessary. The third category of inverse problems is called the *severely ill-posed* inverse problems. They encompass the class of inverse problems in which even limited amounts of noise generate very large undesirable effects. These are the problems in which prior information becomes paramount. Bayesian-based reconstructions are typically useful for such severely ill-posed problems as well as for inverse problems where injectivity no longer holds.

Well-posed inverse problems are relatively rare in practical applications. Examples include Magnetic Resonance Imaging, one dimensional inverse scattering problems in the time domain, and some hybrid inverse problems with internal measurements.

Mildly ill-posed inverse problems form the vast majority of practical inverse problems in geophysical and medical imaging. An incomplete list of examples include X-ray computerized tomography, positron emission tomography, Single photon emission computerized tomography, Ultrasound tomography, inverse kinematic problems, as well as many more recent hybrid imaging modalities such as photo-acoustic tomography, ultrasound modulated tomography, transient elastography.

Finally, severely ill-posed problems involve all the measurement operators involving a diffusion equation that prevents any singularities to propagate.

In the rest of this lecture, we present a simplified modeling of Magnetic Resonance Imaging that gives rise to several possible inverse problems. The last of these inverse problems is nothing but inverting a Fourier transform, which is a well-posed problem. The notion of well-posed versus ill-posed problems is also relatively subjective. As we mentioned, an IP is ill-posed because "noise" generates "unsatisfactory" reconstructions. This notion is quite subjective. A possible definition of what it is for an IP to be ill-posed uses the notion of Fourier transforms and Hilbert scale of functional spaces. We first review the necessary material in section 1.5 and then define possible notions of well-posed and ill-posed problems in section 1.6. We insists once more that these notions are

1.3 Some examples of Measurement Operators

The Measurement Operator will be denoted by $\mathfrak{M}$. It is an operator mapping parameters in $\mathfrak{D}(\mathfrak{M}) \subset \mathfrak{X}$, where $\mathfrak{D}(\mathfrak{M})$ is the domain of definition of $\mathfrak{M}$ and is a subset of a functional space $\mathfrak{X}$, typically a Banach space or a Hilbert space. We assume that $\mathfrak{M}$ takes values in $\mathfrak{Y}$, the space of *data*. So we shall be writing

$$\mathrm{y} = \mathfrak{M}(\mathrm{x}) \quad \text{ for } \mathrm{x} \in \mathfrak{D}(\mathfrak{M}) \subset \mathfrak{X} \text{ and } \mathrm{y} \in \mathfrak{Y}.$$

Solving the inverse problem therefore amounts to finding the point $\mathrm{x} \in \mathfrak{X}$ from knowledge of the data $\mathrm{y} \in \mathfrak{Y}$.

For most measurement operators throughout these lectures, we assume that $\mathfrak{M}$ is **injective**. This means that

$$\mathfrak{M}(\mathrm{x}_1) = \mathfrak{M}(\mathrm{x}_2) \qquad \Longrightarrow \qquad \mathrm{x}_1 = \mathrm{x}_2. \tag{1.1}$$

In other words, data y, if given in the range of $\mathfrak{M}$, uniquely characterize parameter x. However, we shall see that injectivity is not sufficient to make an inverse problem "well-posed". It remains to understand how noise added to the measurement operator are handled by a reconstruction. This is what stability estimates will allow us to quantify in a precise manner.

Example 1. Let $\mathfrak{X} = \mathcal{C}([0,1])$, the space of continuous functions. Let $\mathfrak{Y} = \mathfrak{X}$ and define

$$\mathfrak{M}(f)(x) = \int_0^x f(y)dy.$$

Here, the point x in $\mathfrak{X}$ is the function f, a function traditionally denoted by $f(x)$. So the two "x" appearing in the same line do not have the same meaning. The first x is a point in $\mathfrak{X}$, i.e., a function, while the second x is simply a point in the interval $[0,1]$.

At any rate, $\mathfrak{M}$ is certainly injective since the equality of data $\mathfrak{M}(f) = \mathfrak{M}(g)$ implies that $f = \frac{d}{dx}\mathfrak{M}(f) = \frac{d}{dx}\mathfrak{M}(g) = g$, i.e., the equality of parameters.

However, the operator $\mathfrak{M}$ is "smoothing" since $\mathfrak{M}(f)$ is one derivative more regular than f. So inverting the operator, as indicated above, involves differentiating the data. If "*noise*" is added to the data and "*noise*" is high frequency, the the derivative of "*noise*" will be large and may overwhelm the reconstruction of $f(x)$. The objective of penalization theory will precisely to make sure this *undesirable* fact does not happen.

Example 2. On $\mathfrak{X} = \mathfrak{Y} = \mathcal{C}([0,1])$ and with $\mathfrak{D}(\mathfrak{M}) = \mathcal{C}^1_0([0,1])$, the space of continuously differentiable functions with value 0 at $x = 0$, we can define

$$\mathfrak{M}(f)(x) = f'(x), \qquad \text{the derivative of } f.$$

Note that the derivative of a continuous function is not continuous (and may not exist as a function although it can be defined as a distribution). So here, we found it convenient to define the domain of definition of $\mathfrak{M}$ as a subset of the space of continuous functions. We could obviously have defined $\mathfrak{X}$ directly as the space of continuously differentiable functions vanishing at 0, but then $\mathfrak{X}$ would be different from $\mathfrak{Y}$. For "unbounded" operators such as $\mathfrak{M}$, it is often convenient to keep $\mathfrak{X} = \mathfrak{Y}$ and restrict the domain of definition $\mathfrak{D}(\mathfrak{M})$.

So why do we also insist on $f(0) = 0$ in $\mathfrak{D}(\mathfrak{M})$? It is because otherwise, $\mathfrak{M}$ would *not* be injective. Indeed, antiderivatives are defined up to a constant. By enforcing $f(0) = 0$, the constant is fixed. In $\mathfrak{D}(\mathfrak{M})$, we have from the fundamental theory of calculus that

$$f(x) = f(0) + \int_0^x f'(y)dy = \int_0^x f'(y)dy = \int_0^x \mathfrak{M}(f)(y)dy.$$

Now obviously, $\mathfrak{M}(f) = \mathfrak{M}(g)$ implies $f = g$. From the point of view of inverse problems, this operator is very nice. If "*noise*" is added to $\mathfrak{M}(f)$, it will be *integrated* during the reconstruction, which is a much more stable process than differentiating.

Example 3. With the same setting as in Example 1, consider $\mathfrak{M}(f)(x) = \int_0^1 f(y)dy$. This operator is well defined and very nice. But it is clearly not injective. All we learn about f is its mean on $(0,1)$. This operator corresponds to data that are not very informative.

Example 4. Let $\mathfrak{X} = \mathcal{C}_c(\mathbb{R}^2)$ the space of continuous functions with compact support (functions supported in a bounded domain in $\mathbb{R}^2$, i.e., vanishing outside of that domain). Let $\mathfrak{Y} = \mathcal{C}(\mathbb{R} \times (0, 2\pi))$. We define $l(s, \theta)$ for $s \in \mathbb{R}$ and $\theta \in (0, 2\pi)$ as the *line* with direction perpendicular to $\boldsymbol{\theta} = (\cos\theta, \sin\theta)$ and at a distance $\pm s$ from the origin $(0, 0)$. More precisely, let $\boldsymbol{\theta}^\perp = (-\sin\theta, \cos\theta)$ the rotation of $\boldsymbol{\theta}$ by $\frac{\pi}{2}$. Then

$$l(s, \theta) = \{\mathbf{x} \in \mathbb{R}^2 \text{ such that } \mathbf{x} = s\boldsymbol{\theta} + t\boldsymbol{\theta}^\perp \quad \text{for } t \in \mathbb{R}\}.$$

We define

$$\mathfrak{M}(f)(s, \theta) = \int_{l(s,\theta)} f(x)dl = \int_{\mathbb{R}} f(s\boldsymbol{\theta} + t\boldsymbol{\theta}^\perp)dt,$$

where dl is the line measure along $l(s, \theta)$. In other words, $\mathfrak{M}$ maps a function to the value of its integrals along *any* line. The operator $\mathfrak{M}$ is the **two-dimensional Radon Transform**. We shall see that $\mathfrak{M}$ is injective.

Example 5. Let us conclude with a more involved, though practically very relevant, example. We first introduce the following elliptic partial differential equation

$$\begin{array}{ll} -\nabla \cdot \sigma(x)\nabla u(x) = 0, & x \in X \\ u(x) = g(x) & x \in \partial X, \end{array} \tag{1.2}$$

where X is a smooth domain in $\mathbb{R}^n$, ∂X its boundary, $\sigma(x)$ a smooth coefficient in X bounded above and below by positive constants, and $g(x)$ is a prescribed Dirichlet data for the elliptic problem. This equation is a standard forward problem and it is known that it admits a unique solution. Moreover, the outgoing current

$$j(x) = \sigma(x)\frac{\partial u}{\partial \nu}(x), \quad \text{with } \nu(x) \text{ the outward unit normal to } X \text{ at } x \in \partial X,$$

is a well defined function. This allows us to define the Dirichlet-to-Neumann (aka Poincaré-Steklov) operator

$$\Lambda : \begin{array}{rcl} H^{\frac{1}{2}}(\partial X) & \to & H^{-\frac{1}{2}}(\partial X) \\ g(x) & \mapsto & \Lambda[g](x) = j(x) = \sigma(x)\dfrac{\partial u}{\partial \nu}(x). \end{array} \tag{1.3}$$

Here, $H^s(\partial X)$ are standard Hilbert spaces of distributions defined at the domain's boundary.

Let now $\mathfrak{X} = \mathcal{C}^2(\bar{X})$ and $\mathfrak{Y} = \mathcal{L}(H^{\frac{1}{2}}(\partial X), H^{-\frac{1}{2}}(\partial X))$, where the space $\mathcal{L}(\mathfrak{X}_1, \mathfrak{X}_2)$ means the space of linear bounded (continuous) operators from $\mathfrak{X}_1$ to $\mathfrak{X}_2$. We now define the measurement operator

$$\mathfrak{M} : \quad \mathfrak{X} \ni \sigma \mapsto \mathfrak{M}(\sigma) = \Lambda[\sigma] \in \mathfrak{Y}. \tag{1.4}$$

So the measurement operator maps the unknown conductivity σ to the Dirichlet-to-Neumann operator Λ, which is by construction an functional of σ since the solution u of (1.2) is a functional of σ. The measurement operator therefore lives in a "huge" (admittedly, this is again subjective) space. Acquiring data means acquiring Λ, which means for each and every possible function $g \in H^{\frac{1}{2}}(\partial X)$ at the domain's boundary, perform an experiment that measures $j(x)$.

One of the great milestones of inverse problem theory was to prove that $\mathfrak{M}$ above was an injective operator. The proof belongs to John Sylvester and Gunther Uhlmann and was published in 1987. Another more recent milestone was to prove that we can choose $\mathfrak{X} = L^\infty(X)$ in the two-dimensional setting. The proof belongs to Kari Astala and Lassi Païvarinta and was publised in 2006. What is the "best" (i.e., largest) $\mathfrak{X}$ that can be considered in dimension $n \geq 3$ is an important theoretical open problem.

1.4 Magnetic Resonance Imaging (MRI)

This section presents an extremely simplified version of MRI. By doing so, we will observe that MRI reconstructions may be modeled by (at least) three different inverse problems.

MRI exploits the precession of the spin of protons in a magnetic field $H(\mathbf{x})$, which is a vector in $\mathbb{R}^3$ for each position $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$. The axis of the precession is that of the magnetic field and the frequency of the precession is $\omega(\mathbf{x}) = \gamma H(\mathbf{x})$, where $\gamma = e/(2m)$ is called the gyromagnetic ratio, e is the electric charge of the proton and m its mass.

In a nutshell, MRI works as follows. Assume first that we impose a strong static magnetic field $\mathbf{H}_0 = H_0\mathbf{e}_z$ along the z axis. All protons end up with their spin parallel to $\mathbf{H}_0$ and slightly more so in the direction $\mathbf{H}_0$ than in $-\mathbf{H}_0$. This difference is responsible for a macroscopic *magnetization* $\mathbf{M}$ pointing in the same direction as $\mathbf{H}_0$.

In a second step, we generate radio frequency magnetic waves at the Larmor frequency $\omega_0 = \gamma|\mathbf{H}_0|$. In clinical MRI, the frequency is typically between 15 and 80 MHz (for hydrogen imaging), which corresponds to wavelengths between 20 and 120 m (since $\omega = ck = 2\pi c/\lambda$ and $c \approx 3\,10^8$). So the wavelength is not what governs spatial resolution in MRI. For instance the pulse (assumed to be independent of $\mathbf{x}$ to start with) may be of the form $\mathbf{H}_1(t) = 2H_1\cos(\omega_0 t)\mathbf{e}_x$ and turned on for a duration t_p. Because the field oscillates at the Larmor frequency, the spins of the protons are affected. The resulting effect on the macroscopic magnetization is that it precesses around the axis $\mathbf{e}_z$ at frequency ω_0. The spins make an angle with respect to the direction $\mathbf{e}_z$ given at time t_p by

$$\theta = \gamma H_1 t_p.$$

Generally, t_p is chosen such that $\theta = \pi/2$ or $\theta = \pi$. The corresponding pulses are called 90^0 and 180^0 pulses, respectively. Thus, after a 90^0 pulse, the magnetization oscillates in the xy plane and after a 180^0 pulse, the magnetization is pointing in the direction $-\mathbf{H}_0$.

Once the radio frequency is turned off (but not the static field $\mathbf{H}_0$), protons tend to realign with the static field $\mathbf{H}_0$. By doing so, they emit radio frequency waves at the Larmor frequency ω_0 that can be *measured*. This wave is called the free induction decay (FID) signal. The FID signal after a 90^0 pulse will have the form

$$S(t) = \rho e^{i\omega_0 t} e^{-t/T_2}. \tag{1.5}$$

Here ρ is the density of the magnetic moments and T_2 is the spin-spin relaxation time. (There is also a spin-lattice relaxation time $T_1 \ll T_2$, which cannot be imaged with 90^0 pulses and which we ignore.) The main reason for doing all this is that the density

ρ and the relaxation time T_2 depend on the tissue sample. We restrict ourselves to the reconstruction of ρ here, knowing that similar experiments can be devised to image T_2 (and T_1) as well. To simplify, we assume that measurements are performed over a period of time that is small compared to T_2 so that the exponential term e^{-t/T_2} can be neglected.

Now human tissues are not spatially homogeneous, which makes imaging a lot more useful. The density of magnetic moments $\rho = \rho(\mathbf{x})$ depends on type of tissue at $\mathbf{x} \in \mathbb{R}^3$. This is the parameter we wish to reconstruct.

The physical mechanism that allows for the good spatial resolution of MRI (sub-millimeter resolution for brain imaging) is that only tissue samples under a static magnetic field $\mathbf{H}$ such that $|\mathbf{H}| = \gamma\omega_0$ will be affected by the radio frequency pulse $\mathbf{H}_1(t) = 2H_1 \cos(\omega_0 t)\mathbf{e}_x$. We thus need to make sure that the static field has the correct amplitude in as small a spatial area as possible. To do so, we impose the static field

$$\mathbf{H}(z) = \mathbf{H}_0 + G_z z \mathbf{e}_z.$$

Only those protons in the slice with z is close to 0 will be affected by the pulse $\mathbf{H}_1$ since we have assumed that $|\mathbf{H}_0| = \gamma\omega_0$. As a consequence, the measured signal takes the form

$$S(t) = e^{i\omega_0 t} \int_{\mathbb{R}^2} \rho(x, y, 0) dx dy \quad \text{so that} \quad e^{-i\omega_0 t} S(t) = \int_{\mathbb{R}^2} \rho(x, y, 0) dx dy. \tag{1.6}$$

The above right-hand-side thus obtain the average density in the plane $z = 0$. MRI is thus a tomographic technique (tomos meaning slice in Greek).

By changing $\mathbf{H}_0$ or ω_0, we can obtain the average density in the plane $z = z_0$ for all values of $z_0 \in \mathbb{R}$. Moreover, by rotating the generated magnetic field $\mathbf{H}(z)$, we are ideally able to obtain the average density in any plane in $\mathbb{R}^3$. Planes may be parameterized by their normal vector $\phi \in \mathbb{S}^2$, with $\mathbb{S}^2$ the unit sphere in $\mathbb{R}^3$, and their distance s to the origin $(0, 0, 0) \in \mathbb{R}^3$. Let $P(s, \phi)$ be the corresponding plane. Then what we have obtained is that MRI experiments allowed us to obtain the plane integrals of the density

$$R(s, \phi) = \int_{P(s,\phi)} \rho(\mathbf{x}) d\sigma(\mathbf{x}), \tag{1.7}$$

where $d\sigma(\mathbf{x})$ is the surface (Euclidean) measure on the plane $P(s, \phi)$. Here, $R(s, \phi)$ is the **three-dimensional Radon transform** of $\rho(\mathbf{x})$. This is the first inverse problem we encounter. The measurement operator maps functions defined on $\mathbb{R}^3$ (for instance compactly supported continuous functions) to the Radon transform, which is a function (for instance compactly supported continuous function) of $(s, \phi) \in \mathbb{R} \times \mathbb{S}^2$. We thus have the Inverse Problem:

3D Radon Transform:	Reconstruct density $\rho(\mathbf{x})$ from knowledge of $R(s, \phi)$
Equivalently:	Reconstruct a function from its plane integrals.

There are several issues with the above inverse problem. First of all, the Radon transform is a mildly ill-posed problem as we shall see, whereas in the ideal setting described below, MRI can be modeled as a well-posed inverse problem. More fundamentally from a practical point of view, rotating the whole magnetic field from the

direction e_z to an arbitrary direction ϕ is very challenging technologically. One could also rotate the object of interest rather than the heavy magnet. However, for imaging of human patients, this is not feasible either for rather obvious reasons. Additional modeling is therefore necessary.

So far, we have a vertical discrimination of the proton density. The transversal discrimination is obtained by imposing a static field linearly varying in the x and y directions. Remember that after the 90^0 pulse, the magnetization $\mathbf{M}(x, y, 0)$ rotates with frequency ω_0 in the xy plane (i.e., is orthogonal to $\mathbf{e}_z$), and is actually independent of x and y. Let us now impose a static field $\mathbf{H}(y) = \mathbf{H}_0 + G_y y \mathbf{e}_z$ for a duration T. Since the frequency of precession is related to the magnetic field, the magnetization at position y will rotate with frequency $\omega(y) = \omega_0 + \gamma G_y y$. Therefore, compared to the magnetization at $z = 0$, the magnetization at z will accumulate a phase during the time T the field $G_y y \mathbf{e}_z$ is turned on given by $T(\omega(y) - \omega_0) = T\gamma G_y y$. Once the field $G_y y \mathbf{e}_z$ is turned off, the magnetization will again rotate everywhere with frequency ω_0. However, the phase depends on position y. This part of the process is call *phase encoding*. A measurement of the FID would then give us a radio frequency signal of the form

$$S(t; T) = e^{i\omega_0 t} \int_{\mathbb{R}^2} e^{i\gamma G_y T y} \rho(x, y, 0) dx dy. \tag{1.8}$$

By varying the time T or the gradient G_y, we see that we can obtain the frequency content in y of the density $\rho(x, y, 0)$. More precisely,

$$\int_{\mathbb{R}} \rho(x, y, 0) dx = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ik_y y} S(t; \frac{k_y}{\gamma G_y}) dk_y. \tag{1.9}$$

This provides us with the *line integrals* of $\rho(x, y, 0)$ at $z = 0$ for all the lines that are parallel to the x-axis. Note that the phase encoding was performed by using a field that is linear in the y variable. We can use a field that is linear in the variable $\cos\theta x + \sin\theta y$ instead. Denoting by $\boldsymbol{\theta} = (\cos\theta, \sin\theta) \in \mathbb{S}^1$ a unit vector in $\mathbb{R}^2$, by $\boldsymbol{\theta}^\perp = (-\sin\theta, \cos\theta)$ its rotation by $\frac{\pi}{2}$, and by $l(s, \theta)$ the line of normal $\boldsymbol{\theta}$ at a distance s from the origin $(0, 0) \in \mathbb{R}^2$, we are thus able to measure all line integrals of the function $\rho(x, y, 0)$:

$$R(s, \theta) = \int_{l(s,\theta)} \rho(x, y, 0) dl(x, y) = \int_{\mathbb{R}} \rho(s\boldsymbol{\theta} + t\boldsymbol{\theta}^\perp) dt, \tag{1.10}$$

where $dl(x, y)$ is the line (Euclidean) measure on $l(s, \theta)$. This is the second inverse problem we encounter: we wish to reconstruct $\rho(x, y)$ from knowledge of its line integrals $R(s, \theta)$. This is the **two-dimensional Radon transform** of $\rho(x, y, 0)$. The measurement operator maps functions defined on $\mathbb{R}^2$ (for instance compactly supported continuous functions) to the Radon transform, which is a function (for instance compactly supported continuous function) of $(s, \phi) \in \mathbb{R} \times \mathbb{S}^1$. We thus have the Inverse Problem:

2D Radon Transform: Reconstruct density $\rho(x, y)$ from knowledge of $R(s, \theta)$
Equivalently: Reconstruct a function from its line integrals.

This inverse problem is still mildly ill-posed, as is the three-dimensional Radon transform. Although it is used in practical MRI, the missing information in the x

variable in measurements of the form (1.8) can be obtained by additional modeling. Indeed, nothing prevents us from adding an $x-$dependent static field during the FID measurements. Let us assume that after time T (where we reset time to be $t = 0$), we impose a static field of the form $\mathbf{H}(x) = \mathbf{H}_0 + G_x x \mathbf{e}_z$. The magnetization will now precess around the z axis with $x-$dependent frequency $\omega(x) = \omega_0 + \gamma G_x x$. This implies that the measured signal will be of the form

$$S(t;T) = \int_{\mathbb{R}^2} e^{i\gamma G_y T y} e^{i(\omega_0+\gamma G_x x)t} \rho(x,y,0) dxdy. \tag{1.11}$$

We have thus access to the *measured data*

$$d(k_x,k_y) = e^{-i\omega_0 k_x/(\gamma G_x)} S(\frac{k_x}{\gamma G_x}; \frac{k_y}{\gamma G_y}) = \int_{\mathbb{R}^2} e^{ik_y y} e^{ik_x x} \rho(x,y,0) dxdy. \tag{1.12}$$

By varying T (or G_y) and t and G_x, we can obtain the above information for essentially all values of k_x and k_y. This is our third Inverse Problem:

2D Fourier Transform: Reconstruct density $\rho(x,y,0)$ from knowledge of $d(k_x,k_y)$
Equivalently: Reconstruct a function from its plane wave decomposition.

This is a well-known problem whose solution involves applying the *Inverse Fourier Transform*

$$\rho(x,y,0) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-i(k_x x + k_y y)} d(k_x,k_y) dk_x dk_y. \tag{1.13}$$

Several approximations have been made to obtain this reconstruction formula. Within this framework, we see however that density reconstructions are relatively simple: all we have to do is to invert a Fourier transform. The above procedure can be repeated for all values of z providing the density $\rho($**x** everywhere.

The inverse problems derived above all involve reconstructions in $\mathbb{R}^d$ for $d = 2$ or $d = 3$ that are geometries that invariant by translation. The Fourier transform is therefore an ideal candidate to simplify the reconstructions. We shall see in Chapter 2 that the Radon transforms can indeed be inverted by using the Fourier transform.

1.5 Fourier transforms and well-posedness

We recall in this section some important facts about the Fourier transform and in the next section about some functional (Hilbert) spaces that we will use throughout this course.

Let $f(\mathbf{x})$ be a complex-valued function in $L^2(\mathbb{R}^n)$ for some $n \in \mathbb{N}^*$, which means a (measurable) function on $\mathbb{R}^n$ that is square integrable in the sense that

$$\|f\|^2 = \int_{\mathbb{R}^n} |f(\mathbf{x})|^2 d\mathbf{x} < \infty. \tag{1.14}$$

Here $\|f\|$ is the $L^2(\mathbb{R}^n)$-norm of f and $d\mathbf{x}$ the Lebesgue (volume) measure on $\mathbb{R}^n$. We define the Fourier transform of f as

$$\hat{f}(\mathbf{k}) = [\mathcal{F}_{\mathbf{x}\to\mathbf{k}} f](\mathbf{k}) = \int_{\mathbb{R}^n} e^{-i\mathbf{k}\cdot\mathbf{x}} f(\mathbf{x}) d\mathbf{x}. \tag{1.15}$$

Then we obtain that $\hat{f}(\mathbf{k}) \in L^2(\mathbb{R}^n)$ and the Fourier transform admits an inverse on $L^2(\mathbb{R}^n)$ given by

$$f(\mathbf{x}) = [\mathcal{F}^{-1}_{\mathbf{k}\to\mathbf{x}}\hat{f}](\mathbf{x}) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{i\mathbf{k}\cdot\mathbf{x}}\hat{f}(\mathbf{k})d\mathbf{k}. \tag{1.16}$$

More precisely we have the Parseval relation

$$(\hat{f},\hat{g}) = (2\pi)^n(f,g), \tag{1.17}$$

where the Hermitian product is given by

$$(f,g) = \int_{\mathbb{R}^n} f(\mathbf{x})\overline{g}(\mathbf{x})d\mathbf{x}. \tag{1.18}$$

Here $\overline{g}$ is the complex conjugate to g. This implies that

$$\|\hat{f}\| = (2\pi)^{n/2}\|f\|. \tag{1.19}$$

So up to the factor $(2\pi)^{n/2}$, the Fourier transform and its inverse are isometries.

Important properties of the Fourier transform are how they interact with differentiation and convolutions. Let $\alpha = (\alpha_1,\cdots,\alpha_n)$ be a multi-index of non-negative components $\alpha_j \geq 0$, $1 \leq j \leq n$ and let $|\alpha| = \sum_{i=1}^n \alpha_j$ be the length of the multi-index. We then define the differentiation D^α of degree $|\alpha|$ as

$$D^\alpha = \prod_{i=1}^n \frac{\partial^{\alpha_i}}{\partial x_i^{\alpha_i}}. \tag{1.20}$$

We then deduce from the definition (1.15) that

$$\mathcal{F}_{\mathbf{x}\to\mathbf{k}}[D^\alpha f](\mathbf{k}) \;=\; \Big(\prod_{j=1}^n (ik_j)^{\alpha_j}\Big)\,[\mathcal{F}_{\mathbf{x}\to\mathbf{k}}f](\mathbf{k}). \tag{1.21}$$

Let us now define the convolution as

$$f * g(\mathbf{x}) = \int_{\mathbb{R}^n} f(\mathbf{x}-\mathbf{y})g(\mathbf{y})d\mathbf{y}. \tag{1.22}$$

We then verify that

$$\begin{array}{rclcrcl} \mathcal{F}_{\mathbf{x}\to\mathbf{k}}(f*g) &=& \mathcal{F}_{\mathbf{x}\to\mathbf{k}}f\mathcal{F}_{\mathbf{x}\to\mathbf{k}}g, & \text{i.e.} & \widehat{f*g} &=& \hat{f}\hat{g}, \\ \mathcal{F}^{-1}_{\mathbf{k}\to\mathbf{x}}(\hat{f}*\hat{g}) &=& (2\pi)^n f\,g & \text{i.e.} & \hat{f}*\hat{g} &=& (2\pi)^d\,\widehat{fg}. \end{array} \tag{1.23}$$

So the Fourier transform *diagonalizes* differential operators (replaces them by multiplication in the Fourier domain). However Fourier transforms replace products by non-local convolutions.

The Fourier transform is a well-posed operator from $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n)$ since the inverse Fourier transform is also defined from $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n)$ and is bounded as shown in (1.19). The resulting important effect in practice is that the reconstructions encountered in the two preceding sections are *stable*. Stability is meant with respect to some noise in the data. Let us assume that we measure

$$\hat{d}(\mathbf{k}) = \hat{f}(\mathbf{k}) + \hat{N}(\mathbf{k}),$$

where we know that $\delta = \|\hat{N}\|$ is relatively small. Then the error in the reconstruction will also be of order δ in the $L^2(\mathbb{R}^n)$ norm. Indeed let $d(\mathbf{x})$ be the reconstructed function from the data $d(\mathbf{k})$ and $f(\mathbf{x})$ be the real function we are after. Then we have

$$\|d - f\| = \|\mathcal{F}^{-1}_{\mathbf{k}\to\mathbf{x}}\hat{d} - \mathcal{F}^{-1}_{\mathbf{k}\to\mathbf{x}}\hat{f}\| = \|\mathcal{F}^{-1}_{\mathbf{k}\to\mathbf{x}}(\hat{d} - \hat{f})\| = (2\pi)^{-n/2}\delta. \tag{1.24}$$

In other words, measurements errors can still be seen in the reconstruction. The resulting image is not perfect. However the error due to the noise has not been amplified too drastically.

1.6 Hilbert scale and degrees of ill-posedness

Well-posed Problems and Lipschitz Stability. Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces (nice functional spaces with a norm) and let A be a linear operator from $\mathfrak{X}$ to $\mathfrak{Y}$. For every $y \in \mathfrak{Y}$, we would like to solve the linear problem

$$\text{Find } x \text{ such that} \qquad Ax = y. \tag{1.25}$$

What we mean by a *well posed* problem is a problem such that A is invertible ($A^{-1}y$ is defined for all $y \in \mathfrak{Y}$) and of bounded inverse (i.e., $\|A^{-1}y\|_{\mathfrak{X}} \leq C\|y\|_{\mathfrak{Y}}$ for a constant C that depends on A but not on $y \in \mathfrak{Y}$).

Another way of stating this from the point of view of an inverse problem is to say that the error between two solutions x_1 and x_2 corresponding to two sets of data y_1 and y_2 satisfies that

$$\|x_1 - x_2\|_{\mathfrak{X}} \leq C\|y_1 - y_2\|_{\mathfrak{Y}}. \tag{1.26}$$

Let us stress that the choice of the spaces $\mathfrak{X}$ and $\mathfrak{Y}$ and the norms $\|\cdot\|_{\mathfrak{X}}$ and $\|\cdot\|_{\mathfrak{Y}}$ matters. The definition and the boundedness operator A^{-1} depends upon these choices.

An estimate of the form (1.26) is a *stability* estimate with Lipschitz constant C. We then say that the problem (1.25) is *Lipschitz-stable*. To a certain extent, the constant C in front of (1.26) is not so important conceptually because we can always define an equivalent norm on $\mathfrak{Y}$ equal to C times the existing norm. In other words, the constant is just a reflection of how we *measure* noise level. What is more fundamental with (1.26) from a practical point of view is this: When noise levels are reduced by a factor two, which sometimes may be performed by adding detectors or obtaining higher quality measurements, then (1.26) states that the in the worst possible case (since (1.26) for all possible $x_1 - x_2$, and hence C reflects the amplification of noise in the worst scenario), the error in the reconstruction will also be improved by a factor two.

Ill-posed problems and unbounded operators. The inverse Fourier transform is one of the few *well-posed* inversion problems we will encounter in this course. Most inverse problems are actually considered to be *ill-posed*. Being ill-posed **does not** mean that a problem cannot be solved. And it should be realized that the notion of ill-posedness is *subjective*.

In the case of the linear problem (1.25), ill-posedness means that either A is not invertible on the whole space $\mathfrak{Y}$ (i.e., the range of A defined by $\text{Range}(A) = A(\mathfrak{X})$ is a

proper subset of $\mathfrak{Y}$; that is to say, is not equal to $\mathfrak{Y}$), or that A^{-1} is not bounded from $\mathfrak{Y}$ to $\mathfrak{X}$. In other words, an Lipschitz equality such as (1.26) does not hold.

However, this notion very much depends on our choice of functional space, and hence the subjective aspect of "ill-posedness". The operator A^{-1} could very well be bounded from the other space $\mathfrak{Y}'$ to the other space $\mathfrak{X}'$, in which case A is ill posed from $\mathfrak{X}$ to $\mathfrak{Y}$ but well-posed from $\mathfrak{X}'$ to $\mathfrak{Y}'$.

The main reason an inverse problem is ill-posed is that the forward operator is smoothing from $\mathfrak{X}$ to $\mathfrak{Y}$ so that the image of $\mathfrak{X}$ is in fact a small subset of $\mathfrak{Y}$. Recall that the forward (or direct) operator is the operator that maps what we wish to reconstruct to the noise-free measured data. The operator A in (1.25) is an example of a forward operator. What we mean by smoothing is that Ax is "more regular" than x, in the sense that details (small scale structures) are attenuated by the forward mapping. This does not mean that the details cannot be reconstructed. In many cases they can. It means however that the reconstruction has to undo this smoothing, i.e. has to deregularize. This works as long as no noise is present in the data. However, as soon as the data are noisy (i.e., always in practice), the deregularization process has the effect of amplifying the noise in a way that can potentially be very harmful.

The answer to ill-posedness requires to impose some prior assumptions on the function we wish to reconstruct. The simplest example of an assumption is to assume that the function is sufficiently smooth. The smoothness assumptions are not too drastic for *mildly ill-posed*. They are very drastic for *severely ill-posed* problems.

In order to define what we mean by ill-posedness and quantify the degree of ill-posedness, we need a scale of spaces in which the "smoothing" of A can be measured. Many different scales can be used. We will use what is probably the simplest scale of function spaces, namely the scale of Hilbert spaces.

The scale of Hilbert spaces. Let $s \geq 0$ be a non-negative real-number. We define the scale of Hilbert spaces $H^s(\mathbb{R}^n)$ as the space of measurable functions $f(\mathbf{x})$ such that

$$\|f\|^2_{H^s(\mathbb{R}^n)} = \int_{\mathbb{R}^n} (1 + |\mathbf{k}|^2)^s |\mathcal{F}_{\mathbf{x}\to\mathbf{k}} f|^2(\mathbf{k}) d\mathbf{k} < \infty. \tag{1.27}$$

We verify that $H^0(\mathbb{R}^n) = L^2(\mathbb{R}^n)$ since the Fourier transform is an isometry. We also verify that

$$\{f \in H^1(\mathbb{R}^n)\} \quad \Longleftrightarrow \quad \Big\{ f \in L^2(\mathbb{R}^n) \text{ and } \frac{\partial f}{\partial x_i} \in L^2(\mathbb{R}^n),\ 1 \leq i \leq n \Big\}. \tag{1.28}$$

This results from (1.21). More generally the space $H^m(\mathbb{R}^n)$ for $m \in \mathbb{N}$ is the space of functions such that all partial derivatives of f of order up to m are square integrable. The advantage of the definition (1.27) is that it holds for real values of s. So $H^{\frac{1}{2}}(\mathbb{R}^n)$ is roughly the space of functions such that "half-derivatives" are square integrable. Notice also that s characterizes the degree of smoothness of a function $f(\mathbf{x})$. The larger s, the smoother the function $f \in H^s(\mathbb{R}^n)$, and the faster the decay of its Fourier transform $\hat{f}(\mathbf{k})$ as can be seen from the definition (1.27).

It is also useful to define the Hilbert scale for functions supported on subdomains of $\mathbb{R}^n$. Let X be a sufficiently smooth subdomain of $\mathbb{R}^n$. We define two scales. The first

scale is $H_0^s(X)$, defined as the closure of $C_0^\infty(X)$, the space of functions of class C^∞ with support in X (so these functions and all their derivatives vanish at the boundary of X), for the norm in $H^s(\mathbb{R}^n)$. Thus, $f \in H_0^s(X)$ can be described as the limit of functions $f_n \in C_0^\infty(\mathbb{R})$ uniformly bounded in $H^s(\mathbb{R}^n)$. We also define $H^s(X)$ as the space of functions f on X that can be extended to functions f^* in $H^s(\mathbb{R}^n)$ (i.e., $f = f^*\chi_X$, where χ_X is the characteristic function of X) and $\|f\|_{H^s(X)}$ is the lower bound of $\|f\|_{H^s(\mathbb{R}^n)}$ over all possible extensions. The are several (sometimes not exactly equivalent) ways to define the scale of Hilbert spaces $H^s(X)$.

Finally, it is also convenient to define H^s for negative values of s. We define $H^{-s}(\mathbb{R}^n)$ for $s \geq 0$ as the subspace of $\mathcal{S}'(\mathbb{R}^n)$, the space of *tempered distributions*, such that (1.27) holds. For bounded domains we define $H^{-s}(X)$ as the dual to $H_0^s(X)$ equipped with the norm

$$\|f\|_{H^{-s}(X)} = \sup_{g \in H_0^s(X)} \frac{\int_X f g d\mathbf{x}}{\|g\|_{H_0^s(X)}}. \tag{1.29}$$

We can verify that the two definitions agree when $X = \mathbb{R}^n$, in which case $H_0^s(\mathbb{R}^n) = H^s(\mathbb{R}^n)$.

Let us demonstrate on a simple example how the Hilbert scale can be used to understand the ill-posedness of inverse problems. Let $f(\mathbf{x})$ be a function in $L^2(\mathbb{R})$ and let u be the solution in $L^2(\mathbb{R})$ of the following ODE

$$-u'' + u = f, \qquad x \in \mathbb{R}. \tag{1.30}$$

There is a unique solution in $L^2(\mathbb{R})$ to the above equation given by

$$u(x) = \frac{1}{2}\int_{\mathbb{R}} e^{-|y-x|} f(y) dy = (g * f)(x), \qquad g(x) = \frac{1}{2}e^{-|x|},$$

as can be verified by inspection. Everything simplifies in the Fourier domain as

$$(1 + k^2)\hat{u}(k) = \hat{f}(k).$$

This implies that u is not only in $L^2(\mathbb{R})$ but also in $H^2(\mathbb{R})$ as is easily verified.

The problem is ill-posed... Let us define the operator A as follows

$$A : \begin{array}{rcl} L^2(\mathbb{R}) & \to & L^2(\mathbb{R}) \\ f & \mapsto & Af = u, \end{array} \tag{1.31}$$

where u is the solution to (1.30). As such the operator A is not invertible on $L^2(\mathbb{R})$. Indeed the inverse of the operator A is formally defined by $A^{-1}u = -u'' + u$. However for $u \in L^2(\mathbb{R})$, $-u''$ is not a function in $L^2(\mathbb{R})$ but a distribution in $H^{-2}(\mathbb{R})$. The inverse problem consisting of reconstructing $f(\mathbf{x}) \in L^2(\mathbb{R})$ from $u(\mathbf{x}) \in L^2(\mathbb{R})$ is therefore *ill-posed*. The reason is that the operator A is regularizing.

... or is it? However let us define the "same" operator

$$\tilde{A}: \begin{array}{ccc} L^2(\mathbb{R}) & \to & H^2(\mathbb{R}) \\ f & \mapsto & \tilde{A}f = u. \end{array} \tag{1.32}$$

Now $\tilde{A}$ is invertible from $H^2(\mathbb{R})$ to $L^2(\mathbb{R})$ and its inverse is indeed given by $\tilde{A}^{-1}u = -u'' + u$. So $\tilde{A}$ is *well-posed* (from $L^2(\mathbb{R})$ to $H^2(\mathbb{R})$) as can be easily verified.

So why do we bother defining ill-posed problems? If we assume that our noise (the error between measurement u_1 and measurement u_2) is small in the H^2-norm, so that $\|u_1 - u_2\|_{H^2(\mathbb{R})} \leq \delta$, then there is no problem. The reconstruction will also be accurate in the sense that $\|f_1 - f_2\|_{L^2(\mathbb{R})} \leq C\delta$, where $f_j = \tilde{A}^{-1}u_j$, $j = 1, 2$.

However in general noise is not small in the strong norm $H^2(\mathbb{R})$, but rather in the weaker norm $L^2(\mathbb{R})$. At least this is our *perception*. We have an image, collect data that come with a fair amount of "*noise*" , invert the forward operator, and are not *satisfied* with the answer. Sometimes the reconstructed image looks noisy; in some instances, the image is completely dominated by noise. We conclude that noise was not small in the appropriate norm, here $H^2(\mathbb{R})$. So we have to *come up with another model*. Let us pretend that noise is still "small", but now in the $L^2(\mathbb{R})$ sense. Let us also pretend that an image "looks good" when the error is small in the $L^2(\mathbb{R})$ sense. So now we have a noise in $L^2(\mathbb{R})$ and a metric to estimate the quality of the image which is also $L^2(\mathbb{R})$. The situation is then much more complicated from a mathematical point of view as noise can be arbitrarily amplified by what has become an unbounded operator "A^{-1}" from $L^2(\mathbb{R})$ to $L^2(\mathbb{R})$. The inverse operator $\tilde{A}^{-1}$ cannot be applied directly to the data.

A reasonable definition of mild and severe ill-posednesses. Let us conclude this section by a definition of *mildly* ill-posed problems. Throughout the session, we assume that A is **injective**, which means that one data set in the range of the operator A uniquely characterize the coefficients. So A can be inverted from its range back to $\mathfrak{X}$. It is not that the inverse of A does not exist that causes problems (although when it does not exist, we do have a difficult problem to solve). It is because A is smoothing (is a compact operator) that we need to do something.

The problem (1.25) with $\mathfrak{X} = \mathfrak{Y} = L^2(\mathbb{R}^n)$ is said to be **mildly ill-posed** provided that there exists a positive constant C and $\alpha > 0$ such that

$$\|Af\|_{H^\alpha(\mathbb{R}^n)} \geq C\|f\|_{L^2(\mathbb{R}^n)}. \tag{1.33}$$

We define $\|Af\|_{H^\alpha(\mathbb{R}^n)} = +\infty$ if f does not belong to $H^\alpha(\mathbb{R}^n)$. We say that A is *mildly ill-posed of order* α if α is the smallest real number such that (1.33) holds for some $C = C(\alpha)$. Notice that we can choose any $\alpha \geq 2$ for $\tilde{A}$ so the operator that maps f to u solution of (1.30) is a mildly ill-posed problem of order 2. The operator $\tilde{A}$ in (1.32) essentially integrates twice the function f. Any operator that corresponds to a finite number of integrations is therefore mildly ill-posed (if it is injective: the operator that maps a function to its integral on $(0, 1)$ is certainly smoothing but is not injective).

We call the inversion a **severely ill-posed** problems when no such α exists. Unfortunately, there are many practical instances of such inverse problems. We will see examples of such problems in later Lectures. A typical example is the following operator

$$Bf(x) = \mathcal{F}^{-1}_{\mathbf{k}\to\mathbf{x}}[e^{-k^2}\mathcal{F}_{\mathbf{x}\to\mathbf{k}}f](x). \tag{1.34}$$

Physically this corresponds to solving the heat equation forward in time: a very smoothing operation. We easily verify that the operator B maps $L^2(\mathbb{R})$ to $L^2(\mathbb{R})$. Hoverer it destroys high frequencies exponentially fast. This is faster than any integration would do (m integrations essentially multiply high frequencies by $|k|^{-m}$) so no $\alpha > 0$ in (1.33) exists for B. Notice that it does not mean that B is never invertible. Indeed for sufficiently smooth functions $g(x)$ in the range of B (for instance such that $\hat{g}(k)$ has compact support), we can easily define the inverse operator

$$B^{-1}g(x) = \mathcal{F}^{-1}_{\mathbf{k}\to\mathbf{x}}[e^{k^2}\mathcal{F}_{\mathbf{x}\to\mathbf{k}}f](x).$$

Physically, this corresponds to solving the heat equation backwards in time. It is clear that on a space of sufficiently smooth functions, we have $BB^{-1} = B^{-1}B = Id$. Yet, if a little bit of noise is added to the data, it will be amplified by e^{k^2} in the Fourier domain. This has devastating effects on the reconstruction. You should take your favorite version of matlab and try to add the inverse Fourier transform of $\varepsilon\hat{n}(k)e^{k^2}$ to an image $f(x)$ for $\hat{n}(k) = (1+k^2)^{-1}$, say, so that noise is square integrable. Even for ε machine precision, the image will be quickly lost in the noise even if a relatively small number of Fourier modes is used to perform the Fourier transform.

The following lectures will focus on inverse problems in medical imaging and earth science that will have a structure close to the operator $\tilde{A}$ (for the nice ones) or close to the operator B (for the nasty ones). We will come back later to techniques to solve ill-posed problems numerically.

1.7 One dimensional inverse scattering problem

To conclude this section, here is another example of a "naturally well-posed" problem, i.e., a problem that involves a forward operator that is invertible in fairly "natural" spaces. Inverse problems related to the wave equation are often of this nature. Indeed, the wave equation "transports" the information without regularizing it. Here, we consider a simple one dimensional wave equation. We observe that the inversion again involves the inverse Fourier transform.

Let us consider the one dimensional wave equation

$$\frac{1}{v^2(x)}\frac{\partial^2 U}{\partial t^2} - \frac{\partial^2 U}{\partial x^2} = \delta(t)\delta(x - x_s), \qquad t \in \mathbb{R}, \quad x \in \mathbb{R}, \tag{1.35}$$

with delta source term at time $t = 0$ and position $x = x_0$. We assume causality so that $U(x, t; x_s) = 0$ for $t < 0$ and assume that U is bounded. We measure $U(x_s, t; x_s)$ and want to infer some properties about $v(x)$. Our set-up is such that the geophone is at the same position as the source term.

We want to analyze the problem in the frequency domain. Let us define $u(x, \omega; x_s)$ the *causal Fourier transform* of $U(x, t; x_s)$ in the time variable

$$u(x, \omega; x_s) = \int_0^\infty U(x, t; x_s)e^{i\omega t}dt. \tag{1.36}$$

This transform can be inverted as follows:

$$U(x, t; x_s) = \frac{1}{2\pi}\int_{-\infty}^\infty u(x, \omega; x_s)e^{-i\omega t}d\omega. \tag{1.37}$$

The equation for $u(x,\omega;x_s)$ is the well-known Helmholtz equation

$$\frac{d^2u}{dx^2}+\frac{\omega^2}{v^2(x)}u=-\delta(x-x_s),\qquad \omega\in\mathbb{R},\quad x\in\mathbb{R}, \tag{1.38}$$

augmented with the following *radiation conditions*

$$\frac{du}{dx}\mp\frac{i\omega}{v(x)}u\ \to\ 0,\quad \text{as}\quad x\ \to\ \pm\infty. \tag{1.39}$$

Since $U(x_s,t;x_s)$ is measured, $u(x_s,\omega;x_s)$ is known by Fourier transform.

Let us make a few assumptions. We assume that $v(x)$ is known on $(-\infty,x_s)$ (in Earth profile reconstructions, one is interested in positive depths only) and that we have a good constant approximation c of $v(x)$ on (x_s,∞). We recast the latter assumption as

$$\frac{1}{v^2(x)}=\frac{1}{c^2}(1+\alpha(x)), \tag{1.40}$$

where $\alpha(x)$ is small compared to $v(x)$. In effect we *linearize* the problem of the reconstruction of $v(x)$ from the scattering measurements $u(x_s,\omega;x_s)$. Moreover our linearization is about a spatially independent velocity profile c.

The advantage is that the resulting problem is easy to invert and admits an explicit solution in the asymptotic regime of smallness of $\alpha(x)$. Let us define by u_i (i for incident) the solution of the unperturbed problem

$$\frac{d^2u_i}{dx^2}+\frac{\omega^2}{c^2}u_i=-\delta(x-x_s),\qquad \frac{du_i}{dx}\mp\frac{i\omega}{c}u_i\ \to\ 0,\quad \text{as}\quad x\ \to\ \pm\infty. \tag{1.41}$$

The solution to the above problem is nothing but the *Green function* of the Helmholtz equation with constant coefficients. It is given explicitly by

$$u_i(x,\omega;x_s)=g(x-x_s,\omega)=-\frac{ce^{i\omega|x-x_s|/c}}{2i\omega}. \tag{1.42}$$

This can be verified by inspection. Notice that the radiation conditions are also satisfied.

Let us now decompose the Helmholtz solution as the superposition of the incident field and the *scattered field*:

$$u(x,\omega;x_s)=u_i(x,\omega;x_s)+u_s(x,\omega;x_s).$$

From the equations for u and u_i, we easily verify that u_s satisfies the following equation

$$\frac{d^2u_s}{dx^2}+\frac{\omega^2}{c^2}u_s=-\frac{\omega^2}{c^2}\alpha(x)(u_i+u_s), \tag{1.43}$$

with appropriate radiation conditions. By the principle of superposition, this implies that

$$u_s(x,\omega;x_s)=\omega^2\int_{x_s}^{\infty}\frac{\alpha(y)}{c^2}(u_s+u_i)(y,\omega;x_s)g(x-y,\omega)dy. \tag{1.44}$$

We have not used so far the assumption that $\alpha(x)$ is small. This approximation is called the Born approximation and allows us to obtain from the above equation that u_s is

also small and of order α. This implies that αu_s is of order α^2, hence much smaller than the other contributions in (1.44). So neglecting u_s on the right hand side of (1.44) and replacing u_i and g by their expression in (1.42), we deduce that a good approximation of u_s is

$$u_s(x_s, \frac{ck}{2}; x_s) = -\int_{\mathbb{R}} \frac{\alpha(x)}{4} e^{ikx} dx, \qquad k \in \mathbb{R}. \tag{1.45}$$

This implies that the scattering data $u_s(x_s, \omega; x_s)$ uniquely determine the fluctuation α and that the reconstruction is *stable*: all we have to do is to take the inverse Fourier transform of u_s to obtain $\alpha(x)$. Namely, we have

$$\alpha(x) = -\frac{2}{\pi} \int_{\mathbb{R}} e^{-ikx} u_s(x_s, \frac{ck}{2}; x_s) dk. \tag{1.46}$$

Several assumptions have been made to arrive at this result. However as was the case with the MRI problem, we obtain in the end a very simple reconstruction procedure: all we have to do is to compute an inverse Fourier transform.

Chapter 2

Radon transform, X-Ray tomography, SPECT

In this chapter we consider the simplest example of integral geometry: the integration of a two-dimensional function along all possible lines in the plane (the Radon transform) and its inversion. This is the mathematical framework for one of the most successful medical imaging techniques: computed (or computerized) tomography (CT).

Later in the chapter, we consider the three dimensional Radon transform and a specific example of a weighted two dimensional Radon transform, the attenuated Radon transform and its application in Single Photon Emission Computerized Tomography (SPECT).

2.1 Transmission Tomography

In transmission tomography, objects to be imaged are probed with non-radiating sources such as X-rays in medical imaging. X-rays are composed of high energy photons that propagate in the object along straight lines unless they interact with the underlying medium and get absorbed. Let $\mathbf{x}$ denote spatial position and $\boldsymbol{\theta}$ orientation. We denote by $u(\mathbf{x}, \boldsymbol{\theta})$ the density of X-rays with position $\mathbf{x}$ and orientation $\boldsymbol{\theta}$, and by $a(\mathbf{x})$ the linear attenuation coefficient. Velocity is normalized to 1 so that locally the density $u(\mathbf{x}, \boldsymbol{\theta})$ satisfies the following equation

$$\boldsymbol{\theta} \cdot \nabla_{\mathbf{x}} u(\mathbf{x}, \boldsymbol{\theta}) + a(\mathbf{x}) u(\mathbf{x}, \boldsymbol{\theta}) = 0, \qquad \mathbf{x} \in X, \quad \boldsymbol{\theta} \in \mathbb{S}^1. \tag{2.1}$$

Here X is the physical domain (assumed to be convex) of the object and $\mathbb{S}^1$ is the unit circle. We identify any point $\boldsymbol{\theta} \in \mathbb{S}^1$ with the angle $\theta \in [0, 2\pi)$ such that $\boldsymbol{\theta} = (\cos\theta, \sin\theta)$. The advection operator is given by

$$\boldsymbol{\theta} \cdot \nabla_{\mathbf{x}} = \cos\theta \frac{\partial}{\partial x} + \sin\theta \frac{\partial}{\partial y}$$

and models free convection of X-rays and $a(\mathbf{x})u(\mathbf{x}, \boldsymbol{\theta})$ models the number of absorbed photons per unit distance at $\mathbf{x}$.

The probing source is emitted at the boundary of the domain and takes the form

$$u(\mathbf{x}, \boldsymbol{\theta}) = \delta(\mathbf{x} - \mathbf{x}_0)\delta(\boldsymbol{\theta} - \boldsymbol{\theta}_0), \tag{2.2}$$

where $\mathbf{x}_0 \in \mathbb{R}^2$, say $\mathbf{x}_0 \in \partial X$ and $\boldsymbol{\theta}_0$ is *entering* the domain X, i.e., $\boldsymbol{\theta}_0 \cdot \mathbf{n}(\mathbf{x}_0) < 0$ where $\mathbf{n}$ is the outward normal to X at $x_0 \in \partial X_0$. Above the "delta" functions mean that a unit amount of photons is created at $(\mathbf{x}_0, \boldsymbol{\theta}_0)$ in the sense that for any domain $(\mathbf{x}_0, \boldsymbol{\theta}_0) \in Y \subset \mathbb{R}^2 \times \mathbb{S}^1$, we have

$$\int_Y u(\mathbf{x}, \boldsymbol{\theta}) d\mathbf{x} d\boldsymbol{\theta} = 1.$$

The solution to (2.1)-(2.2), which is a first-order ODE in the appropriate variables, is given by

$$u(\mathbf{x} + t\boldsymbol{\theta}, \boldsymbol{\theta}) \;\; = \;\; u(\mathbf{x}, \boldsymbol{\theta}) \exp\Big(- \int_0^t a(\mathbf{x} + s\boldsymbol{\theta}) ds \Big), \quad \mathbf{x} \in \mathbb{R}^2, \quad \boldsymbol{\theta} \in \mathbb{S}^1. \tag{2.3}$$

Indeed, write $v(t) = u(\mathbf{x} + t\boldsymbol{\theta}, \boldsymbol{\theta})$ and $b(t) = a(\mathbf{x} + t\boldsymbol{\theta})$ so that $\dot{v} + av = 0$ and integrate to obtain the above result. For our specific choice at $(\mathbf{x}_0, \boldsymbol{\theta}_0)$, we thus obtain that

$$u(\mathbf{x}, \boldsymbol{\theta}) = \delta(\mathbf{x} - t\boldsymbol{\theta} - \mathbf{x}_0) \delta(\boldsymbol{\theta} - \boldsymbol{\theta}_0) \exp\Big(- \int_0^t a(\mathbf{x} - s\boldsymbol{\theta}) ds \Big).$$

In other words, on the half line $\mathbf{x}_0 + t\boldsymbol{\theta}_0$ for $t \geq 0$, there is a positive density of photons. Away from that line, the density of photons vanishes.

For $\mathbf{x}_1 = \mathbf{x}_0 + \tau\boldsymbol{\theta}_0 \in \partial X$ different from $\mathbf{x}_0$, if a detector collect the amount of photons reaching that point, it will measure

$$\exp\Big(- \int_0^\tau a(\mathbf{x}_1 - s\boldsymbol{\theta}) ds \Big) = \exp\Big(- \int_0^\tau a(\mathbf{x}_0 + s\boldsymbol{\theta}) ds \Big)$$

which is the amount of photons that reached $\mathbf{x}$ without being absorbed. The travel time (for a particle with rescaled speed equal to 1, so it is really a "distance") from one side of ∂X to the other side depends on $(\mathbf{x}_0, \boldsymbol{\theta}_0)$ and is denoted by $\tau(\mathbf{x}_0, \boldsymbol{\theta}_0) > 0$. By taking the logarithm of the measurements, we have thus access to

$$\int_0^{\tau(\mathbf{x}_0, \boldsymbol{\theta}_0)} a(\mathbf{x}_0 + t\boldsymbol{\theta}_0) dt.$$

This is the integral of a over the line segment $(\mathbf{x}_0, \mathbf{x}_1)$. By varying the point $\mathbf{x}_0$ and the direction of the incidence $\boldsymbol{\theta}_0$, we can have access to integrals of $a(\mathbf{x})$ over all possible segments (and since a can be extended by 0 outside X without changing the measurements, in fact over all possible lines) crossing the domain X.

The question in transmission tomography is thus how one can recover a function $a(\mathbf{x})$ from its integration over all possible lines in the plane $\mathbb{R}^2$. This will be the object of subsequent sections. Of course in practice one needs to consider integration over a finite number of lines. How these lines are chosen is crucial to obtain a rapid and practical inversion algorithm. We do not consider the problems of discretization in this course and refer the reader to the following excellent works in the literature [27, 28].

2.2 Two dimensional Radon transform

We have seen that the problem of transmission tomography consisted of reconstructing a function from its integration along lines. We have considered the two-dimensional problem so far. Since X-rays do not scatter, the three dimensional problem can be treated by using the two-dimensional theory: it suffices to image the object slice by slice using the two dimensional reconstruction, as we did in MRI (Transmission Tomography is indeed a tomographic method, tomos being slice in Greek).

We need to represent (parameterize) lines in the plane in a more convenient way than by describing them as the line joining $\mathbf{x}_0$ and $\mathbf{x}_1$. This is done by defining an origin $\mathbf{0} = (0,0)$, a direction $(\cos\theta, \sin\theta) = \boldsymbol{\theta} \in S^1$, and a scalar s indicating the (signed) distance of the line to $\mathbf{0}$. The line is defined by the set of points $\mathbf{x}$ such that $\mathbf{x} \cdot \boldsymbol{\theta}^\perp = s$, where $\boldsymbol{\theta}^\perp$ is the rotation of $\boldsymbol{\theta}$ by $\pi/2$, i.e., the vector given by $\boldsymbol{\theta}^\perp = (-\sin\theta, \cos\theta)$. More precisely, for a smooth function $f(\mathbf{x})$ on $\mathbb{R}^2$, we define the Radon transform $Rf(s,\theta)$ for $(s,\theta) \in Z = \mathbb{R} \times (0, 2\pi)$ as

$$Rf(s,\theta) = \int_{\mathbb{R}} f(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta})dt = \int_{\mathbb{R}^2} f(\mathbf{x})\delta(\mathbf{x} \cdot \boldsymbol{\theta}^\perp - s)d\mathbf{x}. \tag{2.4}$$

Notice that the cylinder Z is a double covering of the space of lines in the real plane

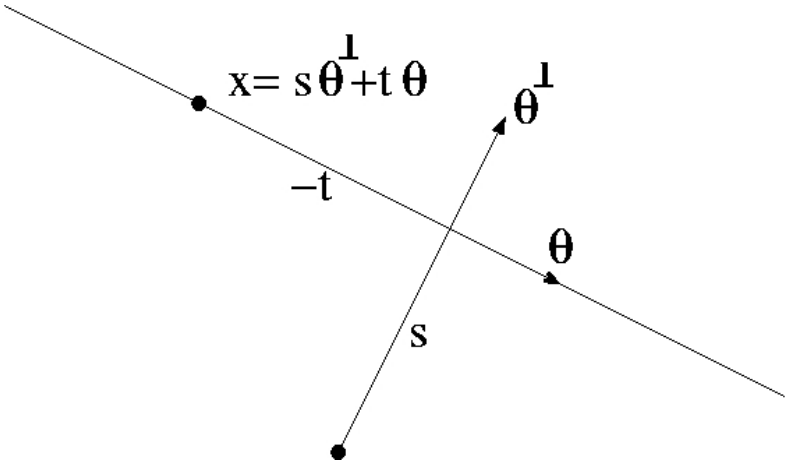

Figure 2.1: Geometry of the Radon transform.

$\mathbb{R}^2$. Indeed one easily verifies that

$$Rf(s,\theta) = Rf(-s,\theta+\pi), \qquad \text{as} \quad \{\mathbf{x} \cdot \boldsymbol{\theta}^\perp = s\} = \{\mathbf{x} \cdot (-\boldsymbol{\theta}^\perp) = (-s)\}.$$

Thus there is a redundancy of order two in the parameterization of lines in the Radon transform.

Radon transform versus X-ray transform. We pause here to make a remark regarding notation. The astute reader may have noticed that the notation introduced here for the Radon transform is not the same as the notation introduced from the Radon transform in Chapter 1! Let us explain the reason for this change of heart. A line in a two dimensional space is a hyperplane, as is a plane in a three dimensional space, a

notion that is used in the definition of the three dimensional Radon transform that we consider below. So it is natural to parameterize hyperplanes by their uniquely defined (up to a sign) normal vector.

Now, we have just seen that the reason why line integrals appear in Computerized Tomography is because particles propagate along straight lines. This is true independent of dimension (although for practical purposes in geophysical and medical imaging, all lines really are lines in the three dimensional Euclidean space). And it is clearly natural to parameterize lines by their main direction $\boldsymbol{\theta}$. This is the point of view of the **X-ray** transform (since X-rays roughly follow straight lines in medical imaging applications).

In three dimensions, hyperplanes and lines are not the same object. So the three dimensional Radon transform and the three dimensional X-ray transform are not the same object either (and there are many more lines than hyperplanes). In two dimensions, however, the X-ray and Radon transforms are the same object as they both correspond to integrals along all possible lines. In Chapter 1, we emphasized the parameterization of hyperplanes and thus modeled lines by their normal vector $\boldsymbol{\theta}$. In this chapter, we emphasize the X-ray transform, which is closer to physics, and parameterize lines by their main direction $\boldsymbol{\theta}$. One can go from one notation to the other by replacing $\boldsymbol{\theta}$ by its 90 degree rotation $\boldsymbol{\theta}^{\perp}$. But this is clearly a source of confusion.

Since the confusion profusely appears in the literature, we want to stress it here, and we shall use the same notation R for both, knowing that R in this chapter is the X-ray transform parameterization and not the Radon transform parameterization although we shall still continue calling it the Radon transform... But again, the Radon transform and the X-ray transforms are the *same* geometrical object in two dimensions: it is the integral of functions along lines. There are (at least) two possible parameterizations of lines, the Radon parameterization with $\boldsymbol{\theta}$ orthogonal to the lines and the X-ray parameterization with $\boldsymbol{\theta}$ parallel to the lines. Once we prove a property in one parameterization, we prove the property for the geometric object, which we shall call the Radon transform.

Some properties of the Radon transform. Let us derive some important properties of the Radon transform. We first define the operator

$$R_\theta f(s) = Rf(s, \theta). \tag{2.5}$$

This notation will often be useful in the sequel. The first important result on the Radon transform is the *Fourier slice theorem*:

Theorem 2.2.1 *Let $f(\mathbf{x})$ be a smooth function. Then for all $\boldsymbol{\theta} \in S^1$, we have*

$$[\mathcal{F}_{s\to\sigma} R_\theta f](\sigma) = \widehat{R_\theta f}(\sigma) = \hat{f}(\sigma\boldsymbol{\theta}^{\perp}), \qquad \sigma \in \mathbb{R}. \tag{2.6}$$

Proof. We have that

$$\widehat{R_\theta f}(\sigma) = \int_{\mathbb{R}} e^{-is\sigma} \int_{\mathbb{R}^2} f(\mathbf{x})\delta(\mathbf{x}\cdot\boldsymbol{\theta}^{\perp} - s)d\mathbf{x}ds = \int_{\mathbb{R}^2} e^{-i\mathbf{x}\cdot\boldsymbol{\theta}^{\perp}\sigma} f(\mathbf{x})d\mathbf{x}.$$

This concludes the proof. □

This result should not be surprising. For a given value of $\boldsymbol{\theta}$, the Radon transform gives us the integration of f over all lines parallel to $\boldsymbol{\theta}$. So obviously the oscillations in the

direction $\boldsymbol{\theta}$ are lost, but not the oscillations in the orthogonal direction $\boldsymbol{\theta}^\perp$. Yet the oscillations of f in the direction $\boldsymbol{\theta}^\perp$ are precisely of the form $\hat{f}(\alpha\boldsymbol{\theta}^\perp)$ for $\alpha \in \mathbb{R}$. It is therefore not surprising that the latter can be retrieved from the Radon transform $R_\theta f$.

Notice that this result also gives us a *reconstruction* procedure. Indeed, all we have to do is to take the Fourier transform of $R_\theta f$ in the variable s to get the Fourier transform $\hat{f}(\sigma\boldsymbol{\theta}^\perp)$. It remains then to obtain the latter Fourier transform for all directions $\boldsymbol{\theta}^\perp$ to end up with the full $\hat{f}(\mathbf{k})$ for all $\mathbf{k} \in \mathbb{R}^2$. Then the object is simply reconstructed by using the fact that $f(\mathbf{x}) = (\mathcal{F}^{-1}_{\mathbf{k}\to\mathbf{x}}\hat{f})(\mathbf{x})$. We will consider other (equivalent) reconstruction methods and explicit formulas later on.

Before doing so, we derive additional properties satisfied by Rf. From the Fourier slice theorem, we deduce that

$$R_\theta[\frac{\partial f}{\partial x_i}](s) = \theta_i^\perp \frac{d}{ds}(R_\theta f)(s). \tag{2.7}$$

Exercise 2.2.1 Verify (2.7).

This is the equivalent for Radon transforms of the property (1.21) for the Fourier transform.

Smoothing properties of the Radon and X-ray transforms. Let us now look at regularizing properties of the Radon transform. To do so we need to introduce a function $\chi(\mathbf{x})$ of class $C_0^\infty(\mathbb{R}^2)$ (i.e., χ is infinitely many times differentiable) and with compact support (i.e. there is a radius R such that $\chi(\mathbf{x}) = 0$ for $|\mathbf{x}| > R$. When we are interested in an object defined on a domain X, we can choose $\chi(\mathbf{x}) = 1$ for $\mathbf{x} \in X$.

As we did for $\mathbb{R}^n$ in the previous chapter, let us now define the Hilbert scale for the cylinder Z as follows. We say that $g(s,\theta)$ belongs to $H^s(Z)$ provided that

$$\|g\|^2_{H^s(Z)} = \int_0^{2\pi}\int_{\mathbb{R}} (1+\sigma^2)^s |\mathcal{F}_{s\to\sigma} g(\sigma,\theta)|^2 d\sigma d\theta < \infty. \tag{2.8}$$

This is a constraint stipulating that the Fourier transform in the s variable decays sufficiently fast at infinity. No constraint is imposed on the directional variable other than being a square-integrable function. We have then the following result:

Theorem 2.2.2 *Let $f(\mathbf{x})$ be a distribution in $H^s(\mathbb{R}^2)$ for some $s \in \mathbb{R}$. Then we have the following bounds*

$$\begin{array}{rcl} \sqrt{2}\|f\|_{H^s(\mathbb{R}^2)} & \leq & \|Rf\|_{H^{s+1/2}(Z)} \\ \|R(\chi f)\|_{H^{s+1/2}(Z)} & \leq & C_\chi \|\chi f\|_{H^s(\mathbb{R}^2)}. \end{array} \tag{2.9}$$

The constant C_χ depends on the function $\chi(\mathbf{x})$, and in particular on the size of its support.

Proof. From the Fourier slice theorem $\widehat{R_\theta w(\sigma)} = \widehat{w}(\sigma\boldsymbol{\theta}^\perp)$, we deduce that

$$\begin{array}{rcl} \displaystyle\int_Z |\widehat{Rw}(\sigma,\theta)|^2 (1+\sigma^2)^{s+1/2} d\sigma d\theta & = & \displaystyle\int_Z |\widehat{w}(\sigma\boldsymbol{\theta}^\perp)|^2 (1+\sigma^2)^{s+1/2} d\sigma d\theta \\ & = & \displaystyle 2\int_{\mathbb{R}^2} |\widehat{w}(\mathbf{k})|^2 \frac{(1+|\mathbf{k}|^2)^{s+1/2}}{|\mathbf{k}|} d\mathbf{k}, \end{array}$$

using the change of variables from polar to Cartesian coordinates so that $d\mathbf{k} = \sigma d\sigma d\theta$ and recalling that $\hat{f}(\sigma\boldsymbol{\theta}) = \hat{f}((-\sigma)(-\boldsymbol{\theta}))$. The first inequality in (2.9) then follows from the fact that $|\mathbf{k}|^{-1} \geq (1+|\mathbf{k}|^2)^{-1/2}$ using $w(\mathbf{x}) = f(\mathbf{x})$. The second inequality is slightly more difficult because of the presence of $|\mathbf{k}|^{-1}$. We now choose $w(\mathbf{x}) = f(\mathbf{x})\chi(\mathbf{x})$. Let us split the integral into $I_1 + I_2$, where I_1 accounts for the integration over $|\mathbf{k}| > 1$ and I_2 for the integration over $|\mathbf{k}| < 1$. Since $(1+|\mathbf{k}|^2)^{1/2} \leq \sqrt{2}|\mathbf{k}|$ for $|\mathbf{k}| > 1$, we have that

$$I_1 \leq 2\sqrt{2} \int_{\mathbb{R}^2} |\widehat{\chi f}(\mathbf{k})|^2 (1+|\mathbf{k}|^2)^s d\mathbf{k} \leq 2\sqrt{2} \|\chi f\|_{H^s}^2.$$

It remains to deal with I_2. We deduce from (1.22) that

$$I_2 \leq C \|\widehat{\chi f}\|_{L^\infty(\mathbb{R}^2)}^2$$

$\psi = 1$ on the support of χ so that $\psi\chi f = \chi f$ and let us define $\psi_{\mathbf{k}}(\mathbf{x}) = e^{-i\mathbf{x}\cdot\mathbf{k}}\psi(\mathbf{x})$. Upon using the definition (1.15), the Parseval relation (1.17) and the Cauchy Schwarz inequality $(f,g) \leq \|f\|\|g\|$, we deduce that

$$|\widehat{\chi f}|(\mathbf{k}) = |\widehat{\psi\chi f}|(\mathbf{k}) = \Big| \int_{\mathbb{R}^2} \psi_{\mathbf{k}}(\mathbf{x})(\chi f)(\mathbf{x}) d\mathbf{x} \Big| = (2\pi)^{-2} \Big| \int_{\mathbb{R}^2} \widehat{\psi_{\mathbf{k}}}(\boldsymbol{\xi}) \widehat{\chi f}(\boldsymbol{\xi}) d\boldsymbol{\xi} \Big|$$
$$= (2\pi)^{-2} \Big| \int_{\mathbb{R}^2} \frac{\widehat{\psi_{\mathbf{k}}}(\boldsymbol{\xi})}{(1+|\boldsymbol{\xi}|^2)^{s/2}} (1+|\boldsymbol{\xi}|^2)^{s/2} \widehat{\chi f}(\boldsymbol{\xi}) d\boldsymbol{\xi} \Big| \leq \|\psi_{\mathbf{k}}\|_{H^{-s}(\mathbb{R}^2)} \|\chi f\|_{H^s(\mathbb{R}^2)}.$$

Since $\psi(\mathbf{x})$ is smooth, then so is $\psi_{\mathbf{k}}$ uniformly in $|\mathbf{k}| < 1$ so that $\psi_{\mathbf{k}}$ belongs to $H^{-s}(\mathbb{R}^2)$ for all $s \in \mathbb{R}$ uniformly $|\mathbf{k}| < 1$. This implies that

$$I_2 \leq C \|\widehat{\chi f}\|_{L^\infty(\mathbb{R}^2)}^2 \leq C_\chi \|\chi f\|_{H^s(\mathbb{R}^2)}^2,$$

where the function depends on ψ, which depends on the support of χ. This concludes the proof. □

The theorem should be interpreted as follows. Assume that the function (or more generally a distribution) $f(\mathbf{x})$ has compact support. Then we can find a function χ which is of class C^∞, with compact support, and which is equal to 1 on the support of f. In that case, we can use the above theorem with $\chi f = f$. The constant C_χ depends then implicitly on the size of the support of $f(\mathbf{x})$.

The above inequalities show that R is a *smoothing* operator. This is not really surprising as an integration over lines is involved in the definition of the Radon transform. However, the result tells us that the Radon transform is smoothing by exactly *one half of a derivative*. The second inequality in (2.9) tells us that the factor $\frac{1}{2}$ is optimal, in the sense that the Radon transform does not regularize more than one half of a derivative. Moreover this corresponds to (1.33) with $\alpha = \frac{1}{2}$, which shows that the inversion of the Radon transform in two dimensions is a mildly ill-posed problem of order $\alpha = \frac{1}{2}$: when we reconstruct f from Rf, a differentiation of order one half is involved.

Filtered-backprojection inversions. Let us now consider such *explicit* reconstruction formulas. In order to do so, we need to introduce two new operators, the adjoint operator R^* and the Hilbert transform H. The adjoint operator R^* to R (with respect

to the usual inner products $(\cdot,\cdot)_{\mathbb{R}^2}$ and $(\cdot,\cdot)_z$ on $L^2(\mathbb{R})$ and $L^2(Z)$, respectively) is given for every smooth function $g(s,\theta)$ on Z by

$$(R^*g)(\mathbf{x}) = \int_0^{2\pi} g(\mathbf{x}\cdot\boldsymbol{\theta}^\perp,\theta)d\theta, \qquad \mathbf{x}\in\mathbb{R}^2. \tag{2.10}$$

That R^* is indeed the adjoint operator to R is verified as follows

$$\begin{aligned}
(R^*g,f)_{\mathbb{R}^2} &= \int_{\mathbb{R}^2} f(\mathbf{x})\int_0^{2\pi} g(\mathbf{x}\cdot\boldsymbol{\theta}^\perp,\theta)d\theta d\mathbf{x}\\
&= \int_{\mathbb{R}^2} f(\mathbf{x})\int_0^{2\pi}\int_{\mathbb{R}} \delta(s-\mathbf{x}\cdot\boldsymbol{\theta}^\perp)g(s,\theta)dsd\theta d\mathbf{x}\\
&= \int_0^{2\pi}\int_{\mathbb{R}} g(s,\theta)\int_{\mathbb{R}^2} f(\mathbf{x})\delta(s-\mathbf{x}\cdot\boldsymbol{\theta}^\perp)d\mathbf{x}dsd\theta\\
&= \int_0^{2\pi}\int_{\mathbb{R}} g(s,\theta)(Rf)(s,\theta)dsd\theta \ = (g,Rf)_Z.
\end{aligned}$$

We also introduce the Hilbert transform defined for smooth functions $f(t)$ on $\mathbb{R}$ by

$$Hf(t) = \frac{1}{\pi}\text{p.v.}\int_{\mathbb{R}}\frac{f(s)}{t-s}ds. \tag{2.11}$$

Here p.v. means that the integral is understood in the principal value sense, which in this context is equivalent to

$$Hf(t) = \lim_{\varepsilon\to 0}\int_{\mathbb{R}\setminus(t-\varepsilon,t+\varepsilon)}\frac{f(s)}{t-s}ds.$$

Both operators turn out to be *local* in the Fourier domain in the sense that they are multiplications in the Fourier domain. More precisely we have the following lemma.

Lemma 2.2.3 *We have in the Fourier domain the following relations:*

$$\begin{aligned}
(\mathcal{F}_{\mathbf{x}\to\boldsymbol{\xi}}R^*g)(\boldsymbol{\xi}) &= \frac{2\pi}{|\boldsymbol{\xi}|}\Big((\mathcal{F}_{s\to\sigma}g)(-|\boldsymbol{\xi}|,\hat{\boldsymbol{\xi}}^\perp)+(\mathcal{F}_{s\to\sigma}g)(|\boldsymbol{\xi}|,-\hat{\boldsymbol{\xi}}^\perp)\Big)\\
(\mathcal{F}_{t\to\sigma}Hf)(\sigma) &= -i\,\text{sign}(\sigma)\mathcal{F}_{\text{t}\to\sigma}\text{f}(\sigma).
\end{aligned} \tag{2.12}$$

We have used the notation $\hat{\boldsymbol{\xi}} = \boldsymbol{\xi}/|\boldsymbol{\xi}|$. For $\boldsymbol{\theta} = (\cos\theta,\sin\theta)$ with $\boldsymbol{\theta}\in S^1$ and $\theta\in(0,2\pi)$, we also identify the functions $g(\theta) = g(\boldsymbol{\theta})$. Assuming that $g(s,\theta) = g(-s,\theta+\pi)$, which is the case in the image of the Radon transform (i.e., when there exists f such that $g = Rf$), and which implies that $\hat{g}(\sigma,\boldsymbol{\theta}) = \hat{g}(-\sigma,-\boldsymbol{\theta})$ we have using shorter notation the equivalent statement:

$$\begin{aligned}
\widehat{R^*g}(\boldsymbol{\xi}) &= \frac{4\pi}{|\boldsymbol{\xi}|}\,\hat{g}(|\boldsymbol{\xi}|,-\hat{\boldsymbol{\xi}}^\perp)\\
\widehat{Hf}(\sigma) &= -i\,\text{sign}(\sigma)\hat{\text{f}}(\sigma).
\end{aligned} \tag{2.13}$$

Proof. Let us begin with R^*g. We compute

$$\begin{aligned}
\widehat{R^*g}(\boldsymbol{\xi}) &= \int e^{-i\mathbf{x}\cdot\boldsymbol{\xi}}g(\mathbf{x}\cdot\boldsymbol{\theta}^\perp,\theta)d\theta d\mathbf{x} = \int e^{-is\boldsymbol{\xi}\cdot\boldsymbol{\theta}^\perp}g(s,\theta)dsd\theta e^{-it\boldsymbol{\xi}\cdot\boldsymbol{\theta}}dt\\
&= \int 2\pi\delta(|\boldsymbol{\xi}|\hat{\boldsymbol{\xi}}\cdot\boldsymbol{\theta})\hat{g}(\boldsymbol{\xi}\cdot\boldsymbol{\theta}^\perp,\theta)d\theta = \int\frac{2\pi}{|\boldsymbol{\xi}|}\delta(\hat{\boldsymbol{\xi}}\cdot\boldsymbol{\theta})\hat{g}(\boldsymbol{\xi}\cdot\boldsymbol{\theta}^\perp,\theta)d\theta\\
&= \frac{2\pi}{|\boldsymbol{\xi}|}\Big(\hat{g}(-|\boldsymbol{\xi}|,\hat{\boldsymbol{\xi}}^\perp)+\hat{g}(|\boldsymbol{\xi}|,-\hat{\boldsymbol{\xi}}^\perp)\Big).
\end{aligned}$$

In the proof we have used that $\delta(\alpha x) = \alpha^{-1}\delta(x)$ and the fact that there are two directions, namely $\hat{\boldsymbol{\xi}}$ and $-\hat{\boldsymbol{\xi}}$ on the unit circle, which are orthogonal to $\hat{\boldsymbol{\xi}}^{\perp}$. When g is in the form $g = Rf$, we have $\hat{g}(-|\boldsymbol{\xi}|, \hat{\boldsymbol{\xi}}^{\perp}) = \hat{g}(|\boldsymbol{\xi}|, -\hat{\boldsymbol{\xi}}^{\perp})$, which explains the shorter notation (2.13).

The computation of the second operator goes as follows. We verify that

$$Hf(t) = \frac{1}{\pi}\Big(\frac{1}{x} * f(x)\Big)(t).$$

So in the Fourier domain we have

$$\widehat{Hf}(\sigma) = \frac{1}{\pi}\widehat{\frac{1}{x}}(\sigma)\hat{f}(\sigma) = -i\text{sign}(\sigma)\hat{f}(\sigma).$$

The latter is a result of the following calculation

$$\frac{1}{2}\text{sign}(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \frac{e^{ix\xi}}{i\xi} d\xi.$$

This concludes the proof of the lemma. □

The above calculations also show that $H^2 = H \circ H = -Id$, where Id is the identity operator, as can easily be seen in from its expression in the Fourier domain. This property is referred to as saying that the Hilbert transform is an *anti-involution*. We are now ready to introduce some reconstruction formulas.

Theorem 2.2.4 *Let $f(\mathbf{x})$ be a smooth function and let $g(s,\theta) = Rf(s,\theta)$ be its Radon transform. Then, f can explicitly be reconstructed from its Radon transform as follows:*

$$f(\mathbf{x}) = \frac{1}{4\pi} R^*\Big(\frac{\partial}{\partial s} Hg(s,\theta)\Big)(\mathbf{x}). \tag{2.14}$$

In the above formula, the Hilbert transform H acts on the s variable.

Proof. The simplest way to verify the inversion formula is to do it in the Fourier domain. Let us denote by

$$w(s,\theta) = \frac{\partial}{\partial s} Hg(s,\theta).$$

Since $g(-s, \theta+\pi) = g(s,\theta)$, we verify that the same property holds for w in the sense that $w(-s,\theta+\pi) = w(s,\theta)$. Therefore (2.14) is equivalent to the statement:

$$\hat{f}(\boldsymbol{\xi}) = \frac{1}{|\boldsymbol{\xi}|}\hat{w}(|\boldsymbol{\xi}|, -\hat{\boldsymbol{\xi}}^{\perp}), \tag{2.15}$$

according to (2.13). Notice that $\hat{w}$ is the Fourier transform of w in the *first variable* only.

Since in the Fourier domain, the derivation with respect to s is given by multiplication by $i\sigma$ and the Hilbert transform H is given by multiplication by $-i\text{sign}(\sigma)$, we obtain that

$$\mathcal{F}^{-1}_{\sigma\to s}\frac{\partial}{\partial s} H \mathcal{F}_{s\to\sigma} = |\sigma|.$$

In other words, we have

$$\hat{w}(\sigma,\theta) = |\sigma|\hat{g}(\sigma,\theta).$$

Thus (2.15) is equivalent to

$$\hat{f}(\boldsymbol{\xi}) = \hat{g}(|\boldsymbol{\xi}|, -\hat{\boldsymbol{\xi}}^{\perp}).$$

This, however, is nothing but the Fourier slice theorem stated in Theorem 2.2.1 since $(-\hat{\boldsymbol{\xi}}^{\perp})^{\perp} = \hat{\boldsymbol{\xi}}$ and $\boldsymbol{\xi} = |\boldsymbol{\xi}|\hat{\boldsymbol{\xi}}$. This concludes the proof of the reconstruction. □
The theorem can equivalently be stated as

$$Id = \frac{1}{4\pi} R^* \frac{\partial}{\partial s} H R = \frac{1}{4\pi} R^* H \frac{\partial}{\partial s} R. \tag{2.16}$$

The latter equality comes from the fact that H and ∂_s commute as can easily be observed in the Fourier domain (where they are both multiplications). Here, Id is the identity operator, which maps a function $f(\mathbf{x})$ to itself $Id(f) = f$.

Here is some additional useful notation in the manipulation of the Radon transform. Recall that $R_\theta f(s)$ is defined as in (2.5) by

$$R_\theta f(s) = Rf(s,\theta).$$

The adjoint R_θ (with respect to the inner products in $L^2_s(\mathbb{R})$ and $L^2_{\mathbf{x}}(\mathbb{R}^2)$) is given by

$$(R^*_\theta g)(\mathbf{x}) = g(\mathbf{x}\cdot\boldsymbol{\theta}^{\perp}). \tag{2.17}$$

Indeed (since $\boldsymbol{\theta}$ is *frozen* here) we have

$$\int_{\mathbb{R}} (R_\theta f)(s) g(s) ds = \int_{\mathbb{R}^2}\int_{\mathbb{R}} f(\mathbf{x})\delta(s - \mathbf{x}\cdot\boldsymbol{\theta}^{\perp}) g(s) d\mathbf{x} ds = \int_{\mathbb{R}^2} g(\mathbf{x}\cdot\boldsymbol{\theta}^{\perp}) f(\mathbf{x}) d\mathbf{x},$$

showing that $(R_\theta f, g)_{L^2(\mathbb{R})} = (f, R^*_\theta g)_{L^2(\mathbb{R}^2)}$. We can then recast the inversion formula as

$$Id = \frac{1}{4\pi}\int_0^{2\pi} \boldsymbol{\theta}^{\perp}\cdot\nabla R^*_\theta H R_\theta d\theta. \tag{2.18}$$

The only thing to prove here compared to previous formulas is that R^*_θ and the derivation commute, i.e., for any function $g(s)$ for $s \in \mathbb{R}$, we have

$$\boldsymbol{\theta}^{\perp}\cdot\nabla(R^*_\theta g)(\mathbf{x}) = (R^*_\theta \frac{\partial}{\partial s} g)(\mathbf{x}).$$

This results from the fact that both terms are equal to $g'(\mathbf{x}\cdot\boldsymbol{\theta}^{\perp})$.

One remark on the smoothing properties of the Radon transform. We have seen that the Radon transform is a smoothing operator in the sense that the Radon transform is half of a derivative smoother than the original function. The adjoint operator R^* enjoys exactly the same property: it regularizes by half of a derivative. It is not surprising that these two half derivatives are exactly canceled by the appearance of a full derivation in the reconstruction formula. Notice that the Hilbert transform (which corresponds to multiplication by the smooth function $i\text{sign}(\sigma)$ in the Fourier domain) is a bounded operator with bounded inverse in $L^2(\mathbb{R})$ (since $H^{-1} = -H$).

Exercise 2.2.2 (easy). Show that

$$f(\mathbf{x}) = \frac{1}{4\pi} \int_0^{2\pi} (Hg')(\mathbf{x} \cdot \boldsymbol{\theta}^\perp, \theta) d\theta.$$

Here g' means first derivative of g with respect to the s variable only.

Exercise 2.2.3 (easy). Show that

$$f(\mathbf{x}) = \frac{1}{4\pi^2} \int_0^{2\pi} \int_{\mathbb{R}} \frac{\frac{d}{ds} g(s, \theta)}{\mathbf{x} \cdot \boldsymbol{\theta}^\perp - s} ds d\theta.$$

This is Radon's original inversion formula [30].

Exercise 2.2.4 (moderately difficult). Starting from the definition

$$f(\mathbf{x}) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{i\mathbf{k}\cdot\mathbf{x}} \hat{f}(\mathbf{k}) d\mathbf{k},$$

and writing it in polar coordinates (with change of measure $d\mathbf{k} = |\mathbf{k}| d|\mathbf{k}| d\hat{\mathbf{k}}$), deduce the above reconstruction formulas by using the Fourier slice theorem.

2.3 Three dimensional Radon transform

Let us briefly mention the case of the Radon transform in three dimensions (generalizations to higher dimensions being also possible). The Radon transform in three dimensions consists of integrating a function $f(\mathbf{x})$ over all possible planes. As we mentioned earlier, the Radon transform is therefore a distinct object from the X-ray transform, which integrates a function along all possible lines.

A plane $\mathcal{P}(s, \boldsymbol{\theta})$ in $\mathbb{R}^3$ is characterized by its direction $\boldsymbol{\theta} \in S^2$, where S^2 is the unit sphere, and by its signed distance to the origin s. Notice again the double covering in the sense that $\mathcal{P}(s, \boldsymbol{\theta}) = \mathcal{P}(-s, -\boldsymbol{\theta})$. The Radon transform is then defined as

$$Rf(s, \boldsymbol{\theta}) = \int_{\mathbb{R}^3} f(\mathbf{x}) \delta(\mathbf{x} \cdot \boldsymbol{\theta} - s) d\mathbf{x} = \int_{\mathcal{P}(s,\theta)} f d\sigma. \tag{2.19}$$

Notice the change of notation compared to the two-dimensional case. The Fourier slice theorem still holds

$$\widehat{Rf}(\sigma, \boldsymbol{\theta}) = \hat{f}(\sigma \boldsymbol{\theta}), \tag{2.20}$$

as can be easily verified. We check that $Rf(s, \boldsymbol{\theta}) = Rf(-s, -\boldsymbol{\theta})$. The reconstruction formula is then given by

$$f(\mathbf{x}) = \frac{-1}{8\pi^2} \int_{S^2} g''(\mathbf{x} \cdot \boldsymbol{\theta}, \boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{2.21}$$

Here $d\boldsymbol{\theta}$ is the usual (Lebesgue) surface measure on the unit sphere.

The result can be obtained as follows. We denote by $S^2/2$ half of the unit sphere (for instance the vectors $\boldsymbol{\theta}$ such that $\boldsymbol{\theta} \cdot \mathbf{e}_z > 0$).

$$f(\mathbf{x}) = \frac{1}{(2\pi)^3} \int_{\frac{S^2}{2}} \int_{\mathbb{R}} \hat{f}(r\boldsymbol{\theta}) e^{ir\boldsymbol{\theta}\cdot\mathbf{x}} |r|^2 dr d\boldsymbol{\theta} = \frac{1}{(2\pi)^3} \int_{\frac{S^2}{2}} \int_{\mathbb{R}} \widehat{Rf}(r, \boldsymbol{\theta}) e^{ir\boldsymbol{\theta}\cdot\mathbf{x}} |r|^2 dr d\boldsymbol{\theta}$$
$$= \frac{1}{(2\pi)^2} \int_{\frac{S^2}{2}} (-g)''(\boldsymbol{\theta}\cdot\mathbf{x}, \boldsymbol{\theta}) d\boldsymbol{\theta} = \frac{1}{2} \frac{-1}{(2\pi)^2} \int_{S^2} g''(\boldsymbol{\theta}\cdot\mathbf{x}, \boldsymbol{\theta}) d\boldsymbol{\theta}.$$

Here we have used the fact that the inverse Fourier transform of $r^2\hat{f}$ is $-f''$.

Exercise 2.3.1 Generalize Theorem 2.2.2 and prove the following result:

Theorem 2.3.1 *There exists a constant C_χ independent of $f(\mathbf{x})$ such that*

$$\begin{aligned} \sqrt{2}\|f\|_{H^s(\mathbb{R}^3)} &\leq \|Rf\|_{H^{s+1}(Z)} \\ \|R(\chi f)\|_{H^{s+1}(Z)} &\leq C_\chi \|\chi f\|_{H^s(\mathbb{R}^3)}, \end{aligned} \tag{2.22}$$

where $Z = \mathbb{R} \times S^2$ and $H^s(Z)$ is defined in the spirit of (2.8).

The above result shows that the Radon transform is more smoothing in three dimensions than it is in two dimensions. In three dimensions, the Radon transform smoothes by a full derivative rather than a half derivative.

Notice however that the inversion of the three dimensional Radon transform (2.21) is *local*, whereas this is not the case in two dimensions. What is meant by local is the following: the reconstruction of $f(\mathbf{x})$ depends on $g(s, \boldsymbol{\theta})$ only for the planes $\mathcal{P}(s, \boldsymbol{\theta})$ that pass through $\mathbf{x}$ (and an infinitely small neighborhood so that the second derivative can be calculated). Indeed, we verify that $\mathbf{x} \in \mathcal{P}(\mathbf{x}\cdot\boldsymbol{\theta}, \boldsymbol{\theta})$ and that all the planes passing by $\mathbf{x}$ are of the form $\mathcal{P}(\mathbf{x}\cdot\boldsymbol{\theta}, \boldsymbol{\theta})$. The two dimensional transform involves the Hilbert transform, which unlike differentiations, is a *non-local* operation. Thus the reconstruction of f at a point $\mathbf{x}$ requires knowledge of *all* line integrals $g(s, \theta)$, and not only for those lines passing through $\mathbf{x}$.

Exercise 2.3.2 Calculate R^*, the adjoint operator to R (with respect to the usual L^2 inner products). Generalize the formula (2.16) to the three dimensional case.

2.4 Attenuated Radon Transform

In the previous sections, the integration over lines for the Radon transform was not weighted. We could more generally ask whether integrals of the form

$$\int_{\mathbb{R}} f(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}) \alpha(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}, \boldsymbol{\theta}) d\boldsymbol{\theta},$$

over all possible lines parameterized by $s \in \mathbb{R}$ and $\boldsymbol{\theta} \in S^1$ and assuming that the weight $\alpha(\mathbf{x}, \boldsymbol{\theta})$ is known, uniquely determine $f(\mathbf{x})$. This is a much more delicate question for which only partial answers are known. We concentrate here on one example of weighted Radon transform, namely the attenuated Radon transform, for which an inversion formula was obtained only fairly recently.

2.4.1 Single Photon Emission Computed Tomography

An important application for the attenuated Radon transform is SPECT, single photon emission computed tomography. The principle is the following: radioactive particles are injected in a domain. These particles emit then some radiation. The radiation propagates through the tissues and gets partially absorbed. The amount of radiation reaching the boundary of the domain can be measured. The imaging technique consists then of reconstructing the location of the radioactive particles from the boundary measurements.

We model the density of radiated photons by $u(\mathbf{x}, \theta)$ and the source of radiated photons by $f(\mathbf{x})$. The absorption of photons (by the human tissues in the medical imaging application) is modeled by $a(\mathbf{x})$. We assume that $a(\mathbf{x})$ is *known* here. The absorption can be obtained, for instance, by transmission tomography as we saw in earlier sections. The density $u(\mathbf{x}, \theta)$ satisfies then the following transport equation

$$\boldsymbol{\theta} \cdot \nabla u(\mathbf{x}, \theta) + a(\mathbf{x}) u(\mathbf{x}, \theta) = f(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}^2, \boldsymbol{\theta} \in S^1. \tag{2.23}$$

We assume that $f(\mathbf{x})$ is compactly supported and impose that no radiation comes from infinity:

$$\lim_{s \to \infty} u(\mathbf{x} - s\boldsymbol{\theta}, \theta) = 0. \tag{2.24}$$

The transport equation (2.23) with conditions (2.24) admits a unique solution that can be obtained by the method of characteristics. Let us define the following *symmetrized* beam transform

$$D_\theta a(\mathbf{x}) = \frac{1}{2} \int_0^\infty [a(\mathbf{x} - t\boldsymbol{\theta}) - a(\mathbf{x} + t\boldsymbol{\theta})] dt = \frac{1}{2} \int_{\mathbb{R}} \operatorname{sign}(t) a(\mathbf{x} - t\boldsymbol{\theta}) dt. \tag{2.25}$$

We verify that $\boldsymbol{\theta} \cdot \nabla D_\theta a(\mathbf{x}) = a(\mathbf{x})$ so that $e^{D_\theta a(\mathbf{x})}$ is an integrating factor for (2.23) in the sense that

$$\boldsymbol{\theta} \cdot \nabla (e^{D_\theta a(\mathbf{x})} u(\mathbf{x}, \theta)) = (e^{D_\theta a} f)(\mathbf{x}, \theta).$$

Therefore the solution $u(\mathbf{x}, \theta)$ is given by

$$e^{D_\theta a(\mathbf{x})} u(\mathbf{x}, \theta) = \int_0^\infty (e^{D_\theta a} f)(\mathbf{x} - t\boldsymbol{\theta}, \theta) dt. \tag{2.26}$$

We recall that $\boldsymbol{\theta} = (\cos\theta, \sin\theta)$ and that $\boldsymbol{\theta}^\perp = (-\sin\theta, \cos\theta)$ and decompose $\mathbf{x} = s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}$. We deduce from (2.26) that

$$\lim_{t \to +\infty} e^{D_\theta a(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta})} u(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}, \theta) = \int_{\mathbb{R}} (e^{D_\theta a} f)(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}, \theta) dt.$$

In the above expression the left hand side is know from the measurements. Indeed $u(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}, \theta)$ is the radiation outside of the domain to image and is thus measured and $e^{D_\theta a(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta})}$ involves the attenuation coefficient $a(\mathbf{x})$ which we have assumed is known. The objective is thus to reconstruct $f(\mathbf{x})$ from the right hand side of the above relation, which we recast as

$$(R_a f)(s, \theta) = (R_{a,\theta} f)(s) = (R_\theta (e^{D_\theta a} f))(s), \tag{2.27}$$

where R_θ is the Radon transform defined for a function of $f(\mathbf{x}, \theta)$ as

$$R_\theta f(s) = \int_{\mathbb{R}} f(s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}, \theta)dt = \int_{\mathbb{R}^2} f(\mathbf{x}, \theta)\delta(\mathbf{x} \cdot \boldsymbol{\theta}^\perp - s)d\mathbf{x}.$$

When $a \equiv 0$, we recover that the measurements involve the Radon transform of $f(\mathbf{x})$ as defined in (2.4). Thus in the absence of absorption, SPECT can be handled by inverting the Radon transform as we saw in earlier sections. When absorption is constant, an inversion formula has been known for quite some time [35]. The inversion formula for non-constant absorption is more recent and was obtained independently by two different techniques [6, 29]. We do not consider here the method of $A-$analytic functions developed in [6]. We will present the method developed in [29] based on the extension of the transport equation in the complex domain and on the solution of a Riemann Hilbert problem.

2.4.2 Riemann Hilbert problem

Riemann Hilbert problems find many applications in complex analysis. We consider here the simplest of Riemann Hilbert problems and refer the reader to [1] for much more general cases and applications.

Let T be a smooth closed curve in the complex plane, which in our application will be the unit circle, i.e., the complex numbers λ such that $|\lambda| = 1$. The reason why we choose the notation λ to represent complex numbers will appear more clearly in the next section. We denote by D^+ the open bounded domain inside the curve T, i.e., in our application the unit disk $\{\lambda \in \mathbb{C}, |\lambda| < 1\}$, and by D^- the open unbounded domain outside of the curve T, i.e., in our application $\{\lambda \in \mathbb{C}, |\lambda| > 1\}$. The orientation of the curve T is chosen so that D^+ is on the "left" of the curve T.

For a smooth function $\phi(\lambda)$ defined on $D^+ \cup D^-$, we denote by $\phi^+(t)$ and $\phi^-(t)$ the traces of ϕ on T from D^+ and D^-, respectively. So in the case where T is the unit circle, we have

$$\phi^+(t) = \lim_{0<\varepsilon\to 0} \phi((1-\varepsilon)t), \qquad \phi^-(t) = \lim_{0<\varepsilon\to 0} \phi((1+\varepsilon)t).$$

We define $\varphi(t)$ on T as the jump of ϕ, i.e.,

$$\varphi(t) = \phi^+(t) - \phi^-(t). \tag{2.28}$$

Let $\varphi(t)$ be a smooth function defined on T. The Riemann Hilbert problem is stated as follows. Find a function $\phi(\lambda)$ on $D^+ \cup D^-$ such that

1. $\phi(\lambda)$ is analytic on D^+ and analytic on D^-
2. $\lambda\phi(\lambda)$ is bounded as $|\lambda| \to \infty$ on D^-
3. the jump of ϕ is given by $\varphi(t) = \phi^+(t) - \phi^-(t)$.

The solution to the above Riemann Hilbert problem is *unique* and is given by the Cauchy formula

$$\phi(\lambda) = \frac{1}{2\pi i}\int_T \frac{\varphi(t)}{t-\lambda}dt, \qquad \lambda \in \mathbb{C}\backslash T = D^+ \cup D^-. \tag{2.29}$$

This is the form of the Riemann Hilbert problem we will use in the sequel. We refer the reader to [1] for the theory.

2.4.3 Inversion of the Attenuated Radon Transform

We now want to apply the theory of Riemann Hilbert problems to invert the attenuated Radon transform (AtRT). The first step is to extend the transport equation to the complex domain as follows. We parameterize the unit circle in the complex plane as

$$\lambda = e^{i\theta}, \qquad \theta \in (0, 2\pi). \tag{2.30}$$

The new parameter takes values on the unit circle T for $\theta \in (0, 2\pi)$. It can also be seen more generally as an arbitrary complex number $\lambda \in \mathbb{C}$. With the notation $\mathbf{x} = (x, y)$, the transport equation (2.23) can be recast as

$$\Big(\frac{\lambda + \lambda^{-1}}{2}\frac{\partial}{\partial x} + \frac{\lambda - \lambda^{-1}}{2i}\frac{\partial}{\partial y} + a(\mathbf{x})\Big)u(\mathbf{x}, \lambda) = f(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}^2, \quad \lambda \in T. \tag{2.31}$$

We can simplify the above equation by identifying $\mathbf{x}$ with $z = x + iy$ and by defining

$$\frac{\partial}{\partial z} = \frac{1}{2}\Big(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\Big), \qquad \frac{\partial}{\partial \overline{z}} = \frac{1}{2}\Big(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\Big). \tag{2.32}$$

The transport equation (2.31) is then equivalent to

$$\Big(\lambda\frac{\partial}{\partial z} + \lambda^{-1}\frac{\partial}{\partial \overline{z}} + a(z)\Big)u(z, \lambda) = f(z), \qquad z \in \mathbb{C}, \quad \lambda \in T. \tag{2.33}$$

The same boundary conditions (2.24) that no information comes from infinity need to be added in the new variables as well.

The above equation can also be generalized to $\lambda \in \mathbb{C}$ instead of T. It is in this framework that the Riemann Hilbert problem theory is used to invert the attenuated Radon transform. This will be done in three steps

(i) We show that $u(z, \lambda)$ is *analytic* in $D^+ \cup D^- = \mathbb{C}\backslash T$ and that $\lambda u(z, \lambda)$ is bounded as $\lambda \to \infty$.

(ii) We verify that $\varphi(\mathbf{x}, \theta) = u^+(\mathbf{x}, \theta) - u^-(\mathbf{x}, \theta)$, the jump of u at $\lambda = e^{i\theta}$ can be written as a function of the measured data $R_a f(s, \theta)$.

(iii) We solve the Riemann Hilbert problem using (2.29) and evaluate (2.33) at $\lambda = 0$ to obtain a reconstruction formula for $f(z) = f(\mathbf{x})$.

2.4.4 Step (i): The $\overline{\partial}$ problem, an elliptic equation

Let us now analyze (2.33). In the absence of absorption the fundamental solution of (2.33) solves the following equation

$$\Big(\lambda\frac{\partial}{\partial z} + \lambda^{-1}\frac{\partial}{\partial \overline{z}}\Big)G(z, \lambda) = \delta(z), \qquad |G(z, \lambda)| \to 0 \text{ as } |z| \to \infty, \tag{2.34}$$

for $\lambda \in \mathbb{C}\backslash(T \cup \{0\})$.

Lemma 2.4.1 *The unique solution to (2.34) is given by*

$$G(z, \lambda) = \frac{\operatorname{sign}(|\lambda| - 1)}{\pi(\lambda\overline{z} - \lambda^{-1}z)}, \qquad \lambda \notin (T \cup \{0\}). \tag{2.35}$$

Proof. The formula can be verified by inspection. A more constructive derivation is the following. Let us define the change of variables

$$\zeta = \lambda^{-1} z - \lambda \bar{z}, \qquad \bar{\zeta} = \overline{\lambda^{-1}} \bar{z} - \bar{\lambda} z. \tag{2.36}$$

Let us assume that $|\lambda| > 1$. The Jacobian of the transformation is $|\lambda|^2 - |\lambda|^{-2}$. We verify that

$$\lambda \frac{\partial}{\partial z} + \lambda^{-1} \frac{\partial}{\partial \bar{z}} = (|\lambda|^{-2} - |\lambda|^2) \frac{\partial}{\partial \bar{\zeta}}.$$

The change of variables (2.36) has been precisely tailored so that the above holds. Denoting $\tilde{G}(\zeta) = G(z)$, we thus obtain

$$\frac{\partial}{\partial \bar{\zeta}} \tilde{G}(\zeta) = \frac{1}{|\lambda|^{-2} - |\lambda|^2} \delta(z(\zeta)) = -\delta(\zeta).$$

So $-\tilde{G}(\zeta)$ is the fundamental solution of the $\overline{\partial}$ operator $\dfrac{\partial}{\partial \bar{\zeta}}$. We verify that

$$\frac{\partial}{\partial \bar{\zeta}} \frac{1}{\zeta} = \pi \delta(\zeta). \tag{2.37}$$

Indeed let $\psi(z)$ be a smooth test function in $C_0^\infty(\mathbb{R}^2)$ and let $d\mu(\zeta)$ be the Lebesgue measure $dxdy$ in $\mathbb{C} \sim \mathbb{R}^2$. Then

$$\begin{aligned} &\int_{\mathbb{C}} \psi(\zeta) \frac{\partial}{\partial \bar{\zeta}} \frac{1}{\zeta} d\mu(\zeta) = -\int_{\mathbb{C}} \frac{\partial \psi}{\partial \bar{\zeta}} \frac{1}{\zeta} d\mu(\zeta) = -\lim_{\varepsilon \to 0} \int_{\mathbb{C} \setminus \{|\zeta| < \varepsilon\}} \frac{\partial \psi}{\partial \bar{\zeta}} \frac{1}{\zeta} d\mu(\zeta) \\ &= -\lim_{\varepsilon \to 0} \int_{\mathbb{C} \setminus \{|\zeta| < \varepsilon\}} \frac{\partial \zeta^{-1} \psi}{\partial \bar{\zeta}} d\mu(\zeta) = \frac{1}{2i} \int_{|\zeta| = \varepsilon} \frac{\psi}{\zeta} d\mu(\zeta), \end{aligned}$$

by the Green formula with complex variables:

$$\int_{\partial X} u dz = \int_{\partial X} (u dx + i u dy) = \int_X (i \frac{\partial u}{\partial x} - \frac{\partial u}{\partial y}) dx dy = 2i \int_X \frac{\partial u}{\partial \bar{z}} d\mu(z).$$

Sending ε to 0, we find in the limit

$$\int_{\mathbb{C}} \psi(\zeta) \frac{\partial}{\partial \bar{\zeta}} \frac{1}{\zeta} d\mu(\zeta) = \frac{1}{2i} 2\pi i \psi(0) = \pi \int_{\mathbb{R}^2} \psi(\zeta) \delta(\zeta) d\mu(\zeta).$$

This implies that $\tilde{G}(\zeta) = (-\pi\zeta)^{-1}$, hence $G(z) = (-\pi\zeta)^{-1} = (\pi(\lambda \bar{z} - \lambda^{-1} z))^{-1}$ This is (2.35) for $|\lambda| > 1$. Now for $|\lambda| < 1$, we verify that the Jacobian of the transformation $z \to \zeta(z)$ now becomes $|\lambda|^{-2} - |\lambda|^2$ so that

$$\frac{\partial}{\partial \bar{\zeta}} \tilde{G}(\zeta) = \frac{1}{|\lambda|^{-2} - |\lambda|^2} \delta(z(\zeta)) = \delta(\zeta).$$

This yields (2.35) for $|\lambda| < 1$. □

The above proof shows that $(\pi z)^{-1}$ is the fundamental solution of the $\bar{\delta}$ operator. This implies that the solution to the following $\bar{\delta}$ problem

$$\frac{\partial}{\partial \bar{z}} f(z) = g(z), \qquad z \in \mathbb{C}, \tag{2.38}$$

such that $f(z)$ vanishes at infinity is given by convolution by the fundamental solution, i.e.,

$$f(z) = \frac{1}{\pi}\int_{\mathbb{C}} \frac{g(\zeta)}{z-\zeta} d\mu(\zeta) = \frac{1}{2\pi i}\int_{\mathbb{C}} \frac{g(\zeta)}{z-\zeta} d\zeta \wedge d\bar{\zeta}. \tag{2.39}$$

Here we have used that $dz \wedge d\bar{z} = (dx + idy) \wedge (dx - idy) = 2idx \wedge dy = 2id\mu(z)$, whence the change of 2-form in the above integrations.

Notice that the Green function $G(z, \lambda)$ tends to 0 as $z \to \infty$ for $\lambda \notin T$. This is clearly not true when $\lambda \in T$, where $G(z, \lambda) = \delta(l_\theta(z))$, where $l_\theta(z)$ is the segment $\{t\boldsymbol{\theta}, t > 0\}$. The reason is that for $\lambda \notin (T \cup \{0\})$,

$$\lambda \frac{\partial}{\partial z} + \lambda^{-1}\frac{\partial}{\partial \bar{z}} \qquad \text{and} \qquad \frac{\partial}{\partial \bar{z}},$$

are elliptic operators, in the sense that in the Fourier domain, their symbol given by $\lambda k_z + \lambda^{-1}\overline{k_z}$ and $\overline{k_z}$, respectively, are positive provided that k_z is not 0. Indeed we verify that

$$\lambda k_z + \lambda^{-1}\overline{k_z} = 0 \quad \text{implies} \quad |\lambda|^2 = 1,$$

when $k_z \neq 0$ since $|k_z| = |\overline{k_z}| \neq 0$.

Let us now define $h(z, \lambda)$ as the solution to

$$\Big(\lambda \frac{\partial}{\partial z} + \lambda^{-1}\frac{\partial}{\partial \bar{z}}\Big) h(z, \lambda) = a(z), \qquad |h(z, \lambda)| \to 0 \text{ as } |z| \to \infty, \tag{2.40}$$

for $\lambda \notin (T \cup \{0\})$. The solution is given by

$$h(z, \lambda) = \int_{\mathbb{R}^2} G(z - \zeta, \lambda) a(\zeta) d\mu(\zeta). \tag{2.41}$$

We now verify that

$$\Big(\lambda \frac{\partial}{\partial z} + \lambda^{-1}\frac{\partial}{\partial \bar{z}}\Big)(e^{h(z,\lambda)} u(z, \lambda)) = e^{h(z,\lambda)} f(z),$$

so that for $\lambda \notin (T \cup \{0\})$, the solution of (2.33) is given by

$$u(z, \lambda) = e^{-h(z,\lambda)} \int_{\mathbb{R}^2} G(z - \zeta, \lambda) e^{h(\zeta,\lambda)} f(\zeta) d\mu(\zeta). \tag{2.42}$$

We verify that $G(z, \lambda)$, $h(z, \lambda)$ and $u(z, \lambda)$ are defined by continuity at $z = 0$ since $G(z, \lambda) = 0$ by continuity. We now verify that $G(z, \lambda)$ is analytic in D^+ (including at $z = 0$) and in D^-. Assuming that $a(z)$ and $f(z)$ are smooth functions, this is also therefore the case for $h(z, \lambda)$ and $u(z, \lambda)$. Moreover we easily deduce from (2.42) that $\lambda u(z, \lambda)$ is bounded on D^-. The solution $u(z, \lambda)$ of the transport equation extended to the complex plane is therefore a good candidate to apply the Riemann Hilbert theorem.

2.4.5 Step (ii): jump conditions

We now want to find the limit of $u(z, \lambda)$ as λ approaches T from above (in D^-) and below (in D^+). Let us write $\lambda = re^{i\theta}$ and let us send $r - 1$ to ∓ 0 on $D^\pm$. The Green function behaves according to the following result

Lemma 2.4.2 *As $r-1 \to \mp 0$, the Green function $G(\mathbf{x},\lambda)$ tends to*

$$G_\pm(\mathbf{x},\theta) = \frac{\pm 1}{2\pi i(\boldsymbol{\theta}^\perp \cdot \mathbf{x} \mp i0\,\mathrm{sign}(\boldsymbol{\theta}\cdot\mathbf{x}))}. \tag{2.43}$$

Proof. Let us assume that $|\lambda|>1$, i.e., $r=1+\varepsilon$ with $\varepsilon>0$. We then find

$$\begin{aligned} G(z,re^{i\theta}) &= \frac{1}{\pi}\frac{1}{re^{i\theta}\bar z - \frac{e^{-i\theta}}{r}z} = \frac{1}{\pi}\frac{1}{(1+\varepsilon)e^{i\theta}\bar z - e^{-i\theta}(1-\varepsilon)z + o(\varepsilon)} \\ &= \frac{1}{2\pi}\frac{1}{-i\mathcal{I}(e^{-i\theta}z)+\varepsilon\mathcal{R}(e^{-i\theta}z)+o(\varepsilon)} = \frac{1}{2i\pi}\frac{-1}{\boldsymbol{\theta}^\perp\cdot\mathbf{x}+i\varepsilon(\boldsymbol{\theta}\cdot\mathbf{x})+o(\varepsilon)}. \end{aligned}$$

Passing to the limit $\varepsilon\to 0$, we obtain

$$G_-(\mathbf{x},\theta) = \frac{1}{2i\pi}\frac{-1}{\boldsymbol{\theta}^\perp\cdot\mathbf{x}+i0\mathrm{sign}(\boldsymbol{\theta}\cdot\mathbf{x})}.$$

Here by convention ± 0 is the limit of $\pm\varepsilon$ as $0<\varepsilon\to 0$. The limit on D^+ is treated similarly. □

We have chosen to define $G_\pm(\mathbf{x},\theta)$ as functions of $\theta\in(0,2\pi)$ instead of functions of $e^{i\theta}$. We have also identified $\mathbf{x}=(x,y)$ with $z=x+iy$. The above lemma gives us a convergence in the sense of distributions. We can equivalently say that for all smooth function $\psi(\mathbf{x})$, we have

$$\int_{\mathbb{R}^2} G_\pm(\mathbf{x}-\mathbf{y},\theta)\psi(\mathbf{y})d\mathbf{y} = \pm\frac{1}{2i}(HR_\theta\psi)(\mathbf{x}\cdot\boldsymbol{\theta}^\perp) + (D_\theta\psi)(\mathbf{x}). \tag{2.44}$$

We recall that the Hilbert transform H is defined in (2.11) and the Radon transform in (2.4)-(2.5).

Proof. The derivation is based on the following result. For any $f(x)\in C_0^\infty(\mathbb{R})$, we have

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}}\frac{f(x)}{ix+\varepsilon}dx = -i\mathrm{p.v.}\int_{\mathbb{R}}\frac{f(x)}{x}dx + \mathrm{sign}(\varepsilon)\pi f(0). \tag{2.45}$$

Exercise 2.4.1 Prove the above limit called Plemelj's formula.

Let us denote $\mathbf{x}=\sigma\boldsymbol{\theta}^\perp+\tau\boldsymbol{\theta}$ and $\mathbf{y}=s\boldsymbol{\theta}^\perp+t\boldsymbol{\theta}$. We have

$$\begin{aligned} &\int_{\mathbb{R}^2} G_+(\mathbf{y})\psi(\mathbf{x}-\mathbf{y})d\mathbf{y} = \frac{1}{2\pi}\int_{\mathbb{R}^2}\frac{\psi((\sigma-s)\boldsymbol{\theta}^\perp+(\tau-t)\boldsymbol{\theta})}{is+0\mathrm{sign}(t)}dsdt \\ &= \frac{1}{2\pi}\int_{\mathbb{R}}\mathrm{p.v.}\int_{\mathbb{R}}\frac{\psi((\sigma-s)\boldsymbol{\theta}^\perp+(\tau-t)\boldsymbol{\theta})}{is}dsdt + \frac{1}{2}\int_{\mathbb{R}}\mathrm{sign}(t)\psi(\sigma\boldsymbol{\theta}^\perp+(\tau-t)\boldsymbol{\theta})dt \\ &= \frac{1}{2\pi}\mathrm{p.v.}\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{\psi((\sigma-s)\boldsymbol{\theta}^\perp+(\tau-t)\boldsymbol{\theta})}{is}dtds + \frac{1}{2}\int_{\mathbb{R}}\mathrm{sign}(t)\psi(\mathbf{x}-t\boldsymbol{\theta})dt \\ &= \frac{1}{2i}(HR_\theta\psi)(\mathbf{x}\cdot\boldsymbol{\theta}^\perp)+(D_\theta\psi)(\mathbf{x}). \end{aligned}$$

A similar derivation yields the limit on D^-. □

We deduce that the function $h(z,\lambda)$ defined in (2.40) admits the limits

$$h_\pm(\mathbf{x},\theta) = \pm\frac{1}{2i}(HR_\theta a)(\mathbf{x}\cdot\boldsymbol{\theta}^\perp)+(D_\theta a)(\mathbf{x}). \tag{2.46}$$

Notice that R_θ and D_θ involve integrations in the direction $\boldsymbol{\theta}$ only so that

$$R_\theta[u(\mathbf{x})v(\mathbf{x}\cdot\boldsymbol{\theta}^\perp)](s) = v(s)R_\theta[u](s), \quad D_\theta[u(\mathbf{x})v(\mathbf{x}\cdot\boldsymbol{\theta}^\perp)](\mathbf{x}) = v(\mathbf{x}\cdot\boldsymbol{\theta}^\perp)D_\theta[u](\mathbf{x}).$$

Using this result and (2.44), we deduce that the limits of the solution $u(z,\lambda)$ to (2.31) are given by

$$\begin{aligned} u^\pm(\mathbf{x},\theta) = e^{-D_\theta a} e^{\frac{\mp 1}{2i}(HR_\theta a)(\mathbf{x}\cdot\boldsymbol{\theta}^\perp)} \frac{\pm 1}{2i} H\Big(e^{\frac{\pm 1}{2i}(HR_\theta a)(s)} R_\theta(e^{D_\theta a} f)\Big)(\mathbf{x}\cdot\boldsymbol{\theta}^\perp) \\ + e^{-D_\theta a} D_\theta(e^{D_\theta a} f)(\mathbf{x}). \end{aligned} \tag{2.47}$$

We recall that $R_\theta(e^{D_\theta a} f) = R_{a,\theta} f(s)$ are our measurements. So whereas u^+ and u^- do not depend only on the measurements (they depend on $D_\theta(e^{D_\theta a} f)(\mathbf{x})$ which is not measured), the difference $u^+ - u^-$ depends *only* on the measurements. This is the property that allows us to invert the attenuated Radon transform. More precisely, let us define

$$\varphi(\mathbf{x},\theta) = (u^+ - u^-)(\mathbf{x},\theta). \tag{2.48}$$

Using (2.47) we deduce that

$$i\varphi(\mathbf{x},\theta) = R^*_{-a,\theta} H_a R_{a,\theta} f(\mathbf{x}), \tag{2.49}$$

where we have defined the following operators

$$\begin{aligned} &R^*_{a,\theta} g(\mathbf{x}) = e^{D_\theta a(\mathbf{x})} g(\mathbf{x}\cdot\boldsymbol{\theta}^\perp), \qquad H_a = C_c H C_c + C_s H C_s \\ &C_c g(s,\theta) = g(s,\theta)\cos(\frac{HR_\theta a(s)}{2}), \quad C_s g(s,\theta) = g(s,\theta)\sin(\frac{HR_\theta a(s)}{2}). \end{aligned} \tag{2.50}$$

Here $R^*_{a,\theta}$ is the formal adjoint operator to $R_{a,\theta}$.

The above derivation shows that $i\varphi(\mathbf{x},\theta)$ is real-valued and of the form $e^{-D_\theta a(\mathbf{x})} M(\mathbf{x}\cdot\boldsymbol{\theta}^\perp,\theta)$ for some function M. We deduce therefore that

$$\boldsymbol{\theta}\cdot\nabla\varphi(\mathbf{x},\theta) + a\varphi(\mathbf{x},\theta) = 0. \tag{2.51}$$

2.4.6 Step (iii): reconstruction formulas

We have seen that $u(z,\lambda)$ is analytic in λ on $D^+\cup D^-$ and is of order $O(z^{-1})$ at infinity. Moreover the jump of $u(z,\lambda)$ across T is given by $\varphi(\mathbf{x},\theta)$ for $0\le\theta<2\pi$. We thus deduce from the Cauchy formula (2.29) that

$$u(\mathbf{x},\lambda) = \frac{1}{2\pi i}\int_T \frac{\varphi(\mathbf{x},t)}{t-\lambda} dt, \qquad \lambda\in D^+\cup D^-. \tag{2.52}$$

where we identify $\varphi(\mathbf{x},t)$ with $\varphi(\mathbf{x},\theta)$ for $t=e^{i\theta}$. We now deduce from (2.33) in the vicinity of $\lambda=0$ that

$$f(\mathbf{x}) = \lim_{\lambda\to 0} \lambda^{-1}\frac{\partial}{\partial\bar{z}} u(\mathbf{x},\lambda). \tag{2.53}$$

Indeed we verify that $u(z,\lambda) = O(\lambda)$ on D^+ so that $a(\mathbf{x})u(\mathbf{x},\lambda)\to 0$ as $\lambda\to 0$. Since $u(\mathbf{x},\lambda)$ is known thanks to (2.52) in terms of the boundary measurements $R_{a,\theta}f(s)$, this is our reconstruction formula. Let us be more specific. We verify that

$$u(\mathbf{x},\lambda) = \frac{1}{2\pi i}\int_T \frac{\varphi(\mathbf{x},t)}{t} dt + \lambda\frac{1}{2\pi i}\int_T \frac{\varphi(\mathbf{x},t)}{t^2} dt + O(\lambda^2). \tag{2.54}$$

We thus deduce from (2.53) and the fact that $u(\mathbf{x},\lambda) = O(\lambda)$ on D^+ that

$$\begin{aligned} 0 &= \frac{1}{2\pi i}\int_T \frac{\varphi(\mathbf{x},t)}{t}dt \\ f(\mathbf{x}) &= \frac{1}{2\pi i}\int_T \frac{\partial\varphi}{\partial\bar{z}}(\mathbf{x},t)\frac{1}{t^2}dt. \end{aligned} \tag{2.55}$$

The second equality is the reconstruction formula we were looking for since φ defined in (2.49) in terms of the measurements $R_{a,\theta}f(s)$. The first equality is a *compatibility* conditions that $i\varphi$ must satisfy in order for the data to be the attenuated Radon transform of a function $f(\mathbf{x})$. This compatibility condition is similar to the condition $g(s,\theta) = g(-s,\theta+\pi)$ satisfied by the Radon transform in the absence of absorption. These compatibility conditions are much more difficult to visualize when absorption does not vanish because the integral along the line $\{s\boldsymbol{\theta}^\perp + t\boldsymbol{\theta}; t\in\mathbb{R}\}$ differs depending on the direction of integration.

Let us recast the reconstruction formula so that it only involves real-valued quantities.

Exercise 2.4.2 Using $t = e^{i\theta}$ and $dt = ie^{i\theta}$, deduce that

$$\begin{aligned} \frac{1}{2\pi i}\int_T \frac{\varphi(\mathbf{x},t)}{t}dt &= \frac{1}{2\pi}\int_0^{2\pi}\varphi(\mathbf{x},\theta)d\theta \\ \frac{1}{2\pi i}\int_T \frac{\partial\varphi}{\partial\bar{z}}(\mathbf{x},t)\frac{1}{t^2}dt &= \frac{1}{4\pi}\int_0^{2\pi}\boldsymbol{\theta}^\perp\cdot\nabla(i\varphi)(\mathbf{x},\theta)d\theta + \frac{1}{4\pi}\int_0^{2\pi}\boldsymbol{\theta}\cdot\nabla\varphi(\mathbf{x},\theta)d\theta. \end{aligned}$$

Use (2.51) and (2.55) to show that

$$f(\mathbf{x}) = \frac{1}{4\pi}\int_0^{2\pi}\boldsymbol{\theta}^\perp\cdot\nabla(i\varphi)(\mathbf{x},\theta)d\theta. \tag{2.56}$$

Let us denote by $g_a(s,\theta) = R_a f(s,\theta)$ the SPECT measurements. From the above results we recast (2.56) as

$$f(\mathbf{x}) = [\mathcal{N}g](\mathbf{x}) \equiv \frac{1}{4\pi}\int_0^{2\pi}\boldsymbol{\theta}^\perp\cdot\nabla(R^*_{-a,\theta}H_a g)(\mathbf{x},\theta)d\theta. \tag{2.57}$$

Exercise 2.4.3 Show that (2.57) simplifies to (2.14) when $a\equiv 0$.

Exercise 2.4.4 (Reconstruction with constant absorption.) We assume that $f(\mathbf{x}) = 0$ for $|\mathbf{x}|\geq 1$ and that $a(\mathbf{x}) = \mu$ for $|\mathbf{x}|<1$. This corresponds thus to the case of a constant absorption coefficient on the support of the source term $f(\mathbf{x})$.
(i) Show that

$$e^{D_\theta a(\mathbf{x})} = e^{\mu\mathbf{x}\cdot\boldsymbol{\theta}}, \qquad |\mathbf{x}|<1.$$

Deduce that

$$\boldsymbol{\theta}^\perp\cdot\nabla(e^{D_\theta a(\mathbf{x})}g(\mathbf{x}\cdot\boldsymbol{\theta}^\perp,\theta)) = e^{\mu\mathbf{x}\cdot\boldsymbol{\theta}}\frac{\partial g}{\partial s}(\mathbf{x}\cdot\boldsymbol{\theta}^\perp).$$

(ii) Verify that the operator H_μ defined by $H_\mu = H_a$ for a constant is diagonal in the Fourier domain and that

$$\widehat{H_\mu u}(\sigma) = -i\mathrm{sign}_\mu(\sigma)\hat{u}(\sigma), \qquad \mathrm{sign}_\mu(\sigma) = \begin{cases} \mathrm{sign}(\sigma) & |\sigma|\geq\mu, \\ 0 & |\sigma|<\mu. \end{cases}$$

(iii) Show that

$$\begin{aligned} g_\mu(s,\theta) &= R_\theta(e^{\mu \mathbf{x}\cdot\boldsymbol{\theta}} f)(s), \\ f(\mathbf{x}) &= \frac{1}{4\pi}\int_0^{2\pi} e^{-\mu \mathbf{x}\cdot\boldsymbol{\theta}}(H_\mu \frac{\partial}{\partial s} g_\mu)(\mathbf{x}\cdot\boldsymbol{\theta}^\perp,\theta)d\theta. \end{aligned} \tag{2.58}$$

Verify that in the limit $\mu \to 0$, we recover the inversion for the Radon transform.

Chapter 3

Diffraction tomography

In this chapter we consider the probing of an object by waves rather than particles. The waves are typically acoustic waves and micro waves. Such waves are characterized by a wavelength small enough so that the features of the object one is interested in can indeed be imaged, but large enough that the high frequency geometrical optics limit (whereby waves are replaced by particles) does not hold. As the frequency increases, the geometrical optics assumption becomes more and more accurate. We will see that the reconstruction in that limit gets closer to the particle reconstruction, i.e., involves inverting a Radon transform.

3.1 Scattering problem

Let us consider the wave equation in dimension $n = 2, 3$ given by

$$\frac{1}{v^2(\mathbf{x})}\frac{\partial^2 U}{\partial t^2} - \Delta U = 0, \qquad \mathbf{x} \in \mathbb{R}^n, \quad t > 0, \tag{3.1}$$

with appropriate initial conditions. The velocity $v(\mathbf{x})$ is unknown. Let us assume that $v(\mathbf{x}) = c$ outside of the object we want to image, where c is a constant. We assume that $U = 0$ for $t < 0$ and pass to the frequency domain by introducing

$$u(\mathbf{x}, \omega) = \int_0^\infty e^{i\omega t} U(\mathbf{x}, t) dt, \qquad U(\mathbf{x}, t) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\omega t} u(\mathbf{x}, \omega) d\omega. \tag{3.2}$$

The equation for u is then the following *Helmholtz* equation

$$\begin{aligned} &(\Delta + \frac{\omega^2}{v^2(\mathbf{x})})u(\mathbf{x}, \omega) = 0, \qquad \mathbf{x} \in \mathbb{R}^n, \quad \omega \in \mathbb{R}, \\ &\hat{\mathbf{x}} \cdot \nabla u(\mathbf{x}, \omega) - i\frac{\omega}{c} u(\mathbf{x}, \omega) = o(|\mathbf{x}|^{-(n-1)/2}). \end{aligned} \tag{3.3}$$

As usual $\hat{\mathbf{x}} = \mathbf{x}/|\mathbf{x}|$ and the second equation is the *radiation* condition, ensuring that no energy comes from infinity (only waves radiating *out* are allowed at infinity). The notation $o(x)$ means a quantity such that $o(x)/x \to 0$ as $0 < x \to 0$. So the decay should be faster than $|\mathbf{x}|^{-1/2}$ in two dimensions and faster than $|\mathbf{x}|^{-1}$ in three dimensions.

Let us now introduce the following expression for the velocity and the frequency:

$$\frac{1}{v^2(\mathbf{x})} = \frac{1}{c^2}(1+\alpha(\mathbf{x})), \qquad k = \frac{\omega}{c}. \tag{3.4}$$

We recast the Helmholtz equation as

$$\begin{aligned} &(\Delta + k^2)u(\mathbf{x},\omega) = -\alpha(\mathbf{x})k^2 u(\mathbf{x},\omega), \\ &\hat{\mathbf{x}}\cdot\nabla u(\mathbf{x},\omega) - iku(\mathbf{x},\omega) = o(|\mathbf{x}|^{-(n-1)/2}). \end{aligned} \tag{3.5}$$

Let $\boldsymbol{\theta} \in S^n$ be a unit vector. We verify that

$$(\Delta + k^2)u_i(\mathbf{x},\omega;\boldsymbol{\theta}) = 0, \qquad \text{where} \quad u_i(\mathbf{x},\omega;\boldsymbol{\theta}) = e^{ik\boldsymbol{\theta}\cdot\mathbf{x}}. \tag{3.6}$$

Thus plane waves with the right wavenumber $k = |k\boldsymbol{\theta}|$ are solutions of the homogeneous Helmholtz equation. Notice however that they *do not* satisfy the radiation conditions (they do radiate out in the direction $\boldsymbol{\theta}$ but certainly not in the direction $-\boldsymbol{\theta}$).

The forward problem we are interested in is the following: we have a plane wave coming from infinity and want to find a solution $u_s(\mathbf{x},\omega)$ modeling the response of the system. Therefore we ask that u_s does not radiate at infinity (i.e., satisfies the radiation condition) and that the whole field $u = u_i + u_s$ satisfies the Helmholtz equation as required by physics. We thus end up with solving the following *scattering* problem

$$\begin{aligned} &(\Delta + k^2)u_s(\mathbf{x},\omega) = -\alpha(\mathbf{x})k^2(u_s+u_i)(\mathbf{x},\omega), \\ &\hat{\mathbf{x}}\cdot\nabla u_s(\mathbf{x},\omega) - iku_s(\mathbf{x},\omega) = o(|\mathbf{x}|^{-(n-1)/2}). \end{aligned} \tag{3.7}$$

In the above equation we have used (3.6). Under some assumptions on $\alpha(\mathbf{x})$, the above equation admits a unique solution [19]. The *inverse scattering* problem consists then of reconstructing $\alpha(\mathbf{x})$ from measurements of u_s at infinity in all possible directions $\mathbf{x}$ for all possible plane waves $\boldsymbol{\theta} \in S^n$. We will not be concerned with this general problem, which is fairly well-understood for $n = 3$ and much less so for $n = 2$. We refer the reader to [19] for details on this difficult theory. What we will do instead is to concentrate on the linearization of the problem about the constant velocity profile $v(\mathbf{x}) = c$. Namely, we will assume that α is small (in some appropriate sense (norm) that we do not describe here). As a consequence, u_s is also small as can be seen from (3.7). We can therefore neglect the term αu_s, which is second-order, as a first approximation. This approximation is called the Born approximation. It also has the advantage of linearizing the problem of reconstructing $\alpha(\mathbf{x})$ from scattering measurements. We are thus now concerned with

$$\begin{aligned} &(\Delta + k^2)u_s(\mathbf{x},\omega) = -\alpha(\mathbf{x})k^2 u_i(\mathbf{x},\omega), \\ &\hat{\mathbf{x}}\cdot\nabla u_s(\mathbf{x},\omega) - iku_s(\mathbf{x},\omega) = o(|\mathbf{x}|^{-(n-1)/2}). \end{aligned} \tag{3.8}$$

This equation can be solved explicitly as

$$u_s(\mathbf{x},\omega) = k^2 \int_{\mathbb{R}^n} \alpha(\mathbf{y})u_i(\mathbf{y},\omega)g_n(\mathbf{x}-\mathbf{y})d\mathbf{y}, \tag{3.9}$$

where g_n is the Green function solving the following equation

$$\begin{aligned} &(\Delta + k^2)g_n(\mathbf{x}) = \delta(\mathbf{x}) \\ &\hat{\mathbf{x}}\cdot\nabla u_s(\mathbf{x},\omega) - iku_s(\mathbf{x},\omega) = o(|\mathbf{x}|^{-(n-1)/2}), \end{aligned} \tag{3.10}$$

and is given for $n = 2, 3$ by

$$g_2(\mathbf{x}) = \frac{i}{4} H_0(k|\mathbf{x}|), \qquad g_3(\mathbf{x}) = \frac{e^{ik|\mathbf{x}|}}{4\pi|\mathbf{x}|}. \tag{3.11}$$

Here, H_0 is the 0th order Hankel function of the first kind, given by

$$H_0(k|\mathbf{x}|) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{1}{\sqrt{k^2 - p^2}} e^{i(px + \sqrt{k^2-p^2}y)} dp, \tag{3.12}$$

where we have decomposed $\mathbf{x} = (x, y)$ in Cartesian coordinates.

3.2 Far field data and reconstruction

The measurements we consider are the far field scattering data. They correspond to the waves propagating outwards at infinity. This simplification amounts to saying that the other component of the radiating field composed of the *evanescent* waves, does not make it to infinity, and thus cannot be measured. Mathematically, this means that we consider the asymptotic limit of u_s as $\mathbf{x} \to \infty$. Let us consider the three dimensional case first. Since $\mathbf{x}$ goes to infinity, $|\mathbf{x} - \mathbf{y}|$ is equal to $|\mathbf{x}|$ plus a smaller order correction. So we have

$$u_s(\mathbf{x}, \omega) = \frac{k^2}{4\pi|\mathbf{x}|} \int_{\mathbb{R}^3} \alpha(\mathbf{y}) e^{ik\boldsymbol{\theta}\cdot\mathbf{y}} e^{ik|\mathbf{x}-\mathbf{y}|} d\mathbf{y} \ + \ \text{l.o.t.} \ .$$

Upon using the following approximation

$$|\mathbf{x} - \mathbf{y}| = |\mathbf{x}||\hat{\mathbf{x}} - \frac{\mathbf{y}}{|\mathbf{x}|}| = |\mathbf{x}|(1 + \frac{|\mathbf{y}|^2}{|\mathbf{x}|^2} - 2\frac{\hat{\mathbf{x}}\cdot\mathbf{y}}{|\mathbf{x}|})^{1/2} = |\mathbf{x}| - \hat{\mathbf{x}}\cdot\mathbf{y} + \ \text{l.o.t.},$$

we obtain

$$u_s(\mathbf{x}, \omega) = \frac{k^2 e^{ik|\mathbf{x}|}}{4\pi|\mathbf{x}|} \int_{\mathbb{R}^3} \alpha(\mathbf{y}) e^{ik(\boldsymbol{\theta}-\hat{\mathbf{x}})\cdot\mathbf{y}} d\mathbf{y} \ + \ \text{l.o.t.} \ .$$

We thus see that

$$\begin{aligned} u_s(\mathbf{x}, \omega) &= \frac{k^2 e^{ik|\mathbf{x}|}}{4\pi|\mathbf{x}|} A(\hat{\mathbf{x}}) + o\Big(\frac{1}{|\mathbf{x}|}\Big), \\ A(\hat{\mathbf{x}}) &= \hat{\alpha}\big(k(\hat{\mathbf{x}} - \boldsymbol{\theta})\big) \ = \int_{\mathbb{R}^3} \alpha(\mathbf{y}) e^{ik(\boldsymbol{\theta}-\hat{\mathbf{x}})\cdot\mathbf{y}} d\mathbf{y}. \end{aligned} \tag{3.13}$$

Recall that $\omega = ck$. So for a plane wave at a given frequency ω, i.e., at a given wavenumber k, and direction $\boldsymbol{\theta}$, the far field measurement is $A(\hat{\mathbf{x}}) = A(\hat{\mathbf{x}}; k, \boldsymbol{\theta})$ in the direction $\hat{\mathbf{x}}$ (obtained by multiplying the measured signal by $4\pi|\mathbf{x}|e^{-ik|\mathbf{x}|}/k^2$).

In two space dimensions ($n = 2$), the final result is similar in the sense that u_s is proportional to $|\mathbf{x}|^{-1/2}$ at infinity with a coefficient of proportionality $A(\hat{\mathbf{x}})$ taking the same expression as given in (3.13).

We observe that each measurement gives us new information about the Fourier transform of the velocity fluctuation $\alpha(\mathbf{x})$. We can distinguish two types of measurements. The first ones correspond to directions of measurements $\mathbf{x}$ such that $\mathbf{x}\cdot\boldsymbol{\theta} > 0$. These

measurements are called *transmission* measurements since they correspond to the radiated wave that have passed through the object to image. The second ones correspond to the directions such that $\mathbf{x} \cdot \boldsymbol{\theta} < 0$ and are called *reflection* measurements. In many settings, we have access to only one type of measurements.

Let us consider transmission measurements first. This means that we have access to $\hat{\alpha}(k(\hat{\mathbf{x}} - \boldsymbol{\theta}))$ for $\hat{\mathbf{x}} \cdot \boldsymbol{\theta} > 0$. In particular we obtain for $\hat{\mathbf{x}} = \hat{\boldsymbol{\theta}}$ the value $\hat{\alpha}(\mathbf{0})$, which is the average of the fluctuation $\alpha(\mathbf{x})$ over the whole domain. More generally as $\hat{\mathbf{x}}$ varies in S^1 such that $\hat{\mathbf{x}} \cdot \boldsymbol{\theta} > 0$, we obtain $\hat{\alpha}(\mathbf{k})$ over a half-circle passing through $\mathbf{0}$, of radius k and symmetric about the axis $\boldsymbol{\theta}$. As $\boldsymbol{\theta}$ varies on the unit circle, we observe that $\hat{\alpha}(k(\hat{\mathbf{x}} - \boldsymbol{\theta}))$ fills in the disk of radius $\sqrt{2}k$. At a fixed value of k, this is therefore all we can get: $\hat{\alpha}(\mathbf{k})$ for $\mathbf{k}$ such that $|\mathbf{k}| \leq \sqrt{2}k$.

The picture in three dimensions is very similar: for a given $\boldsymbol{\theta} \in S^2$, we have access to $\hat{\alpha}(\mathbf{k})$ for $\mathbf{k}$ on a half-sphere of radius $\sqrt{2}k$ passing through $\mathbf{0}$ and invariant by rotation about $\boldsymbol{\theta}$. As $\boldsymbol{\theta}$ varies over the sphere S^2, we thus get $\hat{\alpha}(\mathbf{k})$ for all $\mathbf{k}$ such that $|\mathbf{k}| \leq \sqrt{2}k$, as in the two-dimensional case.

The diffraction inverse problem is therefore *not unique*. All we can reconstruct from the measured data is a low-pass filter of the object $\alpha(\mathbf{x})$. The high frequencies are not measured. The high frequencies of α are encoded in the radiation field u_s. However they are the evanescent part of the waves. They decay therefore much more rapidly than $|\mathbf{x}|^{-1}$ (when $n = 3$), actually exponentially, and thus cannot be measured accurately in practice.

Let us now consider reconstruction formulas. Since frequencies above $\sqrt{2}k$ cannot be reconstructed, we assume that

$$\alpha(\mathbf{x}) = (\mathcal{F}^{-1}_{\mathbf{k}\to\mathbf{x}} \chi_{\sqrt{2}k}(\mathbf{k}) \mathcal{F}_{\mathbf{x}\to\mathbf{k}} \alpha)(\mathbf{x}), \tag{3.14}$$

where $\chi_{\sqrt{2}k}(\mathbf{k}) = 1$ when $|\mathbf{k}| < \sqrt{2}k$ and 0 otherwise,i.e. α does not have high wavenumbers. Then the reconstruction is obviously unique according to what we just saw. Let us consider the two-dimensional case. We want to reconstruct $\alpha(\mathbf{x})$ from $\hat{\alpha}(k(\hat{\mathbf{x}} - \boldsymbol{\theta}))$, where $\hat{\mathbf{x}}$ and $\boldsymbol{\theta}$ run over the unit circle S^1. Notice that we want to reconstruct a two-dimensional function and we have two dimensions of data (one in $\hat{\mathbf{x}}$ and one in $\boldsymbol{\theta}$). The inverse Fourier transform tells us that

$$\alpha(\mathbf{x}) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{i\mathbf{x}\cdot\mathbf{k}} \hat{\alpha}(\mathbf{k}) d\mathbf{k} = \frac{k^2}{(2\pi)^2} \int_0^{2\pi} \int_0^{\sqrt{2}} e^{ik\rho\mathbf{x}\cdot\boldsymbol{\theta}} \hat{\alpha}(k\rho\boldsymbol{\theta}) \rho d\rho d\theta.$$

Observe that as $\boldsymbol{\theta}$ covers the unit circle, all points of the disk $|\mathbf{k}| < \sqrt{2}k$ are covered twice as $\hat{\mathbf{x}}$ varies, once for a point such that $\hat{\mathbf{k}} \cdot \boldsymbol{\theta}^\perp > 0$ and once for a point such that $\hat{\mathbf{k}} \cdot \boldsymbol{\theta}^\perp < 0$. Therefore the information corresponding to $\hat{\mathbf{k}} \cdot \boldsymbol{\theta}^\perp > 0$ is sufficient. This information is parameterized as follows: for a given $\boldsymbol{\theta}$ we write $\hat{\mathbf{x}}$ as

$$\hat{\mathbf{x}}(\phi, \boldsymbol{\theta}) = \sin\phi \boldsymbol{\theta} + \cos\phi \boldsymbol{\theta}^\perp, \qquad 0 \leq \phi \leq \frac{\pi}{2}. \tag{3.15}$$

We thus obtain that

$$\hat{\alpha}(k(\hat{\mathbf{x}} - \boldsymbol{\theta})) = \hat{\alpha}\big(k\rho(\phi)(\frac{\sin\phi - 1}{\rho(\phi)}\boldsymbol{\theta} + \frac{\cos\phi}{\rho(\phi)}\boldsymbol{\theta}^\perp)\big) = \hat{\alpha}\big(k\rho(\phi)\mathcal{R}(\phi)\boldsymbol{\theta}\big),$$

with $\rho(\phi) = \sqrt{2}\sqrt{1-\cos\phi}$ and $\mathcal{R}(\phi)$ an explicitly defined rotation depending on ϕ. Here is the rest of the reconstruction:

$$\begin{aligned}\alpha(\mathbf{x}) &= \frac{k^2}{(2\pi)^2}\int_0^{2\pi}\int_0^{\pi/2} e^{ik\rho(\phi)\mathbf{x}\cdot\boldsymbol{\theta}}\hat{\alpha}(k\rho(\phi)\boldsymbol{\theta})\rho(\phi)\frac{d\rho(\phi)}{d\phi}d\phi d\theta \\ &= \frac{k^2}{(2\pi)^2}\int_0^{2\pi}\int_0^{\pi/2} e^{ik\rho(\phi)\mathbf{x}\cdot\mathcal{R}(\phi)\boldsymbol{\theta}}\hat{\alpha}(k\rho(\phi)\mathcal{R}(\phi)\boldsymbol{\theta})\frac{1}{2}\frac{d\rho^2(\phi)}{d\phi}d\phi d\theta,\end{aligned}$$

so that finally

$$\alpha(\mathbf{x}) = \frac{2k^2}{(2\pi)^2}\int_0^{2\pi}\int_0^{\pi/2} e^{ik\rho(\phi)\mathbf{x}\cdot\mathcal{R}(\phi)\boldsymbol{\theta}}\hat{\alpha}(k|\hat{\mathbf{x}}(\phi,\boldsymbol{\theta})-\boldsymbol{\theta}|)\sin\phi d\phi d\theta. \tag{3.16}$$

This is the reconstruction formula we were after.

Let us conclude the section with a few words about reflection tomography. In that case, we only measure data in directions $\hat{x}$ such that $\hat{x}\cdot\boldsymbol{\theta} < 0$. Following the same techniques as above, we see that we can reconstruct wavenumbers of $\alpha(\mathbf{x})$ in the corona of wavenumbers $\mathbf{k}$ such that $\sqrt{2}k < |\mathbf{k}| < 2k$. The reconstruction from reflection data is therefore by no means unique. We do not get low-frequency components of α and do not get very-high frequencies either. Assuming that the wavenumber content of $\alpha(\mathbf{x})$ is in the above corona, then the reconstruction is unique. A reconstruction formula similar to what we just obtained can also be derived. Notice that when both the transmission and reflection data can be measured, we can reconstruct all wavenumbers $\mathbf{k}$ of $\alpha(\mathbf{x})$ such that $|\mathbf{k}| < 2k$.

All these result are in sharp contrast to the one-dimensional example we saw earlier. There, a given wavenumber k allows us to reconstruct one wavenumber of $\alpha(x)$. All wavenumbers are thus required (i.e. measurements for all frequencies ω) to reconstruct $\alpha(x)$. Here $\alpha(\mathbf{x})$ is also uniquely determined by reconstruction for all values of k (since each value of k allows us to reconstruct all wavenumbers $|\mathbf{k}| < 2k$). However because of the multidimensional nature of the measurements (the variable $\hat{x}$ is discrete in one-dimension instead of living on the unit sphere S^n), measurements for all values of k is very redundant: once we have obtained measurements at a given value of k, all measurements at values less than k are redundant.

3.3 Comparison to X-ray tomography

Let us consider the case of transmission data in two space dimensions. We have seen that wavenumbers of $\alpha(\mathbf{x})$ up to $\sqrt{2}k$ could be reconstructed. However as k tends to ∞, this essentially means that all wavenumbers of $\alpha(\mathbf{x})$ can be reconstructed.

Indeed in that limit we observe that the half circle of radius k becomes the full line orthogonal to $\boldsymbol{\theta}$. That is, as $k\to\infty$, the measurements tend to

$$\hat{\alpha}(\sigma\boldsymbol{\theta}^\perp) = R\alpha(\sigma,\theta+\pi/2).$$

Exercise 3.3.1 Show that the reconstruction formula (3.16) indeed converges to the inverse Radon transform as $k\to\infty$.

In the limit of infinite frequency, we therefore obtain that the transmission measurements tend to the Radon transform of α. We know that the knowledge of $R_\alpha(\sigma, \theta + \pi/2)$ for all values of σ and θ is sufficient to uniquely reconstruct the fluctuation $\alpha(\mathbf{x})$.

So how should we consider the inverse diffraction problem? How ill-posed is it? As we already mentioned, the first problem with diffraction tomography is that for a fixed frequency ω, the function $\alpha(\mathbf{x})$ cannot uniquely be reconstructed. Only the wavenumbers below $\sqrt{2}k$ (below $2k$) in the case of transmission (transmission and reflection) measurements can be reconstructed. However in the class of functions $\alpha(\mathbf{x}) \in L^2(\mathbb{R}^2)$ such that (3.14) holds, we have uniqueness of the reconstruction. In this class we can perform a similar analysis to what we obtained in Theorem 2.2.2.

Let us consider the measurements $d(\phi, \boldsymbol{\theta}) = \hat{\alpha}(k(\hat{\mathbf{x}} - \boldsymbol{\theta}))$ for $\boldsymbol{\theta} \in S^1$ and $0 \leq \phi \leq \pi/2$ using (3.15). We verify that $1 \leq \rho'(\phi) \leq \sqrt{2}$ for $0 \leq \phi \leq \pi/2$.

Let us assume that the error we make is of the same order for every angle ϕ and every angle $\boldsymbol{\theta}$. An estimate of the total error will thus involve

$$\begin{aligned}
&\int_{S^1} \int_0^{\pi/2} |d(\phi, \boldsymbol{\theta})|^2 d\phi d\boldsymbol{\theta} = \int_{S^1} \int_0^{\pi/2} |\hat{\alpha}(k\rho(\phi)\mathcal{R}(\phi)\boldsymbol{\theta})|^2 d\phi d\boldsymbol{\theta} \\
&= \int_{S^1} \int_0^{\sqrt{2}} |\hat{\alpha}(k\rho\boldsymbol{\theta})|^2 (\rho')^{-1} d\rho d\boldsymbol{\theta} \sim \int_{S^1} \int_0^{\sqrt{2}} |\hat{\alpha}(k\rho\boldsymbol{\theta})|^2 d\rho d\boldsymbol{\theta} \\
&\sim \frac{1}{k} \int_{S^1} \int_0^{\sqrt{2}k} |\hat{\alpha}(u\boldsymbol{\theta})|^2 du d\boldsymbol{\theta} \sim \frac{1}{k} \int_{S^1} \int_0^{\sqrt{2}k} \frac{|\hat{\alpha}(u\boldsymbol{\theta})|^2 u \; du \; d\boldsymbol{\theta}}{u} \\
&\geq \frac{1}{k} \|\alpha\|^2_{H^{-1/2}(\mathbb{R}^2)}.
\end{aligned}$$

In some sense, the above formula also shows that the data $d(\phi, \boldsymbol{\theta})$ are more regular than the function $\alpha(\mathbf{x})$ by half of a derivative. This is consistent with the Radon transform in the limit as $k \to \infty$. To be more consistent with the Radon transform, notice in the limit $k \to \infty$ that $k \cos\phi \sim \sigma$ so that $k \sin\phi d\phi \sim d\sigma$ as the half circle converges to the real line. Since $\sin\phi \sim 1$ for most of the wavenumbers σ as $k \to \infty$, this implies that $k d\phi \sim d\sigma$. Therefore a total error in the angular measurements in diffraction tomography consistent with the measurement errors for the Radon transform is given by

$$\int_{S^1} \int_0^{\pi/2} |d(\phi, \boldsymbol{\theta})|^2 k d\phi d\boldsymbol{\theta} \geq \|\alpha\|^2_{H^{-1/2}(\mathbb{R}^2)}.$$

We recover in this limit that the measurements in diffraction tomography smooth the function α by half of a derivative.

We see here again that the ill-posedness of a problem very much depends on the norm in which the error on the data is measured.

Chapter 4

Inverse kinematic problem

This chapter concerns the reconstruction of velocity fields from travel time measurements. The problem can also be recast as the reconstruction of a Riemannian metric from the integration of arc length along its geodesics. We start with a problem with spherical symmetry in which explicit reconstruction formulas are available.

4.1 Spherical symmetry

We want to reconstruct the velocity field $c(r)$ inside the Earth assuming spherical symmetry from travel time measurements. To simplify we assume that the Earth radius is normalized to 1. We assume that $c(1)$ is known. We want to reconstruct $c(r)$ from the time it takes to travel along all possible geodesics. Because the geodesics depend on $c(r)$, the travel time is a *nonlinear* functional of the velocity profile $c(r)$. We thus have to solve a nonlinear inverse problem.

Let us consider the classical mechanics approximation to wave propagation through the Earth. Unlike what we did in Chapter 1, we treat waves here in their geometrical optics limit. This means that we assume that the wavelength is sufficiently large compared to other length scales in the system that the waves can be approximated by particles. The particles satisfy then dynamics given by classical dynamics. They are represented by position $\mathbf{x}$ and wavenumber $\mathbf{k}$ and their dynamics are governed by Hamilton's equations:

$$\begin{aligned} \frac{d\mathbf{x}}{dt} &= \dot{\mathbf{x}} = \nabla_{\mathbf{k}} H(\mathbf{x}(t), \mathbf{k}(t)), \quad \mathbf{x}(0) = \mathbf{x}_0 \\ \frac{d\mathbf{k}}{dt} &= \dot{\mathbf{k}} = -\nabla_{\mathbf{x}} H(\mathbf{x}(t), \mathbf{k}(t)), \quad \mathbf{k}(0) = \mathbf{k}_0. \end{aligned} \tag{4.1}$$

In wave propagation the Hamiltonian is given by

$$H(\mathbf{x}, \mathbf{k}) = c(\mathbf{x})|\mathbf{k}| = c(|\mathbf{x}|)|\mathbf{k}|. \tag{4.2}$$

The latter relation holds because of the assumption of spherical symmetry. Let us denote $\mathbf{x} = r\hat{\mathbf{x}}$ and $\mathbf{k} = |\mathbf{k}|\hat{\mathbf{k}}$. The Hamiltonian dynamics take the form

$$\dot{\mathbf{x}} = c(r)\hat{\mathbf{k}}, \qquad \dot{\mathbf{k}} = -c'(r)|\mathbf{k}|\hat{\mathbf{x}}. \tag{4.3}$$

We are interested in calculating the travel times between points at the boundary of the domain $r = |\mathbf{x}| < 1$. This implies integrating dt along particle trajectories. Since we want to reconstruct $c(r)$, we perform a change of variables from dt to dr. This will allow us to obtain integrals of the velocity $c(r)$ along curves. The objective will then be to obtain a reconstruction formula for $c(r)$.

In order to perform the change of variables from dt to dr, we need to know where the particles are. Indeed the change of variables should only involve position r and no longer time t. This implies to solve the problem $t \mapsto r(t)$. As usual it is useful to find invariants of the dynamical system. The first invariant is the Hamiltonian itself:

$$\frac{dH(\mathbf{x}(t), \mathbf{k}(t))}{dt} = 0,$$

as can be deduced from (4.1). The second invariant is *angular momentum* and is obtained as follows. Let us first introduce the basis $(\hat{\mathbf{x}}, \hat{\mathbf{x}}^\perp)$ for two dimensional vectors (this is the usual basis $(\mathbf{e}_r, \mathbf{e}_\theta)$ in polar coordinates). We decompose $\mathbf{k} = k_r\hat{\mathbf{x}} + k_\theta\hat{\mathbf{x}}^\perp$ and $\hat{\mathbf{k}} = \hat{k}_r\hat{\mathbf{x}} + \hat{k}_\theta\hat{\mathbf{x}}^\perp$. We verify that

$$\dot{r} = c(r)\hat{k}_r \qquad \text{since} \qquad \dot{\mathbf{x}} = \dot{r}\hat{\mathbf{x}} + r\dot{\hat{\mathbf{x}}} = c(r)\hat{\mathbf{k}}. \tag{4.4}$$

We also verify that

$$\frac{d(rk_\theta)}{dt} = \frac{d\mathbf{x}^\perp \cdot \mathbf{k}}{dt} = \dot{\mathbf{x}}^\perp \cdot \mathbf{k} + \mathbf{x} \cdot \dot{\mathbf{k}}^\perp = c(r)\hat{\mathbf{k}}^\perp \cdot \mathbf{k} - c'(r)|\mathbf{k}|\mathbf{x} \cdot \hat{\mathbf{x}}^\perp = 0. \tag{4.5}$$

This is conservation of angular momentum and implies that

$$r(t)k_\theta(t) = k_\theta(0),$$

since $r(0) = 1$.

By symmetry, we observe that the travel time is decomposed into two identical components: the time it takes to go down the Earth until $k_r = 0$, and the time it takes to go back up. On the way up to the surface, k_r is non-negative. Let us denote $p = \hat{k}_\theta(0)$ with $0 < p < 1$. The lowest point is reached when $\hat{k}_\theta = 1$. This means at a point r_p such that

$$\frac{r_p}{c(r_p)} = \frac{p}{c(1)}.$$

To make sure that such a point is uniquely defined, we impose that the function $rc^{-1}(r)$ be *increasing* on $(0, 1)$ since it cannot be decreasing. This is equivalent to the constraint:

$$c'(r) < \frac{c(r)}{r}, \qquad 0 < r < 1. \tag{4.6}$$

This assumption ensures the uniqueness of a point r_p such that $pc(r_p) = c(1)r_p$.

Since the Hamiltonian $c(r)|\mathbf{k}|$ is conserved, we deduce that

$$\dot{r} = c(r)\sqrt{1 - \hat{k}_\theta^2} = c(r)\sqrt{1 - \Big(\frac{\hat{k}_\theta(0)c(r)}{rc(1)}\Big)^2},$$

so that

$$dt = \frac{dr}{c(r)\sqrt{1 - \Big(\frac{\hat{k}_\theta(0)c(r)}{rc(1)}\Big)^2}}. \tag{4.7}$$

Notice that the right-hand side depends only on r and no longer on functions such as $\hat{k}_r$ that depend on time. The travel time as a function of $p = \hat{k}_\theta(0)$ is now given by twice the time it takes to go back to the surface:

$$T(p) = 2\int_{t(r_p)}^{1} dt = 2\int_{r_p}^{1} \frac{dr}{c(r)\sqrt{1 - \Big(\frac{\hat{k}_\theta(0)c(r)}{rc(1)}\Big)^2}}. \tag{4.8}$$

Our measurements are $T(p)$ for $0 < p < 1$ and our objective is to reconstruct $c(r)$ on $(0, 1)$. We need a theory to invert this integral transform. Let us define

$$u = \frac{c^2(1)r^2}{c^2(r)} \quad \text{so that} \quad du = \frac{2rc^2(1)}{c^2(r)}\Big(1 - \frac{rc'(r)}{c(r)}\Big)dr.$$

Upon using this change of variables we deduce that

$$T(p) = 2\int_{p^2}^{1} \Big(\frac{dr}{du}\frac{u}{r}\Big)(u)\frac{du}{\sqrt{u - p^2}}. \tag{4.9}$$

It turns out that the function in parenthesis in the above expression can be reconstructed from $T(p)$. This is an *Abel integral*. Before inverting the integral, we need to ensure that the change of variables $r \mapsto u(r)$ is a diffeomorphism (a continuous function with continuous inverse). This implies that du/dr is positive, which in turn is equivalent to (4.6). The constraint (4.6) is therefore useful both to obtain the existence of a minimal point r_p and to ensure that the above change of variables is admissible. The constraint essentially ensures that no rays are trapped in the dynamics so that energy entering the system will eventually exit it. We can certainly consider velocity profiles such that the energy is attracted at the origin. In such situation the velocity profile cannot be reconstructed.

Let us denote by $f = \frac{dr}{du}\frac{u}{r}$. We will show in the following section that $f(u)$ can be reconstructed from $T(p)$ and is given by

$$f(u) = -\frac{2}{\pi}\frac{d}{du}\int_u^1 \frac{T(\sqrt{p})}{\sqrt{p - u}}dp. \tag{4.10}$$

Now we reconstruct $r(u)$ from the relations

$$\frac{f(u)}{u}du = \frac{dr}{r}, \qquad u(1) = 1, \quad \text{so that} \quad r(u) = \exp\Big(\int_u^1 \frac{f(v)dv}{v}\Big).$$

Upon inverting this diffeomorphism, we obtain $u(r)$ and $g(r) = f(u(r))$. Since

$$g(r) = \frac{1}{2}\frac{1}{1 - rc'/c},$$

we now know rc'/c, hence $(\log c)'$. It suffices to integrate $\log c$ from 1 to 0 to obtain $c(r)$ everywhere. This concludes the proof of the reconstruction.

4.2 Abel integral and Abel transform

For a smooth function $f(x)$ (continuous will do) defined on the interval $(0,1)$, we define the Abel transform as

$$g(x) = \int_x^1 \frac{f(y)}{(y-x)^{1/2}} dy. \tag{4.11}$$

This transform can be inverted as follows:

Lemma 4.2.1 *The Abel transform (4.11) admits the following inversion*

$$f(y) = -\frac{1}{\pi}\frac{d}{dy}\int_y^1 \frac{g(x)}{(x-y)^{1/2}} dx. \tag{4.12}$$

Proof. Let us calculate

$$\int_z^1 \frac{g(x)}{(x-z)^{1/2}} dx = \int_z^1 \int_x^1 \frac{f(y)}{(x-z)^{1/2}(y-x)^{1/2}} dxdy = \int_z^1 dy f(y) k(z,y) dy.$$

The kernel $k(z,y)$ is given by

$$k(z,y) = \int_z^y \frac{dx}{(x-z)^{1/2}(y-x)^{1/2}} = \int_0^1 \frac{dx}{\sqrt{x(1-x)}} = \int_{-1}^1 \frac{dx}{\sqrt{1-x^2}} = \pi.$$

The latter equality comes from differentiating arccos. Thus we have

$$\int_z^1 \frac{g(x)}{(x-z)^{1/2}} dx = \pi \int_z^1 f(y) dy.$$

Upon differentiating both sides with respect to z, we obtain the desired result. □

We can also ask ourselves how well-posed the inversion of the Abel transform is. Since the transforms are defined on bounded intervals, using the Hilbert scale introduced in Chapter 1 would require a few modifications. Instead we will count the number of differentiations. The reconstruction formula shows that the Abel transform is applied once more to g before the result is differentiated. We can thus conclude that the Abel transform regularizes as one half of an integration (since it takes one differentiation to compensate for two Abel transforms). We therefore deduce that the Abel transform is a smoothing operator of order $\alpha = 1/2$ using the terminology introduced in Chapter 1. Inverting the Abel transform is a *mildly ill-posed* problem.

4.3 Kinematic Inverse Source Problem

We have seen that arbitrary spherically symmetric velocity profiles satisfying the condition (4.6) could be reconstructed from travel time measurements. Moreover an explicit reconstruction formula based on the (ill-posed) Abel inverse transform is available.

We now generalize the result to more general velocity profiles that do not satisfy the assumption of spherical symmetry. We start with a somewhat simpler (and linear) problem. Assuming that we have an approximation of the velocity profile and want to obtain a better solution based on travel time measurements. We can linearize the

problem about the known velocity. Once the linearization is carried out we end up with averaging the velocity fluctuations over the integral curves of the known approximation. This is a problem therefore very similar to the Radon transform except that the integration is performed over curves instead than lines.

More generally we give ourselves a set of curves in *two space dimensions* and assume that the integrals of a function $f(\mathbf{x})$ over all curves are known. The question is whether the function $f(\mathbf{x})$ can uniquely be reconstructed. The answer will be yes, except that no explicit formula can be obtained in general. Therefore our result will be a *uniqueness* result.

Let X be a simply connected domain (i.e., $\mathbb{R}^2 \backslash X$ is a connected domain) with smooth boundary $\Sigma = \partial X$. We denote $\overline{X} = X \cup \Sigma$. We parameterize Σ by arc length

$$\mathbf{x} = \boldsymbol{\sigma}(t), \qquad 0 \leq t \leq T.$$

Here T is the total length of the boundary Σ. Obviously, $\boldsymbol{\sigma}(0) = \boldsymbol{\sigma}(T)$.

We give ourselves a *regular family of curves* Γ, i.e., satisfying the following hypotheses:

1. Two points of $\overline{X}$ are joined by a unique curve in Γ.

2. The endpoints of $\gamma \in \Gamma$ belong to Σ, the inner points to X, and the length of γ is uniformly bounded.

3. For every point $\mathbf{x}_0 \in X$ and direction $\boldsymbol{\theta} \in S^1$, there is a unique curve passing through $\mathbf{x}_0$ with tangent vector given by $\boldsymbol{\theta}$. The curve between $\mathbf{x}_0$ and Σ is parameterized by

$$\mathbf{x}(s) = \boldsymbol{\gamma}(\mathbf{x}_0, \boldsymbol{\theta}, s), \qquad 0 \leq s \leq S(\mathbf{x}_0, \boldsymbol{\theta}), \tag{4.13}$$

   where s indicates arc length on γ and S is the distance along γ from $\mathbf{x}_0$ to Σ in direction $\boldsymbol{\theta}$.

4. The function $\boldsymbol{\gamma}(\mathbf{x}_0, \boldsymbol{\theta}, s)$ is of class C^3 of its domain of definition and

$$\frac{1}{s} \boldsymbol{\theta}^{\perp} \nabla_{\boldsymbol{\theta}} \gamma \geq C > 0.$$

These assumptions mean that the curves behave similarly to the set of straight lines in the plane, which satisfy these hypotheses.

Exercise 4.3.1 Show that the lines in the plane satisfy the above hypotheses.

Assuming that we now have the integral of a function $f(\mathbf{x})$ over all curves in Γ joining two points $\boldsymbol{\sigma}(t_1)$ and $\boldsymbol{\sigma}(t_2)$. Let $\gamma(t_1, t_2)$ be this curve and $\mathbf{x}(s; t_1, t_2)$ the points on the curve. The integral is thus given by

$$g(t_1, t_2) = \int_{\gamma(t_1, t_2)} f(\mathbf{x}(s; t_1, t_2)) ds, \tag{4.14}$$

where $ds = \sqrt{dx^2 + dy^2}$ is the usual Lebesgue measure. Our objective is to show that $g(t_1, t_2)$ characterizes $f(\mathbf{x})$. In other words the reconstruction of $f(\mathbf{x})$ is uniquely determined by $g(t_1, t_2)$:

Theorem 4.3.1 *Under the above hypotheses for the family of curves Γ, a function $f \in C^2(\overline{X})$ is uniquely determined by its integrals $g(t_1, t_2)$ given by (4.14) along the curves of Γ. Moreover we have the stability estimate*

$$\|f\|_{L^2(X)} \leq C\|\frac{\partial g(t_1,t_2)}{\partial t_1}\|_{L^2((0,T)\times(0,T))}. \tag{4.15}$$

Proof. We first introduce the function

$$u(\mathbf{x}, t) = \int_{\tilde{\gamma}(\mathbf{x},t)} f ds, \tag{4.16}$$

for $\mathbf{x} \in \overline{X}$ and $0 \leq t \leq T$, where $\tilde{\gamma}(\mathbf{x}, t)$ is the unique segment of curve in Γ joining $\mathbf{x} \in X$ and $\boldsymbol{\sigma}(t) \in \Sigma$. We denote by $\boldsymbol{\theta}(\mathbf{x}, t)$ the tangent vector to $\tilde{\gamma}(\mathbf{x}, t)$ at $\mathbf{x}$. We verify that

$$\boldsymbol{\theta} \cdot \nabla_{\mathbf{x}} u(\mathbf{x}, t) = f(\mathbf{x}). \tag{4.17}$$

This is obtained by differentiating (4.16) with respect to arc length. Notice that $u(\boldsymbol{\sigma}(t_2), t_1) = g(t_1, t_2)$. We now differentiate (4.17) with respect to t and obtain

$$Lu \equiv \frac{\partial}{\partial t}\boldsymbol{\theta} \cdot \nabla_{\mathbf{x}} u = 0. \tag{4.18}$$

As usual we identify $\boldsymbol{\theta} = (\cos\theta, \sin\theta)$. This implies

$$\frac{\partial}{\partial t}\boldsymbol{\theta} = \dot{\theta}\boldsymbol{\theta}^{\perp}, \qquad \frac{\partial}{\partial t}\boldsymbol{\theta}^{\perp} = -\dot{\theta}\boldsymbol{\theta}.$$

Here $\dot{\theta}$ means partial derivative of θ with respect to t. We calculate

$$\boldsymbol{\theta}^{\perp} \cdot \nabla u \frac{\partial}{\partial t}\boldsymbol{\theta} \cdot \nabla u = \dot{\theta}(\boldsymbol{\theta}^{\perp} \cdot \nabla u)^2 + \boldsymbol{\theta}^{\perp} \cdot \nabla u \boldsymbol{\theta} \cdot \nabla(u_t)$$

Here u_t stands for partial derivative of u with respect to t. The same equality with $\boldsymbol{\theta}$ replaced by $\boldsymbol{\theta}^{\perp}$ yields

$$-\boldsymbol{\theta} \cdot \nabla u \frac{\partial}{\partial t}\boldsymbol{\theta}^{\perp} \cdot \nabla u = \dot{\theta}(\boldsymbol{\theta} \cdot \nabla u)^2 - \boldsymbol{\theta} \cdot \nabla u \boldsymbol{\theta}^{\perp} \cdot \nabla(u_t)$$

Upon adding these two equalities, we obtain

$$\begin{aligned} &2\boldsymbol{\theta}^{\perp} \cdot \nabla u \frac{\partial}{\partial t}\boldsymbol{\theta} \cdot \nabla u - \frac{\partial}{\partial t}\Big(\boldsymbol{\theta}^{\perp} \cdot \nabla u \boldsymbol{\theta} \cdot \nabla u\Big) \\ &= \dot{\theta}|\nabla u|^2 + \boldsymbol{\theta} \cdot \nabla(\boldsymbol{\theta}^{\perp} \cdot \nabla u u_t) - \boldsymbol{\theta}^{\perp} \cdot \nabla(\boldsymbol{\theta} \cdot \nabla u u_t). \end{aligned}$$

Why is this relation useful? The answer is the following: one term vanishes thanks to (4.18), one term is positive as soon as u is not the trivial function, and all the other contributions are in divergence form, i.e. are written as derivatives of certain quantities, and can thus be estimated in terms of functions defined at the boundary of the domain (in space $\mathbf{x}$ and arc length t) by applying the Gauss-Ostrogradsky theorem.

More precisely we obtain that

$$\int_0^T \int_X \dot{\theta} |\nabla u|^2 d\mathbf{x} dt = \int_0^T \int_\Sigma [\boldsymbol{\theta}^\perp \cdot \mathbf{n} (\boldsymbol{\theta} \cdot \nabla u) - \boldsymbol{\theta} \cdot \mathbf{n} (\boldsymbol{\theta}^\perp \cdot \nabla u)](\sigma) u_t(\sigma) d\sigma dt.$$

Here $d\sigma$ is the surface (Lebesgue) measure on Σ and $\mathbf{n}$ the outward unit normal vector. Since Σ is a closed curve the term in ∂_t vanishes. The above formula substantially simplifies. Let us assume that Σ has "counterclockwise" orientation. The tangent vector along the contour is $\dot{\boldsymbol{\sigma}}(t)$ so that $\mathbf{n} = -\dot{\boldsymbol{\sigma}}(t)^\perp$, or equivalently $\dot{\boldsymbol{\sigma}}(t) = \mathbf{n}^\perp$. Let us decompose $\boldsymbol{\theta} = \theta_n \mathbf{n} + \theta_\perp \mathbf{n}^\perp$ so that $\boldsymbol{\theta}^\perp = -\theta_\perp \mathbf{n} + \theta_n \mathbf{n}^\perp$. We then verify that

$$\boldsymbol{\theta}^\perp \cdot \mathbf{n} (\boldsymbol{\theta} \cdot \nabla u) - \boldsymbol{\theta} \cdot \mathbf{n} (\boldsymbol{\theta}^\perp \cdot \nabla u) = -\mathbf{n}^\perp \cdot \nabla u = -\dot{\boldsymbol{\sigma}}(t) \cdot \nabla u.$$

However $\dot{\boldsymbol{\sigma}}(t) \cdot \nabla u d\sigma = u_t(\boldsymbol{\sigma}, s) ds$ on Σ so that

$$\int_0^T \int_X \dot{\theta} |\nabla u|^2 d\mathbf{x} dt = -\int_0^T \int_0^T u_t(s) u_t(t) ds dt.$$

This is equivalent to the relation

$$\int_0^T \int_X \dot{\theta} |\nabla u|^2 d\mathbf{x} dt = -\int_0^T \int_0^T \frac{\partial g(t_1, t_2)}{\partial t_1} \frac{\partial g(t_1, t_2)}{\partial t_2} dt_1 dt_2. \tag{4.19}$$

This implies uniqueness of the reconstruction. Indeed, the problem is linear, so uniqueness amounts to showing that $f(\mathbf{x}) \equiv 0$ when $g(t_1, t_2) \equiv 0$. However the above relation implies that $\nabla u \equiv 0$ because $\dot{\theta} > 0$ by assumption on the family of curves. Upon using the transport equation (4.17), we observe that $f(\mathbf{x}) \equiv 0$ as well.

The same reasoning gives us the stability estimate. Indeed we deduce from (4.17) that $|f(\mathbf{x})| \leq |\nabla u(\mathbf{x}, t)|$ so that

$$\int_0^T \int_X \dot{\theta} |f(\mathbf{x})|^2 d\mathbf{x} dt = 2\pi \int_X |f(\mathbf{x})|^2 d\mathbf{x} \leq \int_0^T \int_X \dot{\theta} |\nabla u|^2 d\mathbf{x} dt.$$

Upon using (4.19) and the Cauchy-Schwarz inequality $(f, g) \leq \|f\| \|g\|$, we deduce the result with $C = (2\pi)^{-1/2}$. □

4.4 Kinematic velocity Inverse Problem

Let us return to the construction of the Earth velocity from boundary measurements. Here we consider a "two-dimensional" Earth. Local travel time and distance are related by

$$d\tau^2 = \frac{1}{c^2(\mathbf{x})} (dx^2 + dy^2) = n^2(\mathbf{x}) (dx^2 + dy^2).$$

which defines a Riemannian metric with tensor proportional to the 2×2 identity matrix. We are interested in reconstructing $n(\mathbf{x})$ from the knowledge of

$$G(t_1, t_2) = \int_{\gamma(t_1, t_2)} d\tau = \int_{\gamma(t_1, t_2)} n(\mathbf{x}) \sqrt{dx^2 + dy^2}, \tag{4.20}$$

for every possible boundary points t_1 and t_2, where $\gamma(t_1, t_2)$ is an extremal for the above functional, i.e., a geodesic of the Riemannian metric $d\tau^2$. Notice that since the extremals (the geodesics) of the Riemannian metric depend on $n(\mathbf{x})$, the above problem is *non-linear*.

Let Γ_k for $k = 1, 2$ correspond to measurements for two slownesses n_k, $k = 1, 2$. We then have the following result

Theorem 4.4.1 *Let n_k be smooth positive functions on X such that the family of extremals are sufficiently regular. Then n_k can uniquely be reconstructed from $G_k(t_1, t_2)$ and we have the stability estimate*

$$\|n_1 - n_2\|_{L^2(X)} \leq C\|\frac{\partial}{\partial t_1}(G_1 - G_2)\|_{L^2((0,T)\times(0,T))}. \tag{4.21}$$

Proof. The proof is similar to that of Theorem 4.3.1. The regular family of curves Γ is defined as the geodesics of the Riemannian metric $d\tau^2$. Indeed let us define

$$\tau(\mathbf{x}, t) = \int_{\tilde{\gamma}(\mathbf{x},t)} n ds, \tag{4.22}$$

so that as before $\tau(\boldsymbol{\sigma}(t_1), t_2) = G(t_1, t_2)$. We deduce as before that

$$\boldsymbol{\theta} \cdot \nabla_{\mathbf{x}} \tau = n(\mathbf{x}). \tag{4.23}$$

Because τ is an integration along an extremal curve, we deduce that

$$\boldsymbol{\theta}^\perp \cdot \nabla \tau = 0, \quad \text{so that} \quad \nabla_{\mathbf{x}} \tau = n(\mathbf{x})\boldsymbol{\theta} \quad \text{and} \quad |\nabla_{\mathbf{x}} \tau|^2(\mathbf{x}, t) = n^2(\mathbf{x}).$$

Upon differentiating the latter equality we obtain

$$\frac{\partial}{\partial t}|\nabla_{\mathbf{x}} \tau|^2 = 0.$$

Let us now define $u = \tau_1 - \tau_2$ so that $\nabla u = n_1\boldsymbol{\theta}_1 - n_2\boldsymbol{\theta}_2$. We deduce from the above expression that

$$\frac{\partial}{\partial t}(\nabla u \cdot (\boldsymbol{\theta}_1 + \frac{n_2}{n_1}\boldsymbol{\theta}_2)) = 0.$$

We multiply the above expression by $2\boldsymbol{\theta}_1^\perp \cdot \nabla u$ and express the product in divergence form. We obtain as in the preceding section that

$$\begin{aligned} &2\boldsymbol{\theta}_1^\perp \cdot \nabla u \frac{\partial}{\partial t}\boldsymbol{\theta}_1 \cdot \nabla u - \frac{\partial}{\partial t}\Big(\boldsymbol{\theta}_1^\perp \cdot \nabla u \boldsymbol{\theta}_1 \cdot \nabla u\Big) \\ &= \dot{\theta}_1 |\nabla u|^2 + \boldsymbol{\theta}_1 \cdot \nabla(\boldsymbol{\theta}_1^\perp \cdot \nabla u u_t) - \boldsymbol{\theta}_1^\perp \cdot \nabla(\boldsymbol{\theta}_1 \cdot \nabla u u_t). \end{aligned}$$

We now show that the second contribution can also be put in divergence form. More precisely, we obtain, since n_1 and n_2 are independent of t, that

$$\begin{aligned} &2\boldsymbol{\theta}_1^\perp \cdot \nabla u \frac{\partial}{\partial t}(\frac{n_2}{n_1}\boldsymbol{\theta}_2 \cdot \nabla u) = 2\boldsymbol{\theta}_1^\perp \cdot (n_1\boldsymbol{\theta}_1 - n_2\boldsymbol{\theta}_2)\frac{\partial}{\partial t}(n_2\boldsymbol{\theta}_1 \cdot \boldsymbol{\theta}_2 - \frac{n_2^2}{n_1}) \\ &= -2n_2^2\boldsymbol{\theta}_1^\perp \cdot \boldsymbol{\theta}_2(\boldsymbol{\theta}_1 \cdot \boldsymbol{\theta}_2) = -2n_2^2(\boldsymbol{\theta}_1^\perp \cdot \boldsymbol{\theta}_2)^2\frac{\partial(\theta_1 - \theta_2)}{\partial t} = \\ &-2n_2^2 \sin^2(\theta_1 - \theta_2)\frac{\partial(\theta_1 - \theta_2)}{\partial t} = \frac{\partial}{\partial t}\Big(n_2^2\big[\frac{\sin(2(\theta_1 - \theta_2))}{2} - (\theta_1 - \theta_2)\big]\Big). \end{aligned}$$

The integration of the above term over $X \times (0,T)$ yields thus a vanishing contribution. Following the same derivation as in the preceding section, we deduce that

$$\int_0^T \int_X \frac{\partial \theta_1}{\partial t} |\nabla u|^2 d\mathbf{x} dt = \int_0^T \int_0^T \frac{\partial G(t_1,t_2)}{\partial t_1} \frac{\partial G(t_1,t_2)}{\partial t_2} dt_1 dt_2, \tag{4.24}$$

where we have defined $G = G_1 - G_2$. To conclude the proof, notice that

$$\begin{aligned} &\nabla u \cdot \nabla u = |n_1 \boldsymbol{\theta}_1 - n_2 \boldsymbol{\theta}_2|^2 = n_1^2 + n_2^2 - 2 n_1 n_2 \boldsymbol{\theta}_1 \cdot \boldsymbol{\theta}_2 \\ &\geq n_1^2 + n_2^2 - 2 n_1 n_2 = (n_1 - n_2)^2, \end{aligned}$$

since both n_1 and n_2 are non-negative. With (4.24), this implies that $n_1 = n_2$ when $G_1 = G_2$ and using again the Cauchy-Schwarz inequality yields the stability estimate (4.21). □

Chapter 5

Inverse transport theory

All the inverse problems seen so far were either "well-posed" or "mildly ill-posed" in the sense that they involved a stable inversion operator from a Hilbert space to another Hilbert space. Wave propagation (as in diffraction tomography) or particle propagation along straight lines (as in computerized tomography) thus generate well-posed inverse problems.

What causes an inverse problem to be ill-posed? As we mentioned in the introduction, an inverse problem is ill-posed when the measurement operator is a smoothing/regularizing operator. What are therefore regularizing mechanisms? We have seen that solving the heat equation forward is regularizing. The main reason it is regularizing is "scattering": inherently to the diffusion equation, we assume that "particles" scatterer infinitely often and are represented by some Brownian motion.

Scattering is thus a very powerful mechanism to regularize solutions to PDEs and hence make inverse problems ill-posed. The best equation to see this transition between a non-scattering to a scattering environment is arguably the linear transport equation, also known as the linear Boltzmann equation of the radiative transfer equation.

5.1 Forward transport problem

Physically, we can see the transport equation as modeling the propagation of high frequency waves or particles in scattering environments. In radiative transfer, high frequency wave packets are represented as particles propagating in an environment with changes in the index of refraction and with scatterers. The transport equation then takes the form

$$c(x)\frac{v}{|v|}\cdot\nabla_x u - |v|\nabla c(x)\cdot\nabla_v u + \sigma(x)u = \int_{\mathbb{R}^d} k(x,v',v)u(x,v')\delta(\omega - c(x)|v'|)dv', \quad (5.1)$$

augmented here with boundary conditions at the boundary of a domain of interest. Here, $u(x,v)$ is the density of particles at position x with velocity v. $c(x)$ is the speed of propagation of the high frequency waves. Since we are interested in scattering in this chapter (changes in the index of refraction was considered in the preceding chapter) we assume that $c(x) = 1$ is constant and normalized. The equation thus becomes

$$\frac{v}{|v|}\cdot\nabla_x u + \sigma(x)u = \int_{\mathbb{R}^d} k(x,v',v)u(x,v')\delta(\omega - |v'|)dv'. \quad (5.2)$$

Here, ω is the frequency (e.g. the color) of the propagating waves. Scattering here is elastic so that wave packets with different frequencies satisfy uncoupled equations. Let us normalize $\omega = 1$ so that $|v| = \omega = 1$. In other words, v is now the direction of propagation of the wave packets (photons). We thus have an equation of the form

$$v \cdot \nabla_x u + \sigma(x)u = \int_{\mathbb{S}^{d-1}} k(x, v', v)u(x, v')dv', \tag{5.3}$$

where $x \in \mathbb{R}^d$ and $v \in \mathbb{S}^{d-1}$ the unit sphere in $\mathbb{R}^d$. It remains to describe the main objects participating in scattering: $\sigma(x)$ the attenuation (aka total absorption) coefficient, and $k(x, v', v)$ the scattering coefficient. $\sigma(x)$ models the amount of particles that are either absorbed or scattered per unit distance of propagation. This is the same coefficient already encountered in CT and SPECT. Unlike high-energy CT and SPECT, in lower energies such as for visible light, many "absorbed" photons are re-emitted into another direction, i.e., scattered. Then $k(x, v', v)$ gives the density of particles scattered into direction v from a direction v'. The right-hand side in (5.3) is a *source* term and corresponds to a "creation" of particles into direction v.

In an **inverse transport problem**, $\sigma(x)$ and $k(x, v', v)$ are the unknown coefficients. They have to be reconstructed from knowledge of $u(x, v)$ measured, say, at the boundary of the domain of interest. When $k \equiv 0$, this is Computerized Tomography, where the appropriate logarithm of u provides line integrals of σ.

To be more specific, we need to introduce ways to probe the domain of interest X and to model measurements. In the most general setting, photons can enter into X at any point $x \in \partial X$ and with any incoming direction. In the most general setting, photon densities can then be measured at any point $x \in \partial X$ and for any outgoing direction. The sets of incoming conditions Γ_- and outgoing conditions Γ_+ are defined by

$$\Gamma_\pm = \{(x, v) \in \partial X \times V, \text{ s.t. } \pm v \cdot \nu(x) > 0\}, \tag{5.4}$$

where $\nu(x)$ is the outgoing normal vector to X at $x \in \partial X$ and $V = \mathbb{S}^{d-1}$. Denoting by $g(x, v)$ the incoming boundary conditions, we then obtain the following transport equation

$$\begin{aligned} v \cdot \nabla_x u + \sigma(x)u &= \int_{\mathbb{S}^{d-1}} k(x, v', v)u(x, v')dv', && (x, v) \in X \times V \\ u_{|\Gamma_-}(x, v) &= g(x, v) && (x, v) \in \Gamma_-. \end{aligned} \tag{5.5}$$

From the functional analysis point of view, it is natural to consider the L^1 norm of photon densities, which essentially counts numbers of particles (the L^1 norm of the density on a domain is the number of particles inside that domain). Let us introduce the necessary notation.

We say that the optical parameters (σ, k) are admissible when

$$\begin{aligned} &0 \le \sigma \in L^\infty(X) \\ &0 \le k(x, v', \cdot) \in L^1(V) \text{ a.e. in } X \times V \\ &\sigma_s(x, v') := \int_V k(x, v', v)dv \in L^\infty(X \times V). \end{aligned} \tag{5.6}$$

Here σ_s is also referred to as the scattering coefficient. In most applications, $\sigma_s(x)$ is independent of v'.

We define the times of escape of free-moving particles from X as

$$\tau_\pm(x,v) = \inf\{s>0 | x \pm sv \notin X\} \tag{5.7}$$

and $\tau(x,v) = \tau_+(x,v) + \tau_-(x,v)$. On the boundary sets $\Gamma_\pm$, we introduce the measure $d\xi(x,v) = |v\cdot\nu(x)|d\mu(x)dv$, where $d\mu(x)$ is the surface measure on ∂X.

We define the following Banach space

$$W := \big\{u \in L^1(X\times V) | v\cdot\nabla_x u \in L^1(X\times V), \tau^{-1}u \in L^1(X\times V)\big\}, \tag{5.8}$$

with its natural norm. We recall that τ is defined below (5.7). We have the following trace formula [18]

$$\|f_{|\Gamma_\pm}\|_{L^1(\Gamma_\pm, d\xi)} \le \|f\|_W, \qquad f\in W. \tag{5.9}$$

This allows us to introduce the following lifting operator

$$\mathcal{I}g(x,v) = \exp\Big(-\int_0^{\tau_-(x,v)} \sigma(x-sv,v)ds\Big) g(x-\tau_-(x,v)v, v). \tag{5.10}$$

It is proved in [18] that $\mathcal{I}$ is a bounded operator from $L^1(\Gamma_-, d\xi)$ to W. Note that $\mathcal{I}g$ is the solution u_0 of

$$v\cdot\nabla_x u_0 + \sigma(x)u_0 = 0 \quad (x,v)\in X\times V, \qquad u_0 = g \quad (x,v)\in\Gamma_-.$$

Exercise 5.1.1 *Prove this. This is the same calculation as for the X-ray transform.*

Let us next define the bounded operators

$$\begin{aligned}
\mathcal{K}u(x,v) &= \int_0^{\tau_-(x,v)} \exp\Big(-\int_0^t \sigma(x-sv,v)ds\Big)\int_V k(x-tv,v',v)u(x-tv,v')dv'dt \\
\mathcal{L}S(x,v) &= \int_0^{\tau_-(x,v)} \exp\Big(-\int_0^t \sigma(x-sv,v)ds\Big) S(x-tv,v)dt
\end{aligned} \tag{5.11}$$

for $(x,v)\in X\times V$. Note that $\mathcal{L}S$ is the solution u_S of

$$v\cdot\nabla_x u_S + \sigma(x)u_S = S \quad (x,v)\in X\times V, \qquad u_S = 0 \quad (x,v)\in\Gamma_-.$$

Exercise 5.1.2 *Prove this.*

Note that

$$\mathcal{K}u(x,v) = \mathcal{L}\Big[\int_V k(x,v',v)u(x,v')dv'\Big](x,v),$$

which allows us to handle the right-hand side in (5.5). Looking for solutions in W, the integro-differential equation (5.5) is thus recast as

$$(I-\mathcal{K})u = \mathcal{I}g. \tag{5.12}$$

Exercise 5.1.3 *Prove this.*

Then we have the following result [10, 18].

Theorem 5.1.1 *Assume that*

$$(I - \mathcal{K}) \text{ admits a bounded inverse in } L^1(X \times V, \tau^{-1} dx dv). \tag{5.13}$$

Then the integral equation (5.12) *admits a unique solution* $u \in W$ *for* $g \in L^1(\Gamma_-, d\xi)$. *Furthermore, the albedo operator*

$$\mathcal{A} : L^1(\Gamma_-, d\xi) \to L^1(\Gamma_+, d\xi), \qquad g \mapsto \mathcal{A}g = u_{|\Gamma_+}, \tag{5.14}$$

is a bounded operator.

The invertibility condition (5.13) *holds under either of the following assumptions*

$$\sigma_a := \sigma - \sigma_s \geq 0 \tag{5.15}$$

$$\|\tau \sigma_s\|_\infty < 1. \tag{5.16}$$

We shall not prove this theorem here. The salient features are that the transport equation is well-posed provided that (5.13) is satisfied, which is not necessarily true for arbitrary admissible coefficients (σ, k). The conditions (5.15) or (5.16) are sufficient conditions for (5.13) to be satisfied. The first condition is the most natural for us and states that particles that are "created" by scattering into v by $k(x, v', v)$ are particles that are "lost" for direction v. In other words, the scattering mechanism does not create particles. This is quite natural for photon propagation. In nuclear reactor physics, however, several neutrons may be created by fission for each incoming scattering neutron. There are applications in which (5.15) is therefore not valid. In most medical and geophysical imaging applications, however, (5.15) holds and the transport solution exists. Note that $\sigma_a = \sigma - \sigma_s$ is the *absorption* coefficient, and corresponds to a measure of the particles that are "lost" for direction v' and do not reappear in any direction v (i.e., particles that are absorbed).

Using these operators, we may recast the transport solution as

$$u = \mathcal{I}g + \mathcal{K}\mathcal{I}g + (I - \mathcal{K})^{-1}\mathcal{K}^2\mathcal{I}g, \tag{5.17}$$

where $u_0 := \mathcal{I}g$ is the ballistic component, $u_1 := \mathcal{K}\mathcal{I}g$ the single scattering component and $u_2 := u - u_0 - u_1 = (I - \mathcal{K})^{-1}\mathcal{K}^2\mathcal{I}g$ is the multiple scattering component.

Note that when the problem is subcritical, its solution may be expressed in terms of the following Neumann expansion in $L^1(X \times V)$

$$u = \sum_{m=0}^{\infty} \mathcal{K}^m \mathcal{I}g. \tag{5.18}$$

The contribution $m = 0$ is the ballistic part of u, the contribution $m = 1$ the single scattering part of u, and so on. It is essentially this decomposition of the transport solution into orders of scatterings that allows us to stably reconstruct the optical parameters in the following sections. Note that the above Neumann series expansions has an additional benefit. Since the optical parameters are non-negative, each term in the

above series is non-negative provided that g and S are non-negative so that the transport solution itself is non-negative. A little more work allows us to prove the maximum principle, which states that u in $X \times V$ is bounded a.e. by the (essential) supremum of g in Γ_- when $S \equiv 0$.

Finally, the albedo operator $\mathcal{A}$, which maps incoming conditions to outgoing densities models our measurements. We control the fluxes of particles on Γ_- and obtain information about X by measuring the density of particles on Γ_+. This allows us to define the *measurements operator* of *inverse transport*. Let

$$\mathfrak{X} = \{(\sigma, k) \text{ such that } 0 \leq \sigma \in L^\infty(X),\ 0 \leq k \in L^\infty(X \times V \times V),\ \sigma \geq \sigma_s\}, \tag{5.19}$$

and let $Yf = \mathcal{L}(L^1(\Gamma_-, d\xi), L^1(\Gamma_+, d\xi))$. Then we define the measurement operator $\mathfrak{M}$

$$\mathfrak{M}: \quad \mathfrak{X} \ni (\sigma, k) \mapsto \mathfrak{M}(\sigma, k) = \mathcal{A}[(\sigma, k)] \in \mathfrak{Y}, \tag{5.20}$$

where $\mathcal{A}[(\sigma, k)]$ is the albedo operator constructed in (5.14) with coefficient (σ, k) in (5.5). Note that the measurement operator, as for the Calderón problem in (1.4) is an operator, which to a set of coefficients maps a coefficient-valued operator, the albedo operator.

The main question of inverse transport consists of knowing what can be reconstructed in (σ, k) from knowledge of the full operator $\mathfrak{M}$ or knowledge of only parts of the operator $\mathfrak{M}$. In these notes, we shall mostly be concerned with the full measurement operator.

5.2 Inverse transport problem

One of the main results for inverse transport is the decomposition (5.17). The first term $u_0 := \mathcal{I}g$ is the ballistic component and corresponds to the setting of vanishing scattering. It is therefore the term used in CT and the X-ray transform. It turns out that this term is more *singular*, in a sense that will be made precise below, than the other contributions. It can therefore be extracted from the rest and provide the X-ray transform of the attenuation coefficient σ.

The second term in (5.17) is $u_1 := \mathcal{K}\mathcal{I}g$ and is the the single scattering component of the transport solution. Finally, $u_2 := u - u_0 - u_1 = (I - \mathcal{K})^{-1}\mathcal{K}^2\mathcal{I}g$ is the multiple scattering component, which corresponds to particles that have interacted at least twice with the underlying medium. A fundamental property of the transport equation is that single scattering is also more singular than multiple scattering in dimension three (and higher dimensions), but not in dimension two. We shall describe in more detail below in which sense single scattering is more *singular*. The main conclusion, however, is that the single scattering contribution can also be extracted from the full measurement. As we shall see, single scattering provides a linear operator to invert the scattering coefficient k once σ is known.

Multiple scattering is then *less singular* than ballistic and single scattering. This is difficult to quantify mathematically. Intuitively, this means that multiple scattering contributions are smoother functions. In some sense, after multiple scattering, we do not expect the density to depend too much on the exact location of the scattering events. Multiply scattered particles visit a large domain and hence are less specific about the scenery they have visited.

5.2.1 Decomposition of the albedo operator and uniqueness result

Following (5.17), we decompose the albedo operator as

$$\begin{aligned} \mathcal{A}g &= \mathcal{I}g\big|_{\Gamma_+} + \mathcal{K}\mathcal{I}g\big|_{\Gamma_+} + \mathcal{K}^2(I-\mathcal{K})^{-1}\mathcal{I}g\big|_{|\Gamma_+} \\ &:= \mathcal{A}_0 g + \mathcal{A}_1 g + \mathcal{A}_2 g. \end{aligned} \tag{5.21}$$

We denote by α the Schwartz kernel of the albedo operator $\mathcal{A}$:

$$\mathcal{A}g(x,v) = \int_{\Gamma_-} \alpha(x,v,y,w) g(y,w) d\mu(y) dw.$$

Any linear operator, such as the albedo operator, admits such a decomposition. Knowledge of the operator $\mathcal{A}$ is equivalent to knowledge of its kernel α. The decomposition for $\mathcal{A}$ then translates into the decomposition for α:

$$\alpha = \alpha_0 + \alpha_1 + \alpha_2.$$

Here, α_0 corresponds to the ballistic part of the transport solution, α_1 corresponds to the single scattering part of the transport solution, and α_2 corresponds to the rest of the transport solution.

After some algebra that we shall not reproduce here, we have the following decompositions in the time independent case:

$$\alpha_0(x,v,y,w) = \exp\Big(-\int_0^{\tau_-(x,v)} \sigma(x-sv,v)ds\Big)\delta_v(w)\delta_{\{x-\tau_-(x,v)v\}}(y). \tag{5.22}$$

$$\begin{aligned} \alpha_1(x,v,y,w) = \int_0^{\tau_-(x,v)} \exp\Big(-\int_0^t \sigma(x-sv,v)ds - \int_0^{\tau_-(x-tv,w)} \sigma(x-tv-sw,w)ds\Big) \\ k(x-tv,w,v)\delta_{\{x-tv-\tau_-(x-tv,w)w\}}(y)dt. \end{aligned} \tag{5.23}$$

Exercise 5.2.1 *Prove the above two formulas. Note that the first formula is nothing but the expression for the X-ray transform.*

Note that α_0 and α_1 are distributions: they involve "delta" functions. A difficult result shows that the kernel corresponding to multiple scattering α_2 is in fact a *function*, and when k is bounded, satisfies that

$$|\nu(y)\cdot w|^{-1}\alpha_2(x,v,y,w) \in L^\infty(\Gamma_-, L^p(\Gamma_+, d\xi)), \quad 1 \le p < \frac{d+1}{d}. \tag{5.24}$$

The exact regularity of the function α_2 is not very important for us here. For the mathematical derivation of stability estimates, the fact that we can take $p > 1$ above is important. For a picture of the geometry of the singularities, see Fig. 5.1.

The strategy to recover σ and k in dimension $d \ge 3$ thus goes as follows: We send beams of particles into the medium that concentrate in the vicinity of a point

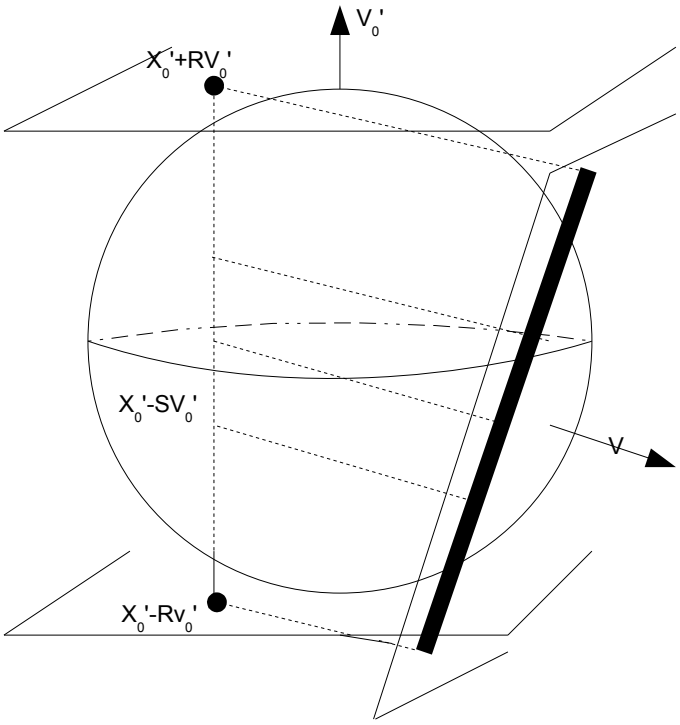

Figure 5.1: Geometry of single scattering for $n = 3$.

$(y_0, v_0) \in \Gamma_-$. More precisely, let g_ε be a sequence of normalized L^1 functions on Γ_- converging as $\varepsilon \to 0$ to $\delta_{v_0}(v)\delta_{\{y_0\}}(y)$.

Since the ballistic term is more singular than the rest, if we place detectors on the support of the ballistic term, then such detectors will overwhelmingly be measuring ballistic particles and very very little scattered particles. Since single scattering is more singular than multiple scattering, we can use the same strategy and place detectors at the location of the support of single scattering. Such detectors will overwhelmingly be measuring singly scattered particles and very very little multiply scattered particles. Mathematically, the details are a little painful. Let us describe them nonetheless.

Recovery of the attenuation coefficient $\sigma(x)$

Let $(y_0, v_0) \in \Gamma_-$ be defined as above and $(x_0, v_0) \in \Gamma_+$ such that $y_0 = x_0 - \tau_-(x_0, v_0)v_0$. Let ϕ_ε be a sequence of bounded functions on Γ_+ equal to 1 in the vicinity of (x_0, v_0) and with vanishing support as $\varepsilon \to 0$. Then we verify [10, 18] that

$$\int_{\Gamma_+\times\Gamma_-} \alpha_m(x, v, y, w)\phi_\varepsilon(x, v)g_\varepsilon(y, w)d\mu(x)dvd\mu(y)dw \xrightarrow{\varepsilon\to 0} 0, \quad m = 1, 2,$$

so that

$$\begin{aligned} \langle \phi_\varepsilon, \mathcal{A}g_\varepsilon\rangle \quad &:= \quad \int_{\Gamma_+\times\Gamma_-} \alpha(x, v, y, w)\phi_\varepsilon(x, v)g_\varepsilon(y, w)d\mu(x)dvd\mu(y)dw \\ &\xrightarrow{\varepsilon\to 0} \quad \exp\Big(-\int_0^{\tau_-(x_0,v_0)} \sigma(x_0 - sv_0, v_0)ds\Big). \end{aligned} \tag{5.25}$$

In other words, the measurements corresponding to the above choices of functions g_ε and ϕ_ε then converge to the function

$$\langle \phi_\varepsilon, \mathcal{A} g_\varepsilon \rangle \xrightarrow{\varepsilon \to 0} E(x,y) := \exp\Big(- \int_0^{|x-y|} \sigma\big(x - s\frac{x-y}{|x-y|}\big) ds \Big). \tag{5.26}$$

This proves that knowledge of the albedo operator $\mathcal{A}$, which allows one to construct $\langle \phi_\varepsilon, \mathcal{A} g_\varepsilon \rangle$, provides knowledge of $E(x,y)$, the exponential of minus the X-ray transform of σ along the segment (x,y). This can be obtained from any segment (x,y) with x and y on ∂X. As a consequence, knowledge of $\mathcal{A}$ provides knowledge of the X-ray transform of σ. We know that σ is then uniquely and stably reconstructed from such knowledge since the X-ray transform is a smoothing operator by (only) one-half of a derivative.

Recovery of the scattering coefficient $k(x,v,w)$

We assume that $\sigma = \sigma(x)$ is now recovered. Let $z_0 \in X$, $v_0 \in V$, and $v_0 \neq w_0 \in V$. Define $x_0 = z_0 + \tau_+(z_0, v_0) v_0$ so that $(x_0, v_0) \in \Gamma_+$ and $y_0 = z_0 - \tau_-(z_0, w_0) w_0$ so that $(y_0, w_0) \in \Gamma_-$. We formally show how the scattering coefficient may be uniquely reconstructed from full knowledge of $\mathcal{A}$.

Let us define g_{ε_1} as before and ϕ_ε as a sequence of bounded functions on Γ_+ equal to a constant in the vicinity of (x_0, v_0) and with vanishing support as $\varepsilon \to 0$. Since $v_0 \neq w_0$, we find that

$$\int_{\Gamma_+ \times \Gamma_-} \alpha_0(x,v,y,w) \phi_\varepsilon(x,v) g_{\varepsilon_1}(y,w) d\mu(x) dv d\mu(y) dw = 0, \quad 0 \leq \varepsilon, \varepsilon_1 < \varepsilon_0(x_0, v_0, y_0, w_0).$$

i.e., the ballistic contribution vanishes with such measurements. Let us define g_{ε_1} such that $|\nu(y_0) \cdot w_0|^{-1} g_{\varepsilon_1}(y,w)$ converges to a delta function. The factor $|\nu(y_0) \cdot w_0|^{-1}$ is here to ensure that the number of emitted particles is independent of y_0 and w_0. The ballistic part of the transport solution is then approximately concentrated on the line passing through y_0 and with direction w_0. Scattering occurs along this line and particles scattered in direction v_0 are approximately supported on the plane with directions v_0 and w_0 passing through x_0. The intersection of that plane with the boundary ∂X is a one-dimensional *curve* $\gamma(x_0, v_0, w_0) \subset X$. In two space dimensions, the curve γ has the same dimension as ∂X. As a consequence, α_1 is a function and therefore is not more singular than α_2 in the time independent setting when $d = 2$.

Let $\phi_\varepsilon(x,v)$ be a bounded test function supported in the $\varepsilon-$vicinity of γ. Because γ is of measure 0 in ∂X when $d \geq 3$, we find using (5.24) that

$$\int_{\Gamma_+ \times \Gamma_-} \alpha_2(x,v,y,w) \phi_\varepsilon(x,v) g_{\varepsilon_1}(y,w) d\mu(x) dv d\mu(y) dw \xrightarrow{\varepsilon, \varepsilon_1 \to 0} 0,$$

i.e., the multiple scattering contribution is asymptotically negligible with such measurements. Now, choosing $\phi_\varepsilon(x,v)$ properly normalized and supported in the ε_2-vicinity of (x_0, v_0) (for $\varepsilon \ll \varepsilon_2 \ll 1$), we find that

$$\langle \phi_\varepsilon, \mathcal{A} g_{\varepsilon_1} \rangle \xrightarrow{\varepsilon, \varepsilon_1, \varepsilon_2 \to 0} E(y_0, z_0) E(z_0, x_0) k(z_0, w_0, v_0),$$

at each point of continuity of $k(z_0, w_0, v_0)$, where $E(x, y)$ is defined in (5.26). Since $\sigma(x)$ and hence $E(x, y)$ are known from knowledge of $\mathcal{A}$, then so is $k(z_0, w_0, v_0)$ at each point of continuity in $X \times V \times V$ thanks to the above formula.

The reconstruction of the attenuation coefficient works for any dimension $d \geq 2$. The reconstruction of the scattering coefficient, however, works only in dimension $d \geq 3$. The reason is that in dimension $d = 2$, the single scattering contribution is also a function, not a distribution, and thus cannot be separated from the multiple scattering component. What we have obtained so far may be summarized as:

Theorem 5.2.1 ([18]) *Let (σ, k) and $(\tilde{\sigma}, \tilde{k})$ be two admissible pairs of optical parameters associated with the same albedo operator $\mathcal{A}$ and such that σ and $\tilde{\sigma}$ are independent of the velocity variable. Then $\sigma = \tilde{\sigma}$ in dimension $d \geq 2$. Moreover, $k = \tilde{k}$ in dimension $d \geq 3$.*

5.2.2 Stability in inverse transport

Let us assume the existence of two types of measurements $\mathcal{A}$ and $\tilde{\mathcal{A}}$, say, corresponding to the optical parameters (σ, k) and $(\tilde{\sigma}, \tilde{k})$, respectively. The question of the stability of the reconstruction is to bound the errors $\sigma - \tilde{\sigma}$ and $k - \tilde{k}$ as a function of $\mathcal{A} - \tilde{\mathcal{A}}$.

We obtain stability estimates in dimension $d \geq 3$. In dimension $d = 2$, only the estimate on $\sigma(x)$ is valid. The construction of the incoming source $\phi_\varepsilon(x, v)$ is such that $\phi_\varepsilon \in C^1(\Gamma_-)$ is supported in the ε_1 vicinity of (x_0, v_0) and normalized so that $\int_{\Gamma_-} \phi_\varepsilon d\xi = 1$. Let ψ be a compactly support continuous function, which models the array of detectors, on Γ_+ such that $\|\psi\|_\infty \leq 1$. Then

$$\Big| \int_{\Gamma_+} \psi(x, v) \big((\mathcal{A} - \tilde{\mathcal{A}}) \phi_\varepsilon \big)(x, v) d\xi(x, v) \Big| \quad \leq \quad \|(\mathcal{A} - \tilde{\mathcal{A}})\|_{\mathcal{L}(L^1)}, \tag{5.27}$$

where now $\| \cdot \|_{\mathcal{L}(L^1)} = \| \cdot \|_{\mathcal{L}(L^1(\Gamma_-, d\xi), L^1(\Gamma_+, d\xi))}$. We still introduce

$$I_m(\psi, \varepsilon) = \int_{\Gamma_+} \psi(x, v) \big((\mathcal{A}_m - \tilde{\mathcal{A}}_m) \phi_\varepsilon \big)(x, v) d\xi(x, v), \qquad m = 0, 1, 2,$$

and obtain that

$$\begin{aligned} \lim_{\varepsilon \to 0+} I_0(\psi, \varepsilon) &= \psi(y_0, v_0) \Big(E(x_0, y_0) - \tilde{E}(x_0, y_0) \Big) \\ \lim_{\varepsilon \to 0+} I_1(\psi, \varepsilon) &= \int_V \int_0^{\tau_+(x_0, v_0)} \psi(x(s) + \tau_+(x(s), v)v, v) \big(E_+ k - \tilde{E}_+ \tilde{k} \big)(x(s), v_0, v) ds dv \end{aligned} \tag{5.28}$$

where we have introduced $x(s) = x_0 + s v_0$.

The estimate (5.24) allows us to show that

$$\big| I_2(\psi, \varepsilon) \big| \leq C \int_V \Big(\int_{\partial X} |\psi(t, x, v)|^{p'} dx \Big)^{\frac{1}{p'}} dv, \qquad p' > d. \tag{5.29}$$

Multiple scattering is therefore still negligible when the support of $\psi := \psi_\lambda$ tends to 0 when $\lambda \to 0$.

The first sequence of functions ψ_λ is chosen to have a small support concentrated in the vicinity of $(y_0, v_0) \in \Gamma_+$. Then the single scattering contribution $I_1(\psi_\lambda) \to 0$ as

$\lambda \to 0$. For w_0 fixed in V, we choose the sequence of functions ψ_λ such that they are concentrated in the vicinity of the curve $\gamma(s)$ on Γ_+ and that they approximately take the value $\text{sign}(E_+k - \tilde{E}_+\tilde{k})(x_0 + sv_0, v_0, w_0)$ along that curve. Since $v_0 \neq w_0$, we verify that $I_0(\psi_\lambda) \to 0$ for such sequence of functions; see [10]. Now, the function ψ_λ has a small support only in dimension $d \geq 3$. Indeed, in dimension $d = 2$, the curve γ has the same dimensionality as the boundary ∂X. When $d = 2$, multiple scattering may no longer be separated from single scattering by using the singular structure of the albedo operator $\mathcal{A}$. This allows us to state the result:

Theorem 5.2.2 ([10]) *Assume that $\sigma(x)$ and $k(x, v', v)$ are continuous on $\bar{X}$ and $\bar{X} \times V \times V$, respectively and that $(\tilde{\sigma}, \tilde{k})$ satisfy the same hypotheses. Let $(x_0, v_0) \in \Gamma_-$ and $y_0 = x_0 + \tau_+(x_0, v_0)v_0$. Then we have for $d \geq 2$ that*

$$\big|E(x_0, y_0) - \tilde{E}(x_0, y_0)\big| \leq \|\mathcal{A} - \tilde{\mathcal{A}}\|_{\mathcal{L}(L^1)}, \tag{5.30}$$

while in dimension $d \geq 3$, we have

$$\int_V \int_0^{\tau_+(x_0,v_0)} \big|E_+k - \tilde{E}_+\tilde{k}\big|(x_0 + sv_0, v_0, v)dsdv \leq \|\mathcal{A} - \tilde{\mathcal{A}}\|_{\mathcal{L}(L^1)}. \tag{5.31}$$

The stability obtained above for the X-ray transform of the absorption coefficient is not sufficient to obtain any stability of σ itself without a priori regularity assumptions on σ. This results from the well known fact that the X-ray transform is a smoothing (compact) operator so that the inverse X-ray transform is an unbounded operator. Let us assume that σ belongs to some space $H^s(\mathbb{R}^d)$ for s sufficiently large and that σ_p defined in (5.6) is bounded. More precisely, define

$$\mathcal{M} = \big\{(\sigma, k) \in C^0(\bar{X}) \times C^0(\bar{X} \times V \times V) | \sigma \in H^{\frac{d}{2}+r}(X), \|\sigma\|_{H^{\frac{d}{2}+r}(X)} + \|\sigma_p\|_\infty \leq M\big\}, \tag{5.32}$$

for some $r > 0$ and $M > 0$. Then, we have the following result.

Theorem 5.2.3 ([10, 11]) *Let $d \geq 2$ and assume that $(\sigma, k) \in \mathcal{M}$ and that $(\tilde{\sigma}, \tilde{k}) \in \mathcal{M}$. Then the following is valid:*

$$\|\sigma - \tilde{\sigma}\|_{H^s(X)} \leq C\|\mathcal{A} - \tilde{\mathcal{A}}\|^\kappa_{\mathcal{L}(L^1)}, \tag{5.33}$$

where $-\frac{1}{2} \leq s < \frac{d}{2} + r$ and $\kappa = \frac{d+2(r-s)}{d+1+2r}$.

When $d \geq 3$, we have

$$\|k - \tilde{k}\|_{L^1(X\times V\times V)} \leq \|\mathcal{A} - \tilde{\mathcal{A}}\|^{\kappa'}_{\mathcal{L}(L^1)}\big(1 + \|\mathcal{A} - \tilde{\mathcal{A}}\|^{1-\kappa'}_{\mathcal{L}(L^1)}\big), \tag{5.34}$$

where $\kappa' = \frac{2(r-r')}{d+1+2r}$ and $0 < r' < r$.

Such estimates show that under additional regularization assumptions on σ, we have explicit stability expression of Hölder type on σ and k. The first stability result (5.33) was first established in [36].

The proof of these theorems is fairly technical and will not be presented here in detail. The reason that such technicalities appear is that we have to estimate the influence of multiple scattering

Chapter 6

Very brief introduction to propagation of singularities

The previous chapter showed that the inverse transport problem was a well-posed problem (or at the very least a mildly ill-posed problem) because the forward transport problem *propagates singularities*. What do we mean by this? We mean that singularities in the object we want to reconstruct, namely the optical parameters, propagated as singularities in the measurement operator, namely the albedo operator. For inverse transport, we actually showed a little more. We showed, at least in dimension $d \geq 3$, that the *whole* Radon transform of $\sigma(x)$ and the *whole* scattering coefficient $k(x, v', v)$ propagated as *singularities* of the albedo operator (respectively describing its ballistic and single scattering components).

It is the existence of singularities in the measurement operator that allowed us to obtain the crucial stability estimates that characterize the behavior of an inverse problem. Assume that the albedo operator is replaced by its angular version, in the following sense: the source $g = g(x)$ is independent of v on Γ_- and the measurements at the detector are of the form $\int_{v\cdot\nu(x)>0} v \cdot \nu(x) u(x, v) dv$ for $x \in \partial X$. In other words, we assume that the source is isotropic at the surface and that the detectors measure the outgoing current of particles. This is typically the setting in Optical Tomography, where both the source and the detector are optical fibers that can be modeled as little discs emitting photons in every direction and collecting photons coming from all directions. Let $\mathcal{B}$ the corresponding measurement operator. Such an operator has a Schwartz kernel given by $\beta(x, y)$ related to the Schwartz kernel $\alpha(x, v, y, w)$ of $\mathcal{A}$ according to the relation

$$\beta(x, y) = \int_{v\cdot\nu(x)>0,\, w\cdot\nu(y)<0} \alpha(x, v, y, w) v \cdot \nu(x) \, |w \cdot \nu(y)| \, dv \, dw. \tag{6.1}$$

Of course, we have a decomposition $\beta = \beta_0 + \beta_1 + \beta_2$ of contributions of ballistic, single scattering, and multiple scattering photons as we have in the angularly resolved case. Yet, it is very clear that $\beta_0(x, y)$ no longer has any of the singularities displayed by $\alpha_0(x, v, y, w)$. Therefore, β_0 can no longer be extracted from β using singularities. This does not mean that the information about σ is not present in β. It is. But it is no longer possible to use singularities of β to extract σ and k. In fact, we can show that the reconstruction of optical parameters from knowledge of β is a *severely ill-posed* problem. We refer the reader to [8] for more details on that inverse problem and other

inverse problems related to inverse transport theory for various (time-dependent, time-independent, angularly resolved, angularly averaged) measurement operators.

6.1 Notion of singularities

So we know that singularities need to propagate for an inverse problem to be well-posed (or mildly ill-posed). When singularities do not propagate, it often means that they are smoothed out. The forward operator is then smoothing and the inverse problem is severely ill-posed. Can we be a little more precise? The answer is yes, but unfortunately the mathematical machinery is far from trivial. A very precise notion of singularities of functions is that of *Wave Front Sets*. We shall not present such a notion here but still would like to present some of its flavors. The presentation will be very heuristic.

Particles and phase space description. Let us think of particles. They propagate along known trajectories and we observe them at known times at the end of the trajectory. Sending them back to their original location is simple: all we have to do is change the sign of their momentum vector and send them back. This is a simple "inverse problem" where singularities (of where particles were originally) propagates to the measurements. What do we need to represent particles? We need position and momentum. With unit speed, this means we need position and direction. The initial positions and directions of the particles become final positions and directions with a one-to-one correspondence, which makes the inverse problem well-posed. We would like to mimic this idea for more complex functions and more complex operator. But we want to retain the idea of localization in space and in directions. For a "function" $f(x)$ to be localized in space and in direction means that the support of f is close to a point x_0 but the Fourier content of $f(x)$ is also close to a plane wave. Think of a localized wave packet with a given orientation. That wave packet is a "function" and yet has the characteristics of a particle, with a rather well-defined position and direction.

Wave packets with scale, position, direction. Now an arbitrary function will be given as a superposition of such wave packets. This is where the mathematical machinery becomes more difficult. It turns out that with some multiscale analysis, we can indeed represent functions as a superposition of wave packets with prescribed *scale*, position, and direction. Then the forward operator will map such wave packets to other "things". If it turns out that these "things" are in fact wave packets in the image space, then we have a very good description of the forward operator: it maps wave packets with a given scale, position, and direction, to other wave packets with typically the same scale but other positions and directions.

It turns out that such a description is sufficiently rich to apply to a lot of problem and "sufficiently simple" to provide very intuitive characterizations of many important inverse problems. There are two main applications where such descriptions are very useful: solutions of wave-type equations and integrals of objects along curves. Integrals along curves are reminiscent of the Radon transform and this is the example we shall treat in detail.

6.2 Radon transform and singularity propagation

We said that a wave packet had a given scale, a position, and an orientation. As the scaling parameter tends to 0, the wave packet is becoming smaller and smaller, and thus becomes a better and better representation of what we can call a "singularity". For a typical singularity, assume that a two dimensional function is equal to 0 for $-1 \ll x_2 < 0$ and is equal to 1 for $0 < x_2 \ll 1$, and this for $|x_1| \ll 1$. In other words, the action takes place close to $x = (x_1, x_2) = 0$ and the jump (singularity) of the function is in the direction e_2. Now translate and rotate the picture and we get a function that can have a singularity at any point and in any direction.

The wave packet for us will be represented by the following function in the so-called parabolic scaling:

$$f_h(x) = \frac{1}{h^{\frac{3}{2}}} f\Big(\frac{x_1}{h}, \frac{x_2}{h^2}\Big). \tag{6.2}$$

Let us assume that f is a smooth function with compact support. Then the above scaling makes that f_h has an extension of order $O(h)$ in the x_1 variable and of order $O(h^2)$ in the x_2 variable. It is therefore localized at $x = 0$ in the plane. But it also clearly displays faster oscillations in the x_2 variable than in the x_1 variable. It is therefore also localize in the "direction" variable, in the sense that it oscillates in the direction e_2. The factor $h^{\frac{3}{2}}$ is here to normalize the function f_h so that its $L^2(\mathbb{R}^2)$ norm is independent of h, say equal to 1 for concreteness.

The space-direction singularity of f_h is therefore given by $(0, e_2)$, i.e., spatial singularity at $x = 0$ and directional singularity in direction e_2. Now let us define

$$\tau_y f(x) = f(x - y), \quad R_{\boldsymbol{\psi}} f(x) = f(R_{-\boldsymbol{\psi}} x),$$

where $R_{\boldsymbol{\psi}} x$ is the rotation of x by ψ. Finally, let

$$f_h(x; y, \psi) = \tau_y R_{\boldsymbol{\psi}} f_h(x), \tag{6.3}$$

the translation-rotation of f_h. Then we observe that f_h has a spatial singularity at $x = y$ and a directional singularity in direction $\boldsymbol{\psi}$. The space-direction singularity of f_h is thus given by $(y, \boldsymbol{\psi})$. This covers the whole three-dimensional space of (microlocal) singularities.

Now we recall that the Radon transform is given by

$$Rf(s, \theta) = \int_{\mathbb{R}} f(s\boldsymbol{\theta}^{\perp} + t\boldsymbol{\theta}) dt, \quad \boldsymbol{\theta} = (\cos\theta, \sin\theta), \quad \boldsymbol{\theta}^{\perp} = (-\sin\theta, \cos\theta).$$

We want to understand the Radon transform of $f_h(x; y, \psi)$ and "see" where the singularities in the (x_1, x_2) variables propagate. Let us start with f_h. We write

$$\begin{aligned} Rf_h(s, \theta) &= \frac{1}{h^{\frac{3}{2}}} \int_{\mathbb{R}} f\Big(\frac{-s\sin\theta + t\cos\theta}{h}, \frac{s\cos\theta + t\sin\theta}{h^2}\Big) dt \\ &= \frac{1}{h^{\frac{3}{2}}} \int_{\mathbb{R}} f(-s_h \sin\theta_h + \tau\cos\theta_h, s_h\cos\theta_h + \tau\sin\theta_h) \frac{dt}{d\tau}(s, \theta) d\tau. \end{aligned}$$

The first line is what we have; the second line what we want. In other words, can we find a change of variables such that the above holds? Of course, we have to be lucky

for this to happen. We want $\frac{dt}{d\tau}$ to be independent of τ in order to write the integral as the Radon transform of the function f. For the Radon transform, we are lucky since we can solve:

$$\frac{s\sin\theta}{h} = s_h \sin\theta_h, \quad \frac{t\cos\theta}{h} = \tau\cos\theta_h, \quad \frac{s\cos\theta}{h^2} = s_h\cos\theta_h, \quad \frac{t\sin\theta}{h^2} = \tau\sin\theta_h.$$

Indeed, we verify that the change of variables is

$$\tan\theta_h = \frac{1}{h}\tan\theta, \qquad s_h = \frac{s}{h^2}\frac{\cos\theta}{\cos\theta_h} = \frac{s}{h}\frac{\sin\theta}{\sin\theta_h}, \quad \tau = \frac{t}{h}\frac{\cos\theta}{\cos\theta_h}. \tag{6.4}$$

We verify that $\frac{dt}{d\tau}$ depends on (s,θ) but not on the variable of integration t.

We have thus found that

$$Rf_h(s,\theta) = \frac{1}{h^{\frac{1}{2}}}\frac{\cos\theta_h}{\cos\theta}Rf(s_h,\theta_h) = \frac{1}{h^{\frac{1}{2}}}\frac{\cos\theta_h}{\cos\theta}Rf\Big(\frac{\cos\theta_h}{\cos\theta}\frac{s}{h^2}, \tan^{-1}\frac{1}{h}\tan\theta\Big). \tag{6.5}$$

Now, from the change of variables, we find that for θ_h of order $O(1)$, then θ is of order $O(h)$. Thus, θ is localized close to 0. Obviously, when h is small, s is of order $O(h^2)$ in order for Rf_h to be bounded away from 0. We thus see that in the variables (s,θ), then $Rf_h(s,\theta)$ has a phase-space (position and direction) singularity given by $(0,0,e_s)$.

So the singularity $(0,0,e_2)$ is mapped to the singularity $(0,0,e_s)$.

How about the other singularities? Quite generally, using the expression for the Radon transform, we find that

$$R(\tau_y R_{\boldsymbol{\psi}} f)(s,\theta) = Rf(s - y\cdot\boldsymbol{\theta}^\perp, \theta - \psi).$$

Applying this to f_h, we thus find that

$$R(f_h(\cdot;y,\psi))(s,\theta) = \frac{1}{h^{\frac{1}{2}}}wRf\Big(w\frac{s - y\cdot\boldsymbol{\theta}^\perp}{h^2}, \tan^{-1}\frac{1}{h}\tan(\theta-\psi)\Big),$$

where $w = w(s,\theta,y,\psi)$ is a weight. In order words, the singularity is at "position" $s = y\cdot\boldsymbol{\theta}^\perp$ and $\theta = \psi$ and is in the direction $d(s - y\cdot\boldsymbol{\theta}^\perp) = ds + y\cdot\boldsymbol{\psi}d\theta$, or equivalently, in the direction (after normalization) $\frac{1}{\sqrt{1+(y\cdot\boldsymbol{\psi})^2}}(e_s + y\cdot\boldsymbol{\psi}e_\theta)$.

Thus the singularity $(y_1,y_2,\boldsymbol{\psi})$ is mapped to $\Big(y\cdot\boldsymbol{\theta}^\perp, \psi, \frac{1}{\sqrt{1+(y\cdot\boldsymbol{\psi})^2}}(e_s + y\cdot\boldsymbol{\psi}e_\theta)\Big)$.

This is an excellent first result: it shows that singularities in the object of interest, namely $f(x)$, propagate to singularities in the data, namely the Radon transform. However, as such, propagation of singularities is not sufficient. We also need to make sure that different singularities go to different places. Otherwise, the inverse map of singularity propagation would not be defined as a function. In other words, the map of singularities needs to be injective. Is this the case? Yes, it is. Assume we have a singularity at $(y\cdot\boldsymbol{\theta}^\perp,\psi)$ with direction $e_s + y\cdot\boldsymbol{\psi}e_\theta$ (we forget about normalization of the unit vector, all that matters is the direction given by the vector, whether it is normalized or not; in fact, all that matters is the line given by the vector since when a function jumps in a direction, it clearly jumps in the opposite direction. This is the reason why projective spaces sometimes make an apparition in the analysis of singularities). We see that ψ is uniquely defined. But so are $y\cdot\boldsymbol{\theta}^\perp$ and $y\cdot\boldsymbol{\psi}$, which also uniquely determine y.

So there is a one-to-one correspondence between singularities. The Radon transform is indeed a nice inverse problem: it propagates singularities and allows us to bring them back in place unambiguously.

We also observe the regularizing effect of the Radon transform. The power of h in the above term is $h^{-\frac{1}{2}}$ rather than $h^{-\frac{3}{2}}$ as was the case in the x variables. The reason is that the wave packet has been attenuated by $h = \sqrt{h^2}$, where h^{-2} is the frequency of the original wave packet. We observe again that the Radon transform is smoothing by half a derivative.

These ideas look deceptively simple. They can in fact provide tremendous information about an inverse problem. Assume that two singularities make it to the same place in theory but that only one of the singularities is present in given data. Then an inversion algorithm will back-propagate the singularity in the data to *both* places without prior information. If you know the bad singularity (ghost singularity) does not exist, then you can remove it from the reconstruction. If you don't know that, then you will get ghost singularities in the reconstruction. The propagation of singularity tools are there to make this concrete. Unfortunately, the mathematics associated to such constructions is often impenetrable. We leave it here at this very superficial level.

Chapter 7

Cauchy Problem and Inverse Diffusion Problem

This chapter introduces two classical examples of severely ill-posed problems for which the measurement operator is still injective. In the absence of any noise, reconstructions would be exact. However, stability is so dire that even minute amounts of moderately high frequency noise will be amplified so much during the reconstruction that the resulting image will often be worthless. We have seen in earlier chapters that solving the heat equation backward was an example of a severely ill-posed problem. Here, we consider two new examples: the Cauchy problem, which is an inverse source problem, and the Inverse Diffusion problem, also known as the Calderón problem, which is a nonlinear inverse problem.

7.1 Cauchy Problem

We start with the Cauchy problem for elliptic equations. As an application in medical imaging, we mention the monitoring of the electrical activity of the heart. The electric problem is modeled by a Laplace equation. We first describe in more detail the medical application and next analyze the Cauchy problem for the Laplace equation.

7.1.1 Electrocardiac potential

Let us consider the application of imaging the electrical activity of the heart. The problem consists of measuring the electric potential on the endocardial surface (the inside of the cardial wall). A possible method consists of sending a probe inside the heart and to measure the potential on the probe. The inverse problem consists then of reconstructing the potential on the endocardial surface from the measurements.

The problem is modeled as follows. Let Γ_0 be a closed smooth surface in $\mathbb{R}^3$ representing the endocardial surface and let Γ_1 be the closed smooth surface inside the volume enclosed by Γ_0 where the measurements are performed. We denote by X the domain with boundary $\partial X = \Gamma_0 \cup \Gamma_1$; see Fig.7.1. The electric potential solves the

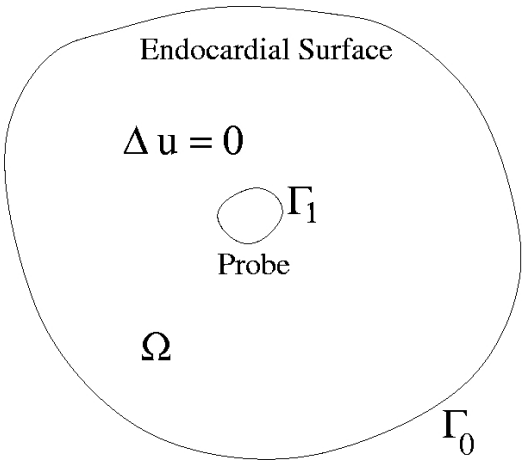

Figure 7.1: Geometry of endocardial measurements

following Laplace equation:

$$\begin{aligned} \Delta u &= 0 && \text{in } X, \\ u &= u_1 && \text{on } \Gamma_1, \\ \frac{\partial u}{\partial \mathbf{n}} &= 0 && \text{on } \Gamma_1. \end{aligned} \tag{7.1}$$

The function u_1 models the measurements at the surface of the probe, which is assumed to be insulated so that $\mathbf{n} \cdot \nabla u = 0$ on Γ_1. The objective is then to find $u = u_0$ on Γ_0. As we will see, this is a severely ill-posed problem.

7.1.2 Half Space Problem

In this section we first consider the case of data given on the real line $y = 0$. In section 7.1.2 we aim at reconstructing the potential on a parallel line $y = \rho$. We show in section 7.1.2 that the same method can be used to analytically continue an analytic function described on the real line.

Electrocardiac application

Let us consider the following two dimensional problem in the upper half plane:

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} &= 0, && x \in \mathbb{R},\, y > 0 \\ u(x,0) &= u_0(x), && x \in \mathbb{R} \\ \frac{\partial u}{\partial y}(x,0) &= g_0(x), && x \in \mathbb{R}. \end{aligned} \tag{7.2}$$

Let us denote by

$$\hat{u}(k_x, y) = (\mathcal{F}_{x \to k_x} u)(k_x, y). \tag{7.3}$$

Upon Fourier transforming (7.2) in the x variable, we obtain

$$\begin{aligned} &-k_x^2\hat{u} + \frac{\partial^2 \hat{u}}{\partial y^2} = 0, && k_x \in \mathbb{R},\ y > 0 \\ &\hat{u}(k_x, 0) = \hat{u}_0(k_x), && k_x \in \mathbb{R}, \\ &\frac{\partial \hat{u}}{\partial y}(k_x, 0) = \hat{g}_0(k_x), && k_x \in \mathbb{R}. \end{aligned} \tag{7.4}$$

The solution of the above ODE is given by

$$\hat{u}(k_x, y) = \hat{u}_0(k_x)\cosh(|k_x|y) + \frac{\hat{g}_0(k_x)}{|k_x|}\sinh(|k_x|y). \tag{7.5}$$

Let us now assume that

$$|k_x|\hat{u}_0(k_x) + \hat{g}_0(k_x) = 0, \tag{7.6}$$

so that

$$\hat{u}(k_x, y) = \hat{u}_0(k_x)e^{-|k_x|y}. \tag{7.7}$$

Upon inverting the Fourier transform, we obtain that the solution $u(x, y)$ is given by

$$u(x, y) = (u_0 * \frac{1}{\pi}\frac{y}{x^2 + y^2})(x) = \frac{1}{\pi}\int_{\mathbb{R}} u_0(x - z)\frac{y}{z^2 + y^2}dz. \tag{7.8}$$

Exercise 7.1.1 (i) Prove (7.8). Hint: Use (1.23) and show that

$$\frac{1}{2\pi}\int_{\mathbb{R}} e^{-y|k_x|}e^{ixk_x}dk_x = \frac{1}{\pi}\frac{y}{x^2 + y^2}.$$

(ii) Show that $\frac{1}{\pi}\frac{y}{x^2+y^2}$ is the fundamental solution of (7.2) with $u_0(x) = \delta(x)$. Calculate the corresponding value of $g_0(x)$.

Notice that the compatibility condition (7.6) can be recast in the physical domain as

$$g_0(x) = -Hu_0'(x), \tag{7.9}$$

where H is the Hilbert transform defined in (2.11). Provided the above compatibility condition is met, the problem (7.2) admits a unique solution and is well-posed, for instance in the sense that

$$\int_{\mathbb{R}} u^2(x, y)dx \leq \int_{\mathbb{R}} u_0^2(x)dx, \qquad \text{for all } y > 0.$$

This is an immediate consequence of (7.7) and the Parseval relation.

However in general, (7.9) is not satisfied. For instance in the electrocardial potential application, we have $g_0 = 0$. In the Fourier domain the solution of the Laplace equation (7.9) with $g_0 = 0$ is thus given by

$$\hat{u}(k_x, y) = \hat{u}_0(k_x)\cosh(|k_x|y). \tag{7.10}$$

This implies that high frequencies are exponentially amplified when $y > 0$. Let us assume that we are interested in the potential at $y = y_0$. The forward problem, in

which $u(x, y_0)$ is known and $u(x, 0) = A[u(x, y_0)]$ is measured, is a well-posed problem in $L^2(\mathbb{R})$. Indeed we easily verify from (7.10) that

$$\|Av\|_{L^2(\mathbb{R})} \leq \|v\|_{L^2(\mathbb{R})}.$$

The inverse problem, which to $u_0(x)$ maps $u(x, y_0)$ is however severely ill-posed. Indeed since $\cosh(|k_x|y_0)$ cannot be bounded by $C(1 + |k_x|^2)^{\alpha/2}$ for any value of α, we see that (1.33) cannot hold for any α. The reason is that the forward operator A is more smoothing than an arbitrary number of anti-differentiations.

The inverse operator A^{-1} can be defined only for sufficiently smooth functions. Indeed let us consider the space of functions

$$X_y(\mathbb{R}) = \{u \in L^2(\mathbb{R}); \cosh(|k_x|y)\hat{u}(k_x) \in L^2(\mathbb{R})\}. \tag{7.11}$$

We verify that A is continuous in $\mathcal{L}(L^2(\mathbb{R}), X_{y_0}(\mathbb{R}))$ and that its inverse A^{-1} is in $\mathcal{L}(X_{y_0}(\mathbb{R}), L^2(\mathbb{R}))$ and is given by

$$A^{-1}u = \mathcal{F}^{-1}_{k_x \to x} \cosh(|k_x|y)\mathcal{F}_{x \to k_x} u. \tag{7.12}$$

Unless noise in the data belongs to X_{y_0}, the noise will be amplified by (7.12) in the reconstruction.

Analytic continuation

Let us now apply the same type of technique to the analytic continuation of an analytic function given on the real line.

Let $f(z) = g(z) + ih(z)$ be an analytic function with $g(z)$ and $h(z)$ real valued-functions. Let us assume that $g(z)$ and $h(z)$ are known on the real line $\Im(z) = 0$. The objective is to find them for arbitrary values of z. We identify $z = x + iy$ and assume that $g(x, 0)$ and $h(x, 0)$ are distributions in $\mathcal{S}'(\mathbb{R})$ so that their Fourier transform is defined. Since $f(z)$ is analytic, we have

$$\frac{\partial f}{\partial \bar{z}} = 0,$$

or equivalently that

$$\frac{\partial g}{\partial x} - \frac{\partial h}{\partial y} = 0, \qquad \frac{\partial g}{\partial y} + \frac{\partial h}{\partial x} = 0. \tag{7.13}$$

These Cauchy-Riemann relations imply that g and h are harmonic, i.e., $\Delta g = \Delta h = 0$. They thus solve the following problems

$$\begin{array}{ll} \Delta g = 0, \qquad y > 0, & \Delta h = 0, \qquad y > 0, \\ \dfrac{\partial g}{\partial y}(x, 0) = -\dfrac{\partial h}{\partial x}(x, 0), & \dfrac{\partial h}{\partial y}(x, 0) = \dfrac{\partial g}{\partial x}(x, 0), \\ g(x, 0) \text{ known}, & h(x, 0) \text{ known}. \end{array} \tag{7.14}$$

Both problems are of the form (7.2). The solutions in the Fourier domain are given by

$$\begin{aligned} \hat{g}(k_x, y) &= \hat{g}(k_x, 0)\cosh(|k_x|y) - i\text{sign}(k_x)\hat{h}(k_x, 0)\sinh(|k_x|y) \\ \hat{h}(k_x, y) &= \hat{h}(k_x, 0)\cosh(|k_x|y) + i\text{sign}(k_x)\hat{g}(k_x, 0)\sinh(|k_x|y) \end{aligned} \tag{7.15}$$

We verify that the problem is well-posed provided that (7.9) is verified, which in this context means

$$\frac{\partial h}{\partial x} = H\frac{\partial g}{\partial x}, \qquad \frac{\partial g}{\partial x} = -H\frac{\partial h}{\partial x}. \tag{7.16}$$

Notice that $H^2 = -I$ so that both equalities above are equivalent. When the above conditions are met, then the analytic continuation is a stable process. When they are not met, we have seen that high frequencies exponentially increase as y increases, which renders the analytic continuation process a severely ill-posed problem.

7.1.3 General two dimensional case

We now consider the arbitrary two-dimensional geometries described in section 7.1.1. We use the Riemann mapping theorem to map such geometries conformally to an annulus (the region lying between two concentric circles) in the plane. We then solve the problem on the annulus. The Riemann mapping gives us a stable way to transform the original domain to an annulus and back. We will see that solving the problem on the annulus is severely ill-posed similarly to what we saw in the preceding section.

Laplace equation on an annulus

Let us begin with a presentation of the problem on the annulus. We assume that the inner circle has radius 1 and the outer circle radius $\rho > 0$. By simple dilation, this is equivalent to the more general case of two circles of arbitrary radius a and b. In polar coordinates the Laplacian takes the form

$$\Delta u = \frac{1}{r}\frac{\partial}{\partial r}\Big(r\frac{\partial u}{\partial r}\Big) + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2}. \tag{7.17}$$

The general solution to the above equation periodic in θ is decomposed in Fourier modes as

$$u(r,\theta) = a_0 + b_0 \ln r + \sum_{n\in\mathbb{N}^*}\Big(\frac{a_n}{2}r^n + \frac{b_n}{2}r^{-n}\Big)e^{in\theta}. \tag{7.18}$$

Since $\mathbf{n}\cdot\nabla u = 0$ on at $r = 1$, we deduce that b_0 and $b_n - a_n$ vanish. We then find that the solution to (7.1) on the annulus is given by

$$u(r,\theta) = \sum_{n\in\mathbb{N}}\Big(\frac{1}{2\pi}\int_0^{2\pi} e^{-in\phi}u_1(\phi)d\phi\Big)\frac{r^n + r^{-n}}{2}e^{in\theta}. \tag{7.19}$$

The above solution holds for all $r > 1$.

Exercise 7.1.2 Find the relation that $u_1(\theta) = u(1,\theta)$ and $g_1(\theta) = \mathbf{n}\cdot\nabla u(1,\theta)$ must satisfy so that the problem

$$\begin{aligned} &\Delta u = 0, \qquad |r| > 1 \\ &u(1,\theta) = u_1(\theta), \quad \mathbf{n}\cdot\nabla u(1,\theta) = g_1(\theta), \qquad 0 \le \theta < 2\pi, \end{aligned}$$

is well-posed (in the sense that the energy of $\theta \to u(\rho,\theta)$ is bounded by the energy of $u_1(\theta)$). Compare to (7.6) and (7.9).

We verify that an error of size δ in the measurement of the coefficient a_n is amplified into an error of order

$$A_n(\rho) = \frac{\delta}{2} e^{n \ln \rho} \qquad \text{at} \quad r = \rho.$$

We verify that A_n cannot be bounded by any Cn^α for all $\alpha > 0$, which would correspond to differentiating the noise level α times. This implies that the reconstruction of $u(\rho, \theta)$ from $u_1(\theta)$ using (7.19) is a severely ill-posed problem.

Riemann mapping theorem

Let X be an open smooth two dimensional domain with smooth boundary having two smooth connected components. We denote by Γ_0 the outer component of ∂X and Γ_1 the inner component; see Fig. 7.1.

For $z \in X \subset \mathbb{C}$ we construct a holomorphic function $\Psi(z)$ (i.e., a function such that $\frac{\partial \Psi}{\partial \bar{z}} = 0$) mapping X to an annulus of the form $1 < r < \rho$. The function is constructed as follows. Let first v be the unique solution to the following Dirichlet problem

$$\Delta v = 0 \ \text{ on } X, \quad v_{|\Gamma_0} = 1, \qquad v_{|\Gamma_1} = 0. \tag{7.20}$$

For some c to be fixed later, let $G = cv$. We verify that

$$I = \int_{\Gamma_1} -\frac{\partial G}{\partial y} dx + \frac{\partial G}{\partial x} dy = -c \int_{\Gamma_1} \frac{\partial v}{\partial \boldsymbol{\nu}} ds > 0$$

by the maximum principle (v is positive inside X and its normal derivative must be negative on Γ_1). We fix c such that $I = 2\pi$. In that case we can define a function $H(z)$, a conjugate harmonic of $G(z)$ on X, by

$$H(z) = \int_p^z -\frac{\partial G}{\partial y} dx + \frac{\partial G}{\partial x} dy, \tag{7.21}$$

where p is an arbitrary point in X. Since G is harmonic, we verify that the definition of H is independent of the path chosen between p and z. Moreover we verify that

$$\frac{\partial H}{\partial x} = -\frac{\partial G}{\partial y}, \qquad \frac{\partial H}{\partial y} = \frac{\partial G}{\partial x},$$

so that $G + iH$ is a holomorphic function on X. Then so is

$$\Psi(z) = e^{G(z) + iH(z)}. \tag{7.22}$$

We verify that $\Psi(z)$ maps Γ_0 to the circle $|z| = e^c$ and Γ_1 to the circle $|z| = 1$. Moreover Ψ is a diffeomorphism between X and the annulus $U_c = \{z \in \mathbb{C},\, 1 \leq |z| \leq e^c\}$. Finally we verify that $\Delta \Psi(z) = 0$ on X since Ψ is holomorphic. Therefore we have replaced the problem (7.1) on X by solving the Laplace equation on U_c with boundary conditions $u_1(\Psi(z))$ on the circle $r = 1$ and vanishing Neumann boundary conditions (we verify that Neumann boundary conditions are preserved by the map Ψ).

7.2 Inverse Diffusion Problem

We now consider the problem of *electrical impedance tomography*, which is a non invasive technique that consists of reconstructing the conductivity coefficient of a diffusion equation from current and voltage boundary measurements. The technical proofs of some theorems are not presented.

7.2.1 Introduction

More precisely consider the Dirichlet problem

$$\begin{aligned} &L_\gamma u(\mathbf{x}) \equiv \nabla \cdot \gamma(\mathbf{x}) \nabla u(\mathbf{x}) = 0, \qquad \mathbf{x} \in X \\ &u(\mathbf{x}) = f(\mathbf{x}), \qquad \mathbf{x} \in \partial X, \end{aligned} \tag{7.23}$$

where $X \subset \mathbb{R}^3$ is a bounded domain with smooth boundary ∂X. The Dirichlet-to-Neumann or voltage-to-current map is given by

$$\Lambda_\gamma(f) = \gamma \frac{\partial u}{\partial \nu}\Big|_{\partial X}. \tag{7.24}$$

The inverse problem consists of reconstructing γ from the knowledge of Λ_γ.

We first assume that γ is a smooth function and perform the following change of variables:

$$\gamma^{-1/2} L_\gamma \gamma^{-1/2} = \Delta - q, \qquad q = \frac{\Delta \gamma^{1/2}}{\gamma^{1/2}}. \tag{7.25}$$

Here Δ is the usual Laplacian operator.

Exercise 7.2.1 Prove (7.25).

Therefore provided that γ is known at the boundary of the domain (we shall come back to this point later), it is sufficient to reconstruct $q(x)$ from the boundary measurements to reconstruct $\gamma(\mathbf{x})$ thanks to (7.25).

7.2.2 Exponential solutions

Let $\boldsymbol{\rho}$ be a complex-valued vector in $\mathbb{C}^n$ for $n \geq 2$ such that $\boldsymbol{\rho} \cdot \boldsymbol{\rho} = 0$. We then verify that $e^{\mathbf{x}\cdot\boldsymbol{\rho}}$ is harmonic, i.e.,

$$\Delta e^{\mathbf{x}\cdot\boldsymbol{\rho}} = \boldsymbol{\rho} \cdot \boldsymbol{\rho}\ e^{\mathbf{x}\cdot\boldsymbol{\rho}} = 0. \tag{7.26}$$

It turns out that there are solutions of $\Delta - q$ that are not very different from $e^{\mathbf{x}\cdot\boldsymbol{\rho}}$. Here is the technical result (without proof)

Theorem 7.2.1 *Let $q \in L^\infty(\mathbb{R}^n)$ such that $q(\mathbf{x}) = 0$ for $|\mathbf{x}| \geq R > 0$. Let $-1 < \delta < 0$. Then there exists $\varepsilon(\delta)$ such that for every $\boldsymbol{\rho} \in \mathbb{C}^n$ satisfying the two constraints:*

$$\boldsymbol{\rho} \cdot \boldsymbol{\rho} = 0, \quad \text{and} \quad |\boldsymbol{\rho}| \geq \frac{1}{\varepsilon(\delta)} \|(1 + |\mathbf{x}|^2)^{1/2} q\|_{L^\infty(\mathbb{R}^n)},$$

there exists a unique solution to

$$(\Delta - q)u = 0,$$

of the form

$$u(\mathbf{x}) = e^{\mathbf{x}\cdot\boldsymbol{\rho}}\big(1+\psi(\mathbf{x};\boldsymbol{\rho})\big), \tag{7.27}$$

where $\psi(\mathbf{x};\boldsymbol{\rho}) \in H^2_\delta(\mathbb{R}^n)$. *Moreover for* $0 \le s \le 1$, *we have*

$$\|\psi(\mathbf{x};\boldsymbol{\rho})\|_{H^s_\delta(\mathbb{R}^n)} \le \frac{C}{|\boldsymbol{\rho}|^{1-s}}. \tag{7.28}$$

We use here the following Hilbert space

$$L^2_\delta(\mathbb{R}^n) = \Big\{ f; \|f\|^2_{L^2_\delta(\mathbb{R}^n)} = \int_{\mathbb{R}^n} (1+|\mathbf{x}|^2)^\delta |f(\mathbf{x})|^2 d\mathbf{x} < \infty \Big\}, \tag{7.29}$$

and $H^s_\delta(\mathbb{R}^n)$ the associated Sobolev space of order s.

The above theorem shows that complex exponentials are almost in the kernel of $(\Delta - q)$, up to a factor of size $|\boldsymbol{\rho}|^{-1}$ in the $L^2_\delta(\mathbb{R}^n)$ sense.

7.2.3 The potential problem

Let $q \in L^\infty(X)$, which we extend to 0 outside X. Let us assume that the dimension $n \ge 3$. We define the Cauchy data of the potential problem as the set

$$\mathcal{C}_q = \Big\{ \big(u\big|_{\partial X}, \frac{\partial u}{\partial \boldsymbol{\nu}}\Big|_{\partial X}\big) \Big\}, \tag{7.30}$$

where $u \in H^1(X)$ is a solution of $(\Delta - q)u = 0$. Then we have the following uniqueness result

Theorem 7.2.2 *Let us assume that* $\mathcal{C}_{q_1} = \mathcal{C}_{q_2}$. *Then* $q_1 = q_2$.

Proof. For $i = 1, 2$ we define $u_i \in H^1(X)$ as a solution to

$$(\Delta - q_i)u_i = 0.$$

We thus deduce from the divergence theorem that

$$\int_X (q_1 - q_2) u_1 u_2 d\mathbf{x} = \int_{\partial X} \big(\frac{\partial u_1}{\partial \boldsymbol{\nu}} u_2 - \frac{\partial u_2}{\partial \boldsymbol{\nu}} u_1\big) d\sigma = 0, \tag{7.31}$$

since $\mathcal{C}_{q_1} = \mathcal{C}_{q_2}$. We want to show that the product of two arbitrary solutions u_1 and u_2 of the above equations is dense in $L^2(\mathbb{R}^n)$. This would indeed imply then that $q_1 = q_2$ which is what we want to prove. The idea is that products of harmonic solutions are indeed dense in $L^2(\mathbb{R}^n)$ for $n \ge 3$. This is done as follows. We choose $\boldsymbol{\rho}_{1,2}$ as

$$\boldsymbol{\rho}_1 = \frac{\mathbf{m}}{2} + i\frac{\mathbf{k}+\mathbf{l}}{2}, \qquad \boldsymbol{\rho}_2 = -\frac{\mathbf{m}}{2} + i\frac{\mathbf{k}-\mathbf{l}}{2}, \tag{7.32}$$

where the three real-valued vectors $\mathbf{k}$, $\mathbf{l}$, and $\mathbf{m}$ are chosen in $\mathbb{R}^n$ such that

$$\mathbf{m}\cdot\mathbf{k} = \mathbf{m}\cdot\mathbf{l} = \mathbf{k}\cdot\mathbf{l} = 0, \qquad |\mathbf{m}|^2 = |\mathbf{k}|^2 + |\mathbf{l}|^2. \tag{7.33}$$

We verify that $\boldsymbol{\rho}_i \cdot \boldsymbol{\rho}_i = 0$ and that $|\boldsymbol{\rho}_i|^2 = \frac{1}{2}(|\mathbf{k}|^2 + |\mathbf{l}|^2)$. The solutions to $(\Delta - q)u_i$ are then chosen of the form

$$u_i(\mathbf{x}) = e^{\mathbf{x}\cdot\boldsymbol{\rho}_i}(1 + \psi_{q_i}(\mathbf{x})),$$

so that (7.31) can be recast as

$$\widehat{q_1 - q_2}(-\mathbf{k}) = -\int_X e^{i\mathbf{x}\cdot\mathbf{k}}(q_1 - q_2)(\psi_{q_1} + \psi_{q_2} + \psi_{q_1}\psi_{q_2})d\mathbf{x}. \tag{7.34}$$

We now let $|\mathbf{l}| \to \infty$ at fixed value of $\mathbf{k}$ and deduce in the limit that $\widehat{q_1 - q_2}(\mathbf{k}) = 0$ for all $\mathbf{k} \in \mathbb{R}^n$. This concludes the proof of the theorem. □

7.2.4 Inverse conductivity problem

Let us come back to the inversion of γ from Λ_γ. We verify from the change of variables (7.25) that

$$\mathcal{C}_{q_i} = \Big\{ \Big(f, \frac{1}{2}\gamma_i^{-1/2}\frac{\partial \gamma_i}{\partial \boldsymbol{\nu}}\Big|_{\partial X} f + \gamma_i^{-1/2}\Big|_{\partial X} \Lambda_{\gamma_i}\big(\gamma_i^{-1/2}\Big|_{\partial X} f\big)\Big), \quad f \in H^{1/2}(\partial X)\Big\}.$$

Therefore we can conclude that $\Lambda_{\gamma_1} = \Lambda_{\gamma_2}$ implies that $\gamma_1 = \gamma_2$ provided that we can show that γ_1 and γ_2 as well as their normal derivatives agree on ∂X. We have the following result

Theorem 7.2.3 *Let us assume that $0 < \gamma_i \in C^m(\bar{X})$ and that $\Lambda_{\gamma_1} = \Lambda_{\gamma_2}$. Then we can show that for all $|\alpha| < m$, we have*

$$\partial^\alpha \gamma_1\Big|_{\partial X} = \partial^\alpha \gamma_2\Big|_{\partial X}. \tag{7.35}$$

This result implies that when γ_i is real analytic (hence defined by its Taylor expansion at the boundary ∂X), the knowledge of Λ_γ uniquely determines γ [26]. The results obtained in the preceding section show that even in the case of $\gamma \in C^2(\bar{X})$, the knowledge of Λ_γ uniquely determines γ. This fundamental result was obtained in [33].

7.2.5 Stability result

We have seen a uniqueness result in the reconstruction of the diffusion coefficient in an elliptic equation. The stability of the reconstruction is however extremely poor. A result expected to be optimal obtained in [4] shows that

$$\|\gamma(\mathbf{x}) - \gamma'(\mathbf{x})\|_{L^\infty(X)} \leq C\big|\log\|\Lambda - \Lambda'\|_X\big|^{-\delta}. \tag{7.36}$$

Here Λ and Λ' are the maps corresponding to the diffusion equations with coefficients γ and γ', respectively, $X = \mathcal{L}(H^{-1/2}(\partial X), H^{1/2}(\partial X))$, and $\delta \in (0,1)$ is a constant that only depends on the spatial dimension (greater than 3). This formula implies that even very small errors of measurements may have quite large effects on the reconstructed diffusion coefficient.

The reconstruction of the conductivity γ at the boundary of the domain ∂X is actually much more stable. We can show the following result. For γ_1 and γ_2 of class $C^\infty(\bar{X})$ such that $0 < E^{-1} \leq \gamma_i \leq E$, there exists C such that

$$\|\gamma_1 - \gamma_2\|_{L^\infty(\partial X)} \leq C\|\Lambda_{\gamma_1} - \Lambda_{\gamma_2}\|_{\frac{1}{2},-\frac{1}{2}}. \tag{7.37}$$

Moreover for all $0 < \sigma < \frac{1}{n+1}$, there exists C_σ such that

$$\|\frac{\partial \gamma_1}{\partial \boldsymbol{\nu}} - \frac{\partial \gamma_2}{\partial \boldsymbol{\nu}}\|_{L^\infty(\partial X)} \leq C_\sigma \|\Lambda_{\gamma_1} - \Lambda_{\gamma_2}\|^\sigma_{\frac{1}{2},-\frac{1}{2}}. \tag{7.38}$$

The next two chapters are devoted to simplifications of the reconstruction problem: since we cannot expect to reconstruct more than a few coefficients from even almost-noisefree data, we should aim at reconstructing the coefficients we are the most interested in. This requires some sort of parameterization of those parameters, or equivalently to some a priori assumptions on the shape of the coefficients we wish to reconstruct.

Chapter 8

Prior Information and Regularization

As we have mentioned several times in these notes, the influence of "*noise*" is largely subjective. This influence is typically stable in a given norm while it is unstable in another (more constraining) norm. In which norms that influence is controlled for a given problem is the role of stability estimates, which we have presented for all the problems considered in these notes.

Once we have decided that "*noise*" had too large an effect in the setting of interest, something must be done. That something inescapably requires that we add *prior information*. Several techniques have been developed to do so. The simplest and most developed is the *regularization* methodology. Typically, such a regularization assumes that the object we are interested in reconstructing has a prior smoothness. We may for instance assume that the object belongs to $H^s(X)$ for some $s > 0$. This assumption indicates that the object of interest decreases rapidly in the Fourier domain. High frequencies, which are not present, thus do not need to be reconstructed with high accuracy. This allows us to mitigate the effect of high frequency noise.

The main drawback of regularization theory is that objects of interest may not necessarily be smooth. Smooth means that the first coefficients in a Fourier series expansion are big while the other coefficients are small. In several settings of interest, the objects may be represented by a few big coefficients and a large number of small coefficients, but not in the basis of Fourier coefficients. In other words, the object may be *sparse* in a different, known basis. The objective of sparse regularization is to devise methods to find these coefficients.

In some settings, the problems are so ill-posed that looking even for the first coefficients in a given basis may not provide sufficient accuracy. Other sorts of prior information may be necessary, for instance assuming that the objects of interest are small inclusions with specific structures. In such settings, reconstructions are typically not very accurate and it is often important to characterize this inaccuracy. A very versatile setting to do so is the statistical Bayesian framework. In such a setting, objects of interest are modeled by a set of possible outcomes with prior probabilities of happening. This is the *prior probability distribution*. Then data are acquired with a given noise model. The probability of such data happening conditioned on given parameters is called the *likelihood probability distribution*. Using the Bayes rule, a *posterior probability*

distribution gives the probability density of the parameters based on availability of the data.

We now briefly consider these three settings, the smoothness regularization methodology, the sparsity regularization methodology, and the Bayesian framework and show their relations.

8.1 Smoothness Regularization

We have seen that many of the inverse problems we have considered so far are ill-posed, either mildly ill-posed or severely ill-posed. We present here some smoothness techniques to regularize them. Such techniques typically work for mildly ill-posed problems but are not sufficient for severely ill-posed problems.

8.1.1 Ill-posed problems and compact operators

Let A be an injective and compact operator defined on an infinite dimensional Hilbert space H with range $\text{Range}(A)$ in H:

$$A : H \rightarrow \text{Range}(A) \subset H. \tag{8.1}$$

We recall that compact operators map the unit ball in H to a subset of H whose closure (with respect to the usual norm in H) is compact, i.e., verifies that every bounded (with respect to the usual norm in H) family of points admits a converging (with respect to the usual norm in H) subsequence in the compact set.

Since A is injective (i.e., $Ax = 0$ implies $x = 0$), we can define the inverse operator A^{-1} with domain of definition $\text{Range}(A)$ and Range H:

$$A^{-1} : D(A^{-1}) = \text{Range}(A) \rightarrow H. \tag{8.2}$$

The problem is that A^{-1} is *never* a continuous operator from $\text{Range}(A)$ to H when both spaces are equipped with the usual norm in H:

Lemma 8.1.1 *For A as above, there exists a sequence x_n such that*

$$\|x_n\|_H = 1, \qquad \|Ax_n\|_H \rightarrow 0. \tag{8.3}$$

The same holds true with $\|x_n\|_H \rightarrow \infty$.

Proof. The proof holds in more complicated settings than Hilbert spaces. The Hilbert structure gives us a very simple proof and is based on the existence of an orthonormal basis in H, i.e., vectors x_n such that $\|x_n\|_H = 1$ and $(x_n, x_m)_H = 0$ for $n \neq m$. Since these vectors belong to the unit ball, we deduce that $y_n = Ax_n$ is a converging sequence (up to taking subsequences), say to $y \in H$. Take now $\tilde{x}_n = 2^{-1/2}(x_n - x_{n+1})$. We verify that $\tilde{x}_n$ satisfies (8.3). Now define $z_n = \tilde{x}_n / \|A\tilde{x}_n\|_H^{1/2}$ when the latter denominator does not vanish and $z_n = n\tilde{x}_n$ otherwise. Then Az_n still converges to 0 while $\|z_n\|_H$ converges to ∞. □

This simple lemma shows that inverting a compact operator can never be a well-posed problem in the sense that A^{-1} is not continuous from $D(A^{-1})$ to H with the H norm.

Indeed take the sequence $y_n = Ax_n/\|Ax_n\|$ in $D(A^{-1})$, where x_n is the sequence in (8.3). Then $\|y_n\|_H = 1$ while $\|A^{-1}y_n\|_H$ tends to ∞.

The implication for our inverse problem is the following. If δy_n is our measurement noise for n large, then $\delta A^{-1}y_n$ will be the error in our reconstruction, which may thus be arbitrarily larger than the norm of the true object we aim to reconstruct. More precisely, if $Ax = b$ is the real problem and $A\tilde{x} = \tilde{b}$ is the exact reconstruction from noisy data, then arbitrarily small errors $\|b - \tilde{b}\|$ in the measurements is still compatible with arbitrary large errors $\|x - \tilde{x}\|$ in the space of objects to reconstruct. This shows that the problem needs to be regularized before any inversion is carried out.

8.1.2 Regularity assumptions and error bound

You should be convinced by the calculations we have carried out in the preceding section that an ill-posed inverse problem cannot satisfactorily be solved if no other *assumptions* on the problem are added. A sometimes reasonable and practically useful assumption is to impose, before we start the reconstruction process, that the object we want to reconstruct is itself *sufficiently smooth*. This allows us to filter out high frequencies that may appear in the reconstruction because we know they are part of the noise and not of the object we want to reconstruct. We present two types of theories that put this idea into a more mathematical framework.

In the first framework, we first introduce the adjoint operator A^* to A, defined from H to Range(A^*) by the relation

$$(Ax, y)_H = (x, A^*y)_H, \qquad \text{for all } x, y \in H.$$

Since A is compact and injective, then so is A^*. We can also define the inverse operator $A^{-*} = (A^*)^{-1}$ from Range(A^*) to H.

We may now assume that x, the object we aim at reconstructing, is sufficiently smooth that is belongs to the range of A^*, i.e., there exists y such that $x = A^*y$. Since A and A^* are compact operators, hence smoothing operators, the above hypothesis means that we assume *a priori* that x is smoother than being merely an element in H. We then define the *stronger* norm

$$\|x\|_1 = \|A^{-*}x\|_H. \tag{8.4}$$

We may also assume that the object x is even smoother than being in the range of A^*. For instance let us assume that x belongs to the range of A^*A, i.e., there exists y such that $x = A^*Ay$. Note that since both A and A^* are smoothing operators (because they are compact), the assumption on x is stronger than simply being in the range of A^*. We define the even stronger norm

$$\|x\|_2 = \|(A^*A)^{-1}x\|_H. \tag{8.5}$$

We want to use these definitions to show that if the solution x is *a priori* bounded for the $\|\cdot\|_1$ or the $\|\cdot\|_2$ norm and "*noise*" is small, then the error in the reconstruction is small. For instance, assume that $y_j = Ax_j$ for $j = 1, 2$ so that $y = Ax$ for $y = y_1 - y_2$ and $x = x_1 - x_2$. If both x_j, $j = 1, 2$ are bounded and y is small, then how small is x? For such questions, we have the following result:

Theorem 8.1.2 *Let $x \in H$ such that $\|x\|_1 \leq E$ and $\|Ax\|_H \leq \delta$. Then we have:*

$$\|x\|_H \leq \sqrt{E\delta}. \tag{8.6}$$

If we now assume that $\|x\|_2 \leq E$ instead, we obtain the better bound

$$\|x\|_H \leq E^{1/3}\delta^{2/3}. \tag{8.7}$$

Proof. Let $y = A^{-*}x$ so that $\|y\|_H \leq E$. We have then

$$\|x\|_H^2 = (x, A^*y) = (Ax, y) \leq \|Ax\|_H\|y\|_H \leq \delta E.$$

This proves (8.6). For the second bound let $z = (A^*A)^{-1}x$ so that $\|z\|_H \leq E$ and compute:

$$\|x\|_H^2 = (x, A^*Az) = (Ax, Az) \leq \delta\|Az\| = \delta(Az, Az)^{\frac{1}{2}} \leq \delta(z, x)^{\frac{1}{2}} \leq \delta E^{\frac{1}{2}}\|x\|_H^{\frac{1}{2}}.$$

This proves the second bound (8.7). □

The theorem should be interpreted as follows. Consider that Ax is the noise level in the measured data and that $\|x\|_1 < E$ or $\|x\|_2 < E$ is *a priori* smoothness information we have on the object we want to reconstruct. Then the worst error we can make on the reconstruction (provided we find an appropriate inversion method; see below) is given by the bounds (8.6) and (8.7). Notice that the latter bound is better (since $\delta^{2/3} \ll \delta^{1/2}$). This results from a more stringent assumption on the image x.

Let us now consider smoothing operators in the framework of the Hilbert scale $H^s(\mathbb{R})$ we have introduced in Chapter 1. Then we have the following result

Theorem 8.1.3 *Let us assume that the operator A is mildly ill-posed of order $\alpha > 0$ so that*

$$\|Af\|_{L^2(\mathbb{R})} \geq m\|f\|_{H^{-\alpha}(\mathbb{R})}. \tag{8.8}$$

Suppose now that the measurement error is small and that the function we want to reconstruct is regular in the sense that

$$\|Af\|_{L^2(\mathbb{R})} \leq \delta m, \qquad \text{and} \qquad \|f\|_{H^\beta(\mathbb{R})} \leq E, \tag{8.9}$$

for some $\delta > 0$, $\beta > 0$, and $E > 0$. Then we have

$$\|f\|_{L^2(\mathbb{R})} \leq \delta^{\frac{\beta}{\alpha+\beta}} E^{\frac{\alpha}{\alpha+\beta}}. \tag{8.10}$$

Proof. The proof is a simple but interesting exercise in interpolation theory. Notice that the hypotheses are

$$\|f\|_{H^\beta(\mathbb{R})} \leq E, \qquad \text{and} \qquad \|f\|_{H^{-\alpha}(\mathbb{R})} \leq \delta,$$

and our objective is to find a bound for $\|f\|_{L^2(\mathbb{R})}$. Let us denote $\langle\boldsymbol{\xi}\rangle = (1 + |\boldsymbol{\xi}|^2)^{1/2}$. We have

$$\begin{aligned}(2\pi)^n\|f\|_{L^2(\mathbb{R})}^2 &= \int_{\mathbb{R}^n} |\hat{f}(\boldsymbol{\xi})|^{2\theta}\langle\boldsymbol{\xi}\rangle^{2\gamma}|\hat{f}(\boldsymbol{\xi})|^{2(1-\theta)}\langle\boldsymbol{\xi}\rangle^{-2\gamma}d\boldsymbol{\xi} \\ &\leq \Big(\int_{\mathbb{R}^n} |\hat{f}(\boldsymbol{\xi})|^2\langle\boldsymbol{\xi}\rangle^{2\gamma/\theta}d\boldsymbol{\xi}\Big)^\theta \Big(\int_{\mathbb{R}^n} |\hat{f}(\boldsymbol{\xi})|^2\langle\boldsymbol{\xi}\rangle^{-2\gamma/(1-\theta)}d\boldsymbol{\xi}\Big)^{1-\theta},\end{aligned}$$

thanks to Hölder's inequality

$$\Big| \int_{\mathbb{R}} f(\mathbf{x}) g(\mathbf{x}) d\mathbf{x} \Big| \leq \|f\|_{L^p(\mathbb{R})} \|g\|_{L^q(\mathbb{R})},$$

which holds for all $p \geq 1$ and $q \geq 1$ such that $p^{-1} + q^{-1} = 1$, where we have defined for all $p \geq 1$,

$$\|f\|_{L^p(\mathbb{R})} = \Big(\int_{\mathbb{R}} |f(\mathbf{x})|^p d\mathbf{x} \Big)^{1/p}. \tag{8.11}$$

Choosing $\theta = \frac{\alpha}{\alpha+\beta}$ and $\gamma = \frac{\alpha\beta}{\alpha+\beta}$ gives (8.10).

Let us briefly recall the proof of the Hölder's inequality [34]. We first verify that

$$x^{1/p} \leq \frac{x}{p} + \frac{1}{q}, \qquad x > 0,$$

for $p^{-1} + q^{-1} = 1$ and $p \geq 1$, since $x^{1/p} - \frac{x}{p}$ attains its maximum at $x = 1$ where it is equal to q^{-1}. For $y > 0$ we use the above inequality for x/y and multiply by y to obtain

$$x^{1/p} y^{1/q} \leq \frac{x}{p} + \frac{y}{q}, \qquad x > 0,\ y > 0. \tag{8.12}$$

Choosing $x = |tf(\mathbf{x})|^p$ and $y = |t^{-1} g(\mathbf{x})|^q$, we deduce that

$$\int_{\mathbb{R}^n} |f(\mathbf{x}) g(\mathbf{x})| d\mathbf{x} \leq \frac{1}{p} \|tf\|^p_{L^p(\mathbb{R})} + \frac{1}{q} \|t^{-1} g\|^q_{L^q(\mathbb{R})} = \frac{t^p}{p} \|f\|^p_{L^p(\mathbb{R})} + \frac{t^q}{q} \|g\|^q_{L^q(\mathbb{R})},$$

for all $t > 0$. Maximizing over t gives the Hölder inequality. □

The last theorem applies to a less general class of operators than compact operators (although it applies to operators that are not necessarily compact) but it gives us more accurate results. We should still consider δ as the noise level and E as an a priori bound we have on the object we want to reconstruct. Then depending on the a priori smoothness of the object, we obtain different possible accuracies in the reconstructions. What is important is the relative regularity of the object compared to the smoothing effect of the operator A. When $\beta = \alpha$, this corresponds to assuming the same regularity as $\|x\|_1 \leq E$ in Theorem 8.1.2. We thus obtain an accuracy of order $\delta^{1/2}$ in the reconstruction. When $\beta = 2\alpha$, this corresponds to $\|x\|_2 \leq E$ since $f = (A^*A)^{-1} g$ for some $g \in L^2(\mathbb{R})$ means that f is twice as regular as A is smoothing. We thus recover the accuracy of order $\delta^{2/3}$ as in Theorem 8.1.2. Theorem 8.1.3 allows us to deal with arbitrary values of β. Notice that as $\beta \to \infty$, we recover that the problem is almost well-posed since the error in the reconstruction is asymptotically of the same order δ as the noise level.

8.1.3 Regularization methods

Now that we know how noise can optimally be controlled in the reconstruction based on the regularity of the object we want to reconstruct, we need to devise algorithms that indeed control noise amplification in the reconstruction.

Since A^{-1} is an unbounded operator with domain of definition Range(A), a proper subset of H, we first need to introduce approximations of the inverse operator. We

denote by R_γ defined from H to H for $\gamma > 0$ a sequence of regularizations of A^{-1} such that

$$\lim_{\gamma \to 0} R_\gamma A x = x \qquad \text{for all } x \in H. \tag{8.13}$$

Under the hypotheses of Lemma 8.1.1, we can show that the sequence of operators R_γ is not uniformly bounded.

Exercise 8.1.1 Prove this.

A uniform bound would indeed imply that A^{-1} is bounded. Thus, $R_\gamma A$ converges to identity *strongly* (since (8.13) is the definition of strong convergence of operators) but not uniformly in the sense that $\|R_\gamma A - I\|$ does not converge to 0.

One of the main objectives of the regularization technique is to handle noise in an optimal fashion. Let us denote by y^δ our measurements and assume that $\|y^\delta - Ax\|_H \leq \delta$. We then define

$$x^{\gamma,\delta} = R_\gamma y^\delta. \tag{8.14}$$

We want to find sequences R_γ that deal with noise in an optimal fashion. For instance assuming that $\|x\|_1 \leq E$ and that $\|y^\delta - Ax\|_H \leq \delta$, we want to be able to show that

$$\|x - x^{\gamma,\delta}\| \leq C\sqrt{E\delta},$$

at least for some values of γ. We know from Theorem 8.1.3 that such a bound is optimal. We will consider three regularization techniques: singular value decomposition, Tikhonov regularization, and Landweber iterations.

The choice of a parameter γ is then obviously of crucial importance as the above bound will not hold independently of γ. More precisely, the reconstruction error can be decomposed as

$$\|x^{\gamma,\delta} - x\|_H \leq \delta \|R_\gamma\|_H + \|R_\gamma A x - x\|_H. \tag{8.15}$$

Exercise 8.1.2 Prove this. The operator norm $\|R_\gamma\|_H$ is defined as the supremum of $\|R_\gamma x\|_H$ under the constraint $\|x\|_H \leq 1$.

We thus observe that two competing effects enter (8.15). The first effect is the ill-posedness effect: as $\gamma \to 0$, the norm $\|R_\gamma\|_H$ tends to ∞ so γ should not be chosen too small. The second effect is the regularization effect: as γ increases, $R_\gamma A$ becomes a less accurate approximation of identity so γ should not be chosen too large. Only intermediate values of γ will provide an optimal reconstruction.

Singular Value Decomposition

For a compact and injective operator A defined on an infinite dimensional Hilbert space H, let us assume that we know it singular value decomposition defined as follows. Let A^* be the adjoint operator to A and $\lambda_j > 0$, $j \in \mathbb{N}$ the eigenvalues of the symmetric operator A^*A. Then the sequence $\mu_j = \sqrt{\lambda_j}$ for $j \in \mathbb{N}$ are called the *singular values* of A. Since $\mu_j \leq \|A\|_H$, we order the singular values such that

$$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n \geq \cdots > 0.$$

Multiple eigenvalues are repeated as many times as their multiplicity (which is necessarily finite since the associated eigenspace for A^*A needs to be compact).

There exists then two orthonormal systems $(x_j)_{j\in\mathbb{N}}$ and $(y_j)_{j\in\mathbb{N}}$ in H such that

$$Ax_j = \mu_j y_j \qquad \text{and} \qquad A^*y_j = \mu_j x_j, \qquad \text{for all} \quad j \in J. \tag{8.16}$$

We call (μ_j, x_j, y_j) the singular system for A. Notice that

$$Ax = \sum_{j=1}^{\infty} \mu_j (x, x_j) y_j, \qquad A^*y = \sum_{j=1}^{\infty} \mu_j (y, y_j) x_j.$$

Here (x, x_j) is the inner product in H, $(x, x_j)_H$. We have then the very useful characterization of the Range of the compact and injective operator A:

Lemma 8.1.4 (Picard) *The equation $Ax = y$ is solvable in H if and only if*

$$\sum_{j\in\mathbb{N}} \frac{1}{\mu_j^2} |(y, y_j)|^2 < \infty, \tag{8.17}$$

in which case the solution is given by

$$x = A^{-1}y = \sum_{j\in\mathbb{N}} \frac{1}{\mu_j} (y, y_j) x_j. \tag{8.18}$$

The ill-posedness of the inverse problem appears very clearly in the singular value decomposition. As $j \to \infty$, the singular values μ_j tend to 0. And they do so all the faster that the inverse problem is ill-posed. We can extend the definition of ill-posed problems in the sense that a compact operator generates a mildly ill-posed inverse problem of order $\alpha > 0$ when the singular values decay like $j^{-\alpha}$ and generates a severely ill-posed problem when the singular values decay faster than any j^{-m} for $m \in \mathbb{N}$.

So in order to regularize the problem, all we have to do is to replace too small singular values by larger values. Let us define $q(\gamma, \mu)$ for $\gamma > 0$ and $\mu \in [0, \|A\|]$ such that

$$|q(\gamma, \mu)| < 1, \quad |q(\gamma, \mu)| \le c(\gamma)\mu, \quad \text{and} \quad q(\gamma, \mu) - 1 \to 0 \quad \text{as} \quad \gamma \to 0, \tag{8.19}$$

(not uniformly in μ obviously). Then we define the regularizing sequence

$$R_\gamma y = \sum_{j\in\mathbb{N}} \frac{q(\gamma, \mu_j)}{\mu_j} (y, y_j) x_j. \tag{8.20}$$

Compare to (8.18). As $\gamma \to 0$, R_γ converges to A^{-1} pointwise. We are interested in estimating (8.15) and showing that the error is optimal based on the assumed regularity of x. The total error is estimated by using

$$\|R_\gamma\|_H \le c(\gamma), \qquad \|R_\gamma A x - x\|_H = \sum_{j=1}^{\infty} \big(q(\gamma, \mu_j) - 1\big)^2 |(x, x_j)|^2. \tag{8.21}$$

Exercise 8.1.3 Prove these relations.

We can now prove the following results

Theorem 8.1.5 *(i) Let us assume that $x = A^*z$ with $\|z\|_H \leq E$ and that $\|y^\delta - Ax\| \leq \delta$, where y^δ is the measurements. Choose $q(\gamma, \mu)$ and γ such that*

$$|q(\gamma,\mu) - 1| \leq C_1 \frac{\sqrt{\gamma}}{\mu}, \qquad c(\gamma) \leq \frac{C_2}{\sqrt{\gamma}}, \qquad \gamma = \frac{C_3\delta}{E}. \tag{8.22}$$

Then we have that

$$\|x^{\gamma,\delta} - x\|_H \leq \Big(\frac{C_2}{\sqrt{C_3}} + C_1\sqrt{C_3}\Big)\sqrt{\delta E}. \tag{8.23}$$

(ii) Let us assume that $x = A^*Az$ with $\|z\|_H \leq E$ and that $\|y^\delta - Ax\| \leq \delta$, where y^δ is the measurements. Choose $q(\gamma, \mu)$ and γ such that

$$|q(\gamma,\mu) - 1| \leq C_4 \frac{\gamma}{\mu^2}, \qquad c(\gamma) \leq \frac{C_5}{\sqrt{\gamma}}, \qquad \gamma = C_6\Big(\frac{\delta}{E}\Big)^{2/3}. \tag{8.24}$$

Then we have that

$$\|x^{\gamma,\delta} - x\|_H \leq \Big(\frac{C_5}{\sqrt{C_6}} + C_4C_6\Big)\delta^{2/3}E^{1/3}. \tag{8.25}$$

Proof. Since $x = A^*z$, we verify that $(x, x_j) = \mu_j(y, y_j)$ so that

$$\|R_\gamma Ax - x\|_H^2 = \sum_{j=1}^{\infty} \big(q(\gamma,\mu_j) - 1\big)^2 |(z, y_j)|^2 \leq C_1^2\gamma\|z\|_H^2.$$

This implies that

$$\delta\|R_\gamma\|_H + \|R_\gamma Ax - x\|_H \leq \frac{C_2\delta}{\sqrt{\gamma}} + C_1\sqrt{\gamma}E.$$

Using (8.15) and the expression for γ yields (8.23).

Exercise 8.1.4 Using similar arguments, prove (8.25).

This concludes the proof. □

We have thus defined an optimal regularization scheme for the inversion of $Ax = y$. Indeed from the theory in Theorem 8.1.2 we know that up to some multiplicative constants, the above estimates are optimal.

It remains to find filters $q(\gamma, \mu)$ satisfying the above hypotheses. We propose two:

$$q(\gamma,\mu) = \frac{\mu^2}{\gamma + \mu^2}, \tag{8.26}$$

$$q(\gamma,\mu) = \begin{cases} 1, & \mu^2 \geq \gamma, \\ 0, & \mu^2 < \gamma. \end{cases} \tag{8.27}$$

Exercise 8.1.5 Show that the above choices verify the hypotheses of Theorem 8.1.5.

Tikhonov Regularization

One of the main drawbacks of the theory presented in the preceding section is that in most cases, the singular value decomposition of the operator is not analytically available (although it is for the Radon transform; see [27, 28]), and is quite expensive to compute numerically once the continuous problem has been discretized. It is therefore useful to consider regularization techniques that do not depend on the SVD. One of the most famous regularization techniques is the Tikhonov-Phillips regularization technique.

Solving $Ax = y$ corresponds to minimizing $\|Ax - y\|_H$. Instead one may want to minimize the regularized Tikhonov functional

$$J_\gamma(x) = \|Ax - y\|_H^2 + \gamma\|x\|_H^2, \qquad x \in H. \tag{8.28}$$

For $\gamma > 0$ and A a linear bounded operator on H, we can show that the above functional admits a unique minimizer x^γ solving the following normal equations

$$A^*Ax^\gamma + \gamma x^\gamma = A^*y. \tag{8.29}$$

We can thus define the regularizing sequence

$$R_\gamma = (\gamma + A^*A)^{-1}A^*. \tag{8.30}$$

The operator is bounded in H by $\|R_\gamma\|_H \leq C\gamma^{-1/2}$ for all $\gamma > 0$. Notice that for a compact operator A with singular system (μ_i, x_i, y_i), we verify that the singular value decomposition of R_γ is

$$R_\gamma y = \sum_{j=1}^{\infty} \frac{\mu_j}{\gamma + \mu_j^2}(y, y_j)x_j. \tag{8.31}$$

This means that the Tikhonov regularization corresponds to the SVD regularization with filter given by (8.26) and implies that the Tikhonov regularization is optimal to inverse problem with a priori regularity $\|x\|_1 \leq E$ or $\|x\|_2 \leq E$. It is interesting to observe that the Tikhonov regularization is no longer optimal when the a priori regularity of x is better than $\|x\|_2 \leq E$ (see [25]).

Let us make this observation more explicit. Let us consider the operator A given in the Fourier domain by multiplication by $\langle \boldsymbol{\xi} \rangle^{-\alpha}$ for some $\alpha > 0$. We verify that $A^* = A$ and that R_γ is given in the Fourier domain by

$$R_\gamma = \mathcal{F}^{-1}_{\boldsymbol{\xi}\to\mathbf{x}} \frac{\langle \boldsymbol{\xi} \rangle^{-\alpha}}{\langle \boldsymbol{\xi} \rangle^{-2\alpha} + \gamma} \mathcal{F}_{\mathbf{x}\to\boldsymbol{\xi}}, \qquad \text{so that} \qquad \|R_\gamma\| \leq \frac{1}{2\sqrt{\gamma}}.$$

Indeed, we check that $x/(x^2 + \gamma) \leq 1/(2\sqrt{\gamma})$ and attains its maximum at $x = \sqrt{\gamma}$. We now verify that

$$I - R_\gamma A = \mathcal{F}^{-1}_{\boldsymbol{\xi}\to\mathbf{x}} \frac{\gamma}{\langle \boldsymbol{\xi} \rangle^{-2\alpha} + \gamma} \mathcal{F}_{\mathbf{x}\to\boldsymbol{\xi}},$$

so that for a function $f \in H^\beta(\mathbb{R}^n)$, we have

$$\|f - R_\gamma A f\| \leq \sup_{\langle \boldsymbol{\xi} \rangle \geq 1} \frac{\gamma \langle \boldsymbol{\xi} \rangle^{-\beta}}{\langle \boldsymbol{\xi} \rangle^{-2\alpha} + \gamma} \|f\|_{H^\beta(\mathbb{R}^n)}.$$

Moreover the inequality is sharp in the sense that there exists functions f such that the reverse inequality holds (up to a multiplicative constant independent of γ; Check this). For $\beta > 2\alpha$, the best estimate we can have for the above multiplier is that it is of order $O(\gamma)$ (choose for instance $\langle \boldsymbol{\xi} \rangle = 1$).

Exercise 8.1.6 Using (8.12) show that

$$\frac{\gamma \langle \boldsymbol{\xi} \rangle^{-2\alpha\theta}}{\langle \boldsymbol{\xi} \rangle^{-2\alpha} + \gamma} \leq \gamma^{\theta}, \qquad 0 \leq \theta \leq 1.$$

Show that the above inequality is sharp.

Let $Af = g$ be the problem we want to solve and g^{δ} the measurements so that $\|Af - g^{\delta}\|_{L^2(\mathbb{R}^n)} \leq \delta$. Let us assume that f belongs to $H^{\beta}(\mathbb{R}^n)$. We verify using (8.15) that the error of the regularized problem is given by

$$\|f - R_{\gamma} g^{\delta}\| \leq \frac{\delta}{2\sqrt{\gamma}} + \gamma^{\frac{\beta}{2\alpha} \wedge 1} \|f\|_{H^{\beta}(\mathbb{R}^n)}. \tag{8.32}$$

Here, $a \wedge b = \min(a, b)$. This implies that

$$\|f - R_{\gamma} g^{\delta}\| \leq C \delta^{\frac{\beta}{\alpha+\beta} \wedge \frac{2}{3}} \|f\|^{\frac{\alpha}{\alpha+\beta} \wedge \frac{1}{3}}_{H^{\beta}(\mathbb{R}^n)}, \tag{8.33}$$

for a universal constant C. We therefore obtain that the Tikhonov regularization is optimal according to Theorem 8.1.3 when $0 < \beta \leq 2\alpha$. However, for all $\beta > 2\alpha$, the error between the Tikhonov regularization and the exact solution will be of order $\delta^{2/3}$ instead of $\delta^{\frac{\beta}{\beta+\alpha}}$.

Exercise 8.1.7 More generally, consider an operator A with symbol $a(\boldsymbol{\xi})$, i.e.,

$$A = \mathcal{F}^{-1}_{\boldsymbol{\xi} \to \mathbf{x}} a(\boldsymbol{\xi}) \mathcal{F}_{\mathbf{x} \to \boldsymbol{\xi}},$$

such that $0 < a(\boldsymbol{\xi}) \in C^{\infty}(\mathbb{R}^n)$ and for some $\alpha > 0$ and $a_{\infty} \neq 0$,

$$\frac{a(\boldsymbol{\xi})}{\langle \boldsymbol{\xi} \rangle^{\alpha}} \to a_{\infty}, \qquad \text{as} \qquad |\boldsymbol{\xi}| \to \infty. \tag{8.34}$$

(i) Show that A^*, the adjoint of A for the $L^2(\mathbb{R}^n)$ inner product, satisfies the same hypothesis (8.34).
(ii) Show that R_{γ} and $S_{\gamma} = R_{\gamma} A - I$ are bounded operator with symbols given by

$$r_{\gamma}(\boldsymbol{\xi}) = (|a(\boldsymbol{\xi})|^2 + \gamma)^{-1} \bar{a}(\boldsymbol{\xi}), \qquad s_{\gamma}(\boldsymbol{\xi}) = \gamma (|a(\boldsymbol{\xi})|^2 + \gamma)^{-1},$$

respectively.
(iii) Assuming that $f \in H^{\beta}(\mathbb{R})$, show that (8.33) holds.

These results show that for the Radon transform, an a priori regularity of the function $f(\mathbf{x})$ in $H^1(\mathbb{R}^2)$ is sufficient to obtain an error of order $\delta^{2/3}$. When the function is smoother, a different technique from Tikhonov regularization is necessary to get a more accurate reconstruction.

Landweber iterations

The drawback of the Tikhonov regularization is that it requires to invert the regularization of the normal operator $\gamma + A^*A$. This inversion may be very costly in practice. The Landweber iteration method is an iterative technique in which no inversion is necessary. It is defined to solve the equation $Ax = y$ as follows

$$x_0 = 0, \qquad x_{n+1} = (I - rA^*A)x_n + rA^*y, \qquad n \geq 0, \tag{8.35}$$

for some $r > 0$. By induction, we verify that $x_n = R_n y$, where

$$R_n = r\sum_{k=0}^{n-1}(I - rA^*A)^k A^*, \qquad n \geq 1. \tag{8.36}$$

Consider a compact operator A with singular system (μ_j, x_j, y_j). We thus verify that

$$R_n y = \sum_{j=1}^{\infty} \frac{1}{\mu_j}\big(1 - (1 - r\mu_j^2)^n\big)(y, y_j)x_j. \tag{8.37}$$

Exercise 8.1.8 Check (8.37).

This implies that R_n is of the form R_γ in (8.20) with $\gamma = n^{-1}$ and

$$q(\gamma, \mu) = 1 - (1 - r\mu^2)^{1/\gamma}.$$

Exercise 8.1.9 *Show that the above filter verifies the hypotheses (8.19) and those of Theorem 8.1.5.*

This implies that the Landweber iteration method is an optimal inversion method by Theorem 8.1.5.

Exercise 8.1.10 Show that the hypotheses of Theorem 8.1.5 are met provided that the number of iterations n is chosen as

$$n = c\frac{E}{\delta}, \qquad n = c\Big(\frac{E}{\delta}\Big)^{2/3},$$

when $\|x\|_1 \leq E$ and $\|x\|_2 \leq E$, respectively.

The above result shows that the number of iterations should be chosen carefully: when n is too small, then $R_\gamma A$ is a poor approximation of I, and when n is too large, then $\|R_\gamma\|_H$ is too large. Unlike the Tikhonov regularization, we can show that the Landweber iteration method is also optimal for stronger regularity assumptions on x than those given in Theorem 8.1.5 (see [25] for instance).

Let us come back to the operator A with symbol $a(\boldsymbol{\xi}) = \langle\boldsymbol{\xi}\rangle^{-\alpha}$. We verify that R_n and $S_n = R_nA - I$ have respective symbols

$$r_n(\boldsymbol{\xi}) = \frac{1 - (1 - r\langle\boldsymbol{\xi}\rangle^{-2\alpha})^n}{\langle\boldsymbol{\xi}\rangle^{-\alpha}}, \qquad s_n(\boldsymbol{\xi}) = -(1 - r\langle\boldsymbol{\xi}\rangle^{-2\alpha})^n.$$

Exercise 8.1.11 (i) Show that $s_n(\mathbf{x})\langle\boldsymbol{\xi}\rangle^{-\beta}$ is bounded by $Cn^{-\beta/(2\alpha)}$ for $\langle\boldsymbol{\xi}\rangle$ of order $n^{1/(2\alpha)}$. Deduce that for $f \in H^\beta(\mathbb{R}^n)$, we have

$$\|S_n f\| \leq Cn^{-\frac{\beta}{2\alpha}}\|f\|_{H^\beta(\mathbb{R}^n)}.$$

(ii) Show that provided that n is chosen as

$$n = C\delta^{\frac{-2\alpha}{\alpha+\beta}}\|f\|_{H^\beta(\mathbb{R}^n)}^{\frac{2\alpha}{\alpha+\beta}},$$

we have the estimate

$$\delta\|R_n\| + \|S_n f\| \leq C\delta^{\frac{\beta}{\alpha+\beta}}\|f\|_{H^\beta(\mathbb{R}^n)}^{\frac{\alpha}{\alpha+\beta}}. \tag{8.38}$$

(iii) Deduce that the Landweber iteration method is an optimal regularization technique for all $\beta > 0$.
(iv) Generalize the above results for the operators described in Exercise 8.1.7.

We have thus the striking result that unlike the Tikhonov regularization method described in (8.33), the Landweber iteration regularization can be made optimal (by choosing the appropriate number of iterations n) for all choices on the regularity in $H^\beta(\mathbb{R}^n)$ of the object f.

Let us conclude this section by the following summarizing remark. The reason why regularization was necessary was because *you as a user* decided that noise was too amplified during not regularized inversions. Smoothness priors were then considered to restore well-posedness. This corresponds to a choice of the factor β and a bound E. *In addition*, we need to choose a *regularization* parameter, γ in the Tikhonov regularization algorithm and the stopping iteration n in the Landweber iteration. How these choices are made depends on the bound E but also on the *estimated error* δ. Various techniques have been developed to choose γ or n a posteriori (Morozov principle, L-curve). All these techniques require that δ be known. There is no free lunch. Regularization does require prior information about the solution to mitigate the perceived lack of information in the available data.

8.2 Sparsity and other Regularization Priors

The regularization methods considered in the preceding section have a huge advantage: the replace ill-posed linear systems of equations by well-posed, at least better-posed (better-conditioned), *linear* systems as well. For instance, the inversion of A has been replaced by that of $A^*A + \gamma I$ in the simplest version of Tikhonov regularization. Their main disadvantage is that they render the regularized solution typically smoother than the "exact" solution. Such a smoothing is unavoidable with such regularizations.

The reason why smooth objects are well reconstructed by the smoothing regularization method is that such objects can be represented by a small number of large coefficients (e.g., the first Fourier modes in a Fourier series expansion). It turns out that some objects are better represented in other bases. For instance, an image tends to have sharp discontinuities, for instance between a bright area and a dark area. Some bases, such as for instance those based on wavelets, will be much better than Fourier bases to represent this type of information.

A general framework to account for such prior information is to recast the inverse problem as seeking the minimum of an appropriate functional. Solving the inverse problem then amounts to solving a *optimization problem*. Let us assume that we have a problem of the form

$$\mathfrak{M}(u) = v,$$

and assume that v_d are given data. Let us assume that a functional $u \mapsto \mathcal{R}(u)$ incorporates the prior information about u in the sense that $\mathcal{R}(u)$ is small when u satisfies the constraints. Let us assume also that $\rho(u, v)$ is a function that we wish to use to quantify the error between the available data v_d and the forward model $\mathfrak{M}(u)$. Then we want to minimize $\rho(v, v_d)$ and at the same time minimize $\mathcal{R}(u)$. Both constraints can be achieved by introducing a *regularization parameter* α and minimizing the sum

$$\mathcal{F}_\alpha(u) = \rho\big(\mathfrak{M}(u), v_d\big) + \alpha\mathcal{R}(u). \tag{8.39}$$

Solving the inverse problem consists of minimizing the above functional to get

$$\hat{u}_\alpha = \ \text{argmin}\ \mathcal{F}_\alpha(u). \tag{8.40}$$

These minimization problems often go by the name of Tikhonov regularization and may be seen as generalizations of (8.28).

The main objective of regularization (or sparsity) theory is then to devise functions ρ, $\mathcal{R}$ and a regularization parameter α, that best fits our prior information about the problem of interest. Once such a problem has been formulated, it remains to devise a method to solve such an optimization problem numerically. There is a vast literature on the subject; for recent references see the two excellent books [31, 32].

8.2.1 Smoothness Prior and Minimizations

The simplest class of regularizations consists of choosing $\rho(u, v) = \frac{1}{2}\|D(u - v)\|_H^2$ in the $H = L^2$ sense for an operator D that may be identity or an operator of differentiation if small errors on derivatives of the solution matter in practice, and choosing $\mathcal{R}(u)$ also as a quadratic functional, for instance $\mathcal{R}(u) = \frac{1}{2}\|Ru\|_H^2$, again for an operator R that may be identity or a differential operator. Then associated to the linear problem $Au = v$, we have the minimization problem:

$$\mathcal{F}_\alpha(u) = \frac{1}{2}\|D(Au - v_d)\|_H^2 + \frac{\alpha}{2}\|Ru\|_H^2. \tag{8.41}$$

The main advantage of the above quadratic expression is that the Euler-Lagrange equations associated to the above minimization problem is the following linear system of equations

$$\big((DA)^*(DA) + \alpha R^*R\big)u = (DA)^*Dv_d. \tag{8.42}$$

When $D = R = I$, this is nothing but (8.29). Provided that R is an invertible matrix, then the above problem can be solved for all $\alpha > 0$. We have seen in the preceding section how the method converged (at least when $D = R = I$) as the noise in the data and the regularization parameter α tend to 0.

8.2.2 Sparsity Prior and Minimizations

Choosing $\rho(u,v) = \frac{1}{2}\|D(u-v)\|_H^2$ for the misfit to the data is relatively "natural" as it corresponds to measuring noise in the $H = L^2$ sense. The quadratic functional $\mathcal{R}(u) = \frac{1}{2}\|Ru\|_H^2$ is, however, much less justified in many settings.

Sometimes, prior knowledge about the object we wish to reconstruct shows that the latter is sparse in a given representation (a given basis, say). Sparse means here that the object is represented by a small number of large coefficients. For instance, an audio signal may be represented by a finite number of frequencies. Images typically display sharp edges that can be represented in more economical fashion than pixel by pixel values.

Let us assume that u is discrete and A a matrix to simplify the presentation. Let us also assume that Bu is sparse, where B is a known matrix. Sparsity will be encoded by the fact that the l_1 norm of Bu is small. Penalizing the residual and the l_1 norm yields the minimization of

$$\mathcal{F}_\mu(u) = \|Bu\|_{l_1} + \frac{\mu}{2}\|Au - v_d\|_{l_2}^2. \tag{8.43}$$

This and similar minimization problems have been applied very successfully for a large class of imaging problems.

However, minimizing $\mathcal{F}_\mu$ above is significantly more difficult than solving (8.42). Several algorithms have been developed to solve such minimization problems. We present one strategy, called the split Bregman iteration, that is both efficient and relatively easy to explain. The main idea is that when B and A are the identity operators, then the above minimization can be performed for each component of u separately. In the general case, we introduce

$$d = Bu,$$

and replace the above minimization by

$$\min_{u,d} \ \|d\|_{l_1} + \frac{\mu}{2}\|Au - v_d\|_{l_2}^2 + \frac{\lambda}{2}\|d - Bu\|_{l_2}^2. \tag{8.44}$$

Choosing λ sufficiently large provides a good approximation of the problem we wish to solve. Alternatively, we can solve a series of problems of the above form and show that we minimize (8.43) in the limit; we do not present the details here and refer the reader to [24].

Now the minimization of (8.44) can be performed iteratively by successively minimizing for u and for d. The minimization for u becomes a linear problem while the minimization for d can be performed for each coordinate independently (this is called soft shrinkage). The iterative algorithm then converges [24]. More precisely, the solution of

$$\min_u \ \frac{\mu}{2}\|Au - v_d\|_{l_2}^2 + \frac{\lambda}{2}\|d - Bu\|_{l_2}^2,$$

is given by

$$(\mu A^*A + \lambda B^*B)u = \mu A^* v_d + \lambda d. \tag{8.45}$$

This is a linear problem that admits a unique solution. Now the solution of

$$\min_d \ \|d\|_{l_1} + \frac{\lambda}{2}\|d - Bu\|_{l_2}^2 = \min_d \sum_{j=1}^J |d_j| + \frac{\lambda}{2}|d_j - (Bu)_j|^2 = \sum_{j=1}^J \min_{d_j} |d_j| + \frac{\lambda}{2}|d_j - (Bu)_j|^2.$$

Now each element in the sum is minimized separately. We find that the solution of

$$\min_d |d| + \frac{\lambda}{2}|d-a|^2, \tag{8.46}$$

is given by the soft thresholding

$$d = \mathrm{sgn}(a)\ \max\Big(|a| - \frac{1}{\lambda}, 0\Big).$$

We have presented this algorithm to show that replacing a smoothness regularization as in (8.41) by a sparsity regularization as in (8.43) increased the computational complexity of the reconstruction algorithm: instead of solving one linear system, we have to iteratively solve linear systems of the form (8.45) and soft thresholdings given by (8.46). When the sparsity assumptions are valid, however, these methods have shown to be greatly beneficial in many practical settings of medical imaging.

8.3 Bayesian framework and statistical regularization

The penalty regularization framework seen in the preceding two sections is very efficient when the data are sufficiently informative. When the data are very informative and noise relatively small, then no real regularization is necessary. When data are less informative but still quite informative, prior information becomes necessary and smoothness and sparsity type priors allow us to still obtain very accurate reconstructions. When data are even less informative, for instance because noise is very large, or because the problem is severely ill-posed, then sparsity priors are typically no longer sufficient. What one typically obtains as a result is a function that resembles the minimum of the penalization term. In some cases, that may not be desirable. Also, prior information may also sometimes be known, for instance that next to a black pixel, there is never a blue pixel. Such information is difficult to include in a penalization method.

A fairly versatile methodology to include various prior informations is the Bayesian framework. To simplify the presentation slightly, let us assume that the problem of interest is

$$y = \mathfrak{M}(x) + n, \tag{8.47}$$

where $\mathfrak{M}$ is the measurement operator, x the unknown set of coefficients, y the measurements and n models additive noise in the data.

The main assumption of Bayesian inversions is to assume that x belongs to a class of possible models $\mathfrak{X}$ and that each $x \in \mathfrak{X}$ is given an *a priori* probability of being the "true" coefficient. The associated probability (density) $\pi(x)$ is called the *prior* distribution. A second ingredient in the Bayesian framework is the model for the noise n. We denote by $\pi_n(n)$ the probability distribution of n.

Let us now define

$$\pi(y|x) = \frac{\pi(x,y)}{\pi(x)}$$

the conditional probability density of y *knowing* x with $\pi(x,y)$ the probability density of x *and* y. Note that $\pi(x) = \int \pi(x,y)dy$ as a marginal density so that the above

conditional probability density is indeed a probability density (integrating to 1 in y). Note that knowledge of $\pi(y|x)$ is equivalent to knowledge of π_n since

$$\pi(y|x) = \pi_n(y - \mathfrak{M}(x)) \quad \text{for each fixed } x.$$

Bayes' rule then essentially states that

$$\pi(x|y)\pi(y) = \pi(y|x)\pi(x) = \pi(x, y). \tag{8.48}$$

In our inverse problem where y is the measured data and x the unknown coefficients, this means

$$\pi(x|y) = \frac{1}{\pi(y)}\pi(y|x)\pi(x) \propto \pi(y|x)\pi(x), \tag{8.49}$$

where $\propto$ means proportional to, i.e., up to a normalizing constant (here $1/\pi(y)$). In other words, if we know the prior density $\pi(x)$ and the likelihood function $\pi(y|x)$, then by Bayes' rule, we know $\pi(x|y)$, which is the *posterior* probability (density).

Let us recapitulate the main ingredients of the Bayesian formalism. We assume the *prior* distribution $\pi(x)$ known as an indication of our prior beliefs about the coefficients before data are acquired. We assume knowledge of the likelihood function $\pi(y|x)$, which as we have seen is a statement for the noise model in the experiment. From these two prior assumptions, we use Bayes' rule to infer the *posterior* distribution $\pi(x|y)$ for x knowing the data.

8.3.1 Penalization methods and Bayesian framework

Before going into the advantages and drawbacks of the method, we first show that penalization methods can be seen as an application of the Bayesian framework. Let $\mathcal{R}(x)$ be a given function and assume that the prior is given by the Gibbs distribution:

$$\pi(x) \propto e^{-\mathcal{R}(x)}.$$

Now assume that the likelihood function is of the form

$$\pi(y|x) \propto e^{-\rho(y-\mathfrak{M}(x))},$$

where ρ is a distance function. Then by Bayes' rule, we find that

$$\pi(x|y) \propto e^{-\left(\rho(y-\mathfrak{M}(x))+\mathcal{R}(x)\right)}.$$

The Maximum A Posteriori (MAP) x_{MAP} is the parameter that maximizes the posterior distribution, or equivalently the minimum of the functional

$$\mathcal{F}(x) = \rho(y - \mathfrak{M}(x)) + \mathcal{R}(x).$$

Therefore, for appropriate choices of the prior and likelihood function, we retrieve the penalization methods seen in the preceding section.

Note that the minimization problem is solved by linear algebra when both ρ and $\mathcal{R}$ are quadratic functionals. For instance if $\pi_n(n) \sim \mathcal{N}(0, \Sigma)$ a multivariate Gaussian with

correlation matrix Σ, then we have $\rho(n) \propto e^{-\frac{1}{2}n^t\Sigma^{-1}n}$. Similarly, for $\pi(x) \sim \mathcal{N}(0,\Gamma)$, then $\mathcal{R}(x) \propto e^{-\frac{1}{2}x^t\Gamma^{-1}x}$ so that we need to minimize

$$\mathcal{F}(x) = \frac{1}{2}(y - \mathfrak{M}x)^t\Sigma^{-1}(y - \mathfrak{M}x) + \frac{1}{2}x^t\Gamma^{-1}x.$$

If $\mathfrak{M}$ is a linear operator, then the solution to the minimization problem is, as we already saw, solution of

$$(\mathfrak{M}^*\Sigma^{-1}\mathfrak{M} + \Gamma^{-1})x = \mathfrak{M}^*\Sigma^{-1}y.$$

The Bayesian framework can then be used to recover the Tikhonov regularization of linear equations. Moreover, it gives an explicit characterization of the correlation matrices Σ and Γ as the co-variance functions of the measurement noise and of the prior assumptions on the coefficients, respectively.

Note that the l_1 minimization corresponds to a choice $\mathcal{R}(x) = \sum_i |x_i|$. This corresponds to assuming that each pixel value satisfies independent and identically distributed random variables with a Laplace distribution. We thus also recover the sparsity regularizations using the Bayesian framework. If we expect nearby pixels to be correlated, then more complex prior models or functionals $\mathcal{R}(x)$ need to be constructed. This is a still very active area of research. Although the derivation of the "best" functional is often more an art than grounded in first principles, the Bayesian framework sometimes allows for very "pleasing" reconstructions to the eye.

8.3.2 Computational and psychological costs of the Bayesian framework

We have seen above that the Bayesian framework reduced to an optimization problem when the Maximum A Posteriori (MAP) x_{MAP} is what we are looking for. The Bayesian framework allows one to obtain much more information, at least in theory, since the outcome of the procedure is the whole posterior distribution $\pi(x|y)$ and not only its argmax.

In practice, however, we are faced with daunting tasks: first of all, how do we sample what is often a very high dimensional distribution $\pi(x|y)$? And second of all, even if sampling is possible, how does one represent such a huge object practically? These two questions severely limit the applicability of Bayesian frameworks in practice.

A third, and in some sense more fundamental and structural, question pertains to the choice of the prior $\pi(x)$. Where should that prior information come from? There is no "good" answer to this fundamental yet ill-formulated question. In some sense, we have already partially answered it: since *you as a user* decided that adding no prior information was not working, in the sense that "*noise*" had too large an effect on the reconstruction, then *you as a user* have to come up with *another* model. The priori gives you a chance to come up with such a prior model. There is no such a thing as a "non-informative" prior, since *you as a user* decided that a prior was necessary. If data alone are sufficient to obtain "good" reconstructions, then the Bayesian framework is not necessary. If data are not sufficient, then the Bayesian framework provides a very versatile framework for *you as a user* to show that you know something about the problem that helps you compensate what is not present in the data. Some researchers will not be satisfied with this way of addressing the inverse problem and the notion of

"compensating" for the lack of data. This is a perfectly reasonable position. However, the Bayesian framework at least has this very appealing feature: it provides a logical mapping from the prior information, namely the prior distribution and the likelihood function, to the outcome of the procedure, namely the posterior distribution. If nothing else, it can therefore serve as a very valuable tool to guide intuition and to search what types of prior informations are necessary for a given set of constraints on the posterior distribution.

Now that the psychological cost of the Bayesian framework has been taken into account, let us come back to the computational cost for which we are at least on safer theoretical grounds and which still poses enormous challenges. Let us first address the representation of the posterior distribution. Typically, moments of the posterior distribution are what we are interested in. For instance, one may be interested in the first moment (a vector) and the variance (a matrix)

$$x_m = \int x\pi(x|y)dx, \qquad \Gamma_x = \int x \otimes x\, \pi(x|y)dx - x_m \otimes x_m. \tag{8.50}$$

Of interest are also various quantiles of the posterior distribution, for instance the probability that x_j be larger than a number γ: $\int \pi(x|y)\chi(x_j > \gamma)dx$, where $\chi(X)$ is the indicatrix function of the set X equal to 1 on X and to 0 otherwise.

For each of these moments of the posterior distribution, we need to be able to sample $\pi(x|y)dx$. In a few cases, the sampling of $\pi(x|y)dx$ may be straightforward, for instance when $\pi(y|x)$ has a Gaussian structure. In most cases, however, sampling is a difficult exercise. The most versatile method to perform such a sampling is arguably the Markov Chain Monte Carlo (MCMC) method. The objective of MCMC samplers is to generate a Markov chain X^i for $i \in \mathbb{N}$ whose invariant distribution (distribution at convergence when the algorithm converges) is the posterior distribution. There are two main MCMC samplers, the Gibbs sampler and the Metropolis-Hastings sampler. The latter is defined as follows. Let us assume that we want to sample a distribution $\pi(x|y)$

Let $q(x, x')$ be a given, positive, transition density from the vector x to the vector x' (it thus sums to 1 integrated in all possible vectors x' for each x). Let us then define

$$\alpha(x, x') := \min\Big(\frac{q(x, x')\pi(x'|y)}{q(x', x)\pi(x|y)}, 1\Big). \tag{8.51}$$

Note that the above quantity, which is all we need about $\pi(x|y)$ in the Metropolis-Hastings sampler, depends only on $\frac{\pi(x'|y)}{\pi(x|y)}$ and thus is independent of the normalizing constant of $\pi(x|y)$, which is typically not known in the Bayesian framework.

Let X^i the current state of the Markov chain. Let $\tilde{X}^{i+1}$ be drawn from the transition kernel $q(X^i, x')$. Then with probability $\alpha(X^i, \tilde{X}^{i+1})$, we accept the transition and set $X^{i+1} = \tilde{X}^{i+1}$ while with probability $1 - \alpha(X^i, \tilde{X}^{i+1})$, we reject the transition and set $X^{i+1} = X^i$.

The transition probability of the chain from x to x' is thus $p(x, x') = \alpha(x, x')q(x, x')$ while the probability to stay put at x is $1 - \int p(x, x')dx'$. The construction is such that $\pi(x|y)p(x, x') = p(x', x)\pi(x'|y)$, which means that $\pi(x|y)dy$ is indeed the invariant distribution of the Markov chain. In practice, we want independent samples of $\pi(x|y)$ so that the following Monte Carlo integration follows from an application of the law of

large numbers (ergodicity), for instance:

$$\int f(x)\pi(x|y)dx \sim \frac{1}{|I|}\sum_{i\in I} f(X^i), \tag{8.52}$$

for any reasonable (continuous) functional f. Such a rule is accurate if the X^i are sampled according to $\pi(x|y)dy$ and are sufficiently independent. This is for instance achieved by choosing $I = \{1 \leq i \leq i_{\max}, \quad i = Nj, \quad j \in \mathbb{N}\}$. For instance, we can take $i_{\max} = 10^7$ and $N = 1000$ so that I is composed of $|I| = 10^4$ points chosen every 1000 points in the Metropolis-Hastings Markov chain. For an accuracy equal to $\sqrt{|I|} = 0.01$ (as an application of the central limit theorem to estimate the error in (8.52)), we thus need 10^7 evaluations of $\pi(x|y)$. Using Bayes' rule, this is proportional to $\pi(x)\pi(y|x)$, where the latter likelihood function requires that we solve a forward problem (for a given x drawn from the prior $\pi(x)$) to estimate the law of the "data" y. In other words, the construction of the above statistical moment with an accuracy of order 10^{-2} requires that we solve 10^7 forward problems. In many practical situations, this is an unsurmountable computational cost. Moreover, this assumes that the transition $q(x, x')$ has been chosen in a way that every 1000 samples X^i are indeed sufficiently independent. This is very difficult to achieve in practice and is typically obtained by experienced users rather than from sound, physics- or mathematics- based principles.

Note that in practice, it has been observed that I in (8.52) should be the set of all runs $1 \leq i \leq i_{\max}$. In other words, there is no gain in throwing away points X^i in the evaluation of the integrals. However, the above heuristics are correct: the error in the approximation (8.52) is indeed proportional to the square root of the number of independent components in $\{X^i\}$ and not $\sqrt{i_{\max}}$.

Many methodologies have been developed to improve the efficiency of MCMC algorithms. It is however fair to say that even with nowadays computational capabilities, many problems of interest are totally out of reach using the standard Bayesian framework. That said, it is still a very versatile methodology that goes a long way to address our main problem, which I recall was: *you as a user* decided that adding no prior information was not working and thus something had to be done.

Chapter 9

Examples of geometrical prior information

We have seen in the preceding chapter that the Bayesian framework allowed for a versatile method to incorporate prior information into reconstructions. The main drawback of such methods is that a deterministic reconstruction is replaced by a measure of possible reconstructions, which is often daunting to sample and represent from a numerical point of view (computational cost) and requires prior information that different practitioners may disagree on (psychological cost).

Other, more physically motivated, methods allow us to incorporate (strong) prior information. In one such method, we aim to reconstruct the support of an inclusion rather than its full description and assume that such an inclusion is embedded into a *known* medium. Several techniques have been developed to do this and we shall focus on a method called the *factorization* method. The factorization method is a functional analytical tool that allows us to separate the influence of the unknown inclusion from that of the known background.

In another method, we assume that the inclusion is *small* and perform asymptotic expansions in its volume to understand the leading influence of such an inclusion on available measurements. The reconstruction then focuses on the first coefficients appearing in the asymptotic expansion assuming the surrounding background known. We now present the two methods in a simple setting, knowing that both methods have been extended in various ways.

9.1 Reconstructing the domain of inclusions

The reconstruction of physical parameters in an elliptic equation from boundary measurements, such as the Neumann-to-Dirichlet map, is a severely ill-posed problem. One should therefore not expect to reconstruct much more than a few coefficients modeling the physical parameters, such as for instance the first Fourier modes in a Fourier series expansion.

In certain applications, knowing the first few coefficients in a Fourier series expansion is not what one is interested in. One may have additional *a priori* knowledge about the object to be reconstructed and may want to use this knowledge to look for more appropriate parameters. In this chapter, we assume that the physical parameters are

given by a background, which is known, and an inclusion, from which we only know that it differs from the background. Moreover, we are not so much interested in the detailed structure of the inclusion as in its location. We thus wish to reconstruct an interface separating the background from the inclusion.

To reconstruct this interface, we use the method of *factorization*. The method provides a constructive method to obtain the support of the inclusion from the Neumann-to-Dirichlet (NtD) boundary measurements. Notice that the NtD measurements allow us a priori to reconstruct much more than the support of the inclusion. However, because we restrict ourselves to this specific reconstruction, we can expect to obtain more accurate results on location of the inclusion than by directly reconstructing the physical parameters on the whole domain.

9.1.1 Forward Problem

We consider here the problem in impedance tomography. The theory generalizes to a certain extent to problems in optical tomography.

Let $\gamma(\mathbf{x})$ be a conductivity tensor in an open bounded domain $X \in \mathbb{R}^n$ with Lipschitz boundary ∂X. We define Σ, a smooth surface in X, as the boundary of the inclusion. We denote by D the simply connected bounded open domain such that $\Sigma = \partial D$. This means that D is the domain "inside" the surface Σ. We also define $D^c = X \backslash \overline{D}$, of boundary $\partial D^c = \partial X \cup \Sigma$. We assume that $\gamma(\mathbf{x})$ is a smooth known background $\gamma(\mathbf{x}) = \gamma_0(\mathbf{x})$ on D^c, and that γ and γ_0 are smooth but different on D. For γ_0 a smooth known tensor on the full domain X, this means that γ jumps across Σ so that Σ is the surface of discontinuity of γ. More precisely, we assume that the $n \times n$ symmetric tensor $\gamma_0(\mathbf{x})$ is of class $C^2(\overline{X})$ and positive definite such that $\xi_i \xi_j \gamma_{0ij}(\mathbf{x}) \geq \alpha_0 > 0$ uniformly in $\mathbf{x} \in X$ and in $\{\xi_i\}_{i=1}^n = \boldsymbol{\xi} \in S^{n-1}$, the unit sphere in $\mathbb{R}^n$. Similarly, the $n \times n$ symmetric tensor $\gamma(\mathbf{x})$ is of class $C^2(\overline{D}) \otimes C^2(\overline{D^c})$ (in the sense that $\gamma(\mathbf{x})|_D$ can be extended as a function of class $C^2(D)$ and same thing for $\gamma(\mathbf{x})|_{D^c}$) and positive definite such that $\xi_i \xi_j \gamma_{ij}(\mathbf{x}) \geq \alpha_0 > 0$ uniformly in $\mathbf{x} \in X$ and in $\{\xi_i\}_{i=1}^n = \boldsymbol{\xi} \in S^{n-1}$.

The equation for the electric potential $u(\mathbf{x})$ is given by

$$\begin{aligned} \nabla \cdot \gamma \nabla u &= 0, && \text{in } X \\ \mathbf{n} \cdot \gamma \nabla u &= g && \text{on } \partial X \\ \int_{\partial X} u \, d\sigma &= 0. \end{aligned} \tag{9.1}$$

Here, $\mathbf{n}(\mathbf{x})$ is the outward unit normal to X at $\mathbf{x} \in \partial X$. We also denote by $\mathbf{n}(\mathbf{x})$ the outward unit normal to D at $\mathbf{x} \in \Sigma$. Finally $g(\mathbf{x})$ is a mean-zero current, i.e., $\int_{\partial X} g d\sigma = 0$, imposed at the boundary of the domain.

The above problem admits a unique solution $H_0^1(X)$, the space of functions in $u \in H^1(X)$ such that $\int_{\partial X} u d\sigma = 0$. This results from the variational formulation of the above equation

$$b(u, \phi) \equiv \int_X \gamma \nabla u \cdot \nabla \phi d\mathbf{x} = \int_{\partial X} g \phi d\sigma(\mathbf{x}) \equiv l(\phi), \tag{9.2}$$

holding for any test function $\phi \in H_0^1(X)$. Indeed from a Poincaré-like inequality, we deduce that $b(u, v)$ is a coercive and bounded bilinear form on $H_0^1(X)$ and the existence

result follows from the Lax-Milgram theorem. Classical trace estimates show that $u|_{\partial X} \in H_0^{1/2}(\partial X)$, the space of functions $v \in H^{1/2}(\partial X)$ such that $\int_{\partial X} v d\sigma = 0$.

We define the Neumann-to-Dirichlet operator Λ_Σ, depending on the location of the discontinuity Σ, as

$$\Lambda_\Sigma \ : \ H_0^{-1/2}(\partial X) \ \longrightarrow \ H_0^{1/2}(\partial X), \qquad g \ \longmapsto \ u_{|\partial X}, \tag{9.3}$$

where $u(\mathbf{x})$ is the solution to (9.1) with boundary normal current $g(\mathbf{x})$. Similarly, we introduce the "background" Neumann-to-Dirichlet operator Λ_0 defined as above with γ replaced by the known background γ_0. To model that the inclusion has a different conductivity from the background, we assume that γ satisfies either one of the following hypotheses

$$\gamma(\mathbf{x}) - \gamma_0(\mathbf{x}) \geq \alpha_1 > 0 \quad \text{on } D, \qquad \gamma_0(\mathbf{x}) = \gamma(\mathbf{x}), \quad \text{on } D^c, \tag{9.4}$$

$$\gamma_0(\mathbf{x}) - \gamma(\mathbf{x}) \geq \alpha_1 > 0 \quad \text{on } D, \qquad \gamma_0(\mathbf{x}) = \gamma(\mathbf{x}), \quad \text{on } D^c, \tag{9.5}$$

for some constant positive definite tensor α_1. The tensor inequality $\gamma_1 \geq \gamma_2$ is meant in the sense that $\xi_i \xi_j (\gamma_{1,ij} - \gamma_{2,ij}) \geq 0$ for all $\boldsymbol{\xi} \in \mathbb{R}^n$.

9.1.2 Factorization method

The purpose of the factorization method is to show that

$$\Lambda_0 - \Lambda_\Sigma = L^* F L, \tag{9.6}$$

where L and L^* are operators in duality that depend only on $\gamma|_D = (\gamma_0)|_D$ and $\pm F$ is an operator that generates a coercive form on $H_0^{1/2}(\Sigma)$ when (9.4) and (9.5) are satisfied, respectively. The operators are constructed as follows. Let v and w be the solutions of

$$\begin{array}{rclll} \nabla \cdot \gamma \nabla v & = & 0, & \text{in } D^c & \\ \mathbf{n} \cdot \gamma \nabla v & = & \phi & \text{on } \partial X & \\ \mathbf{n} \cdot \gamma \nabla v & = & 0 & \text{on } \Sigma & \\ \displaystyle\int_\Sigma v d\sigma & = & 0, & & \end{array} \qquad \begin{array}{rclll} \nabla \cdot \gamma \nabla w & = & 0, & \text{in } D^c & \\ \mathbf{n} \cdot \gamma \nabla w & = & 0 & \text{on } \partial X & \\ \mathbf{n} \cdot \gamma \nabla w & = & -\phi & \text{on } \Sigma & \\ \displaystyle\int_\Sigma w d\sigma & = & 0. & & \end{array} \tag{9.7}$$

These equations are well-posed in the sense that they admit solutions in $H^1(D^c)$ with traces in $H^{1/2}(\Sigma)$ and in $H^{1/2}(\partial X)$ at the boundary of D^c. We then define the operator L, which maps $\phi \in H_0^{-1/2}(\partial X)$ to $v|_\Sigma \in H_0^{1/2}(\Sigma)$, where v is the unique solution to the left equation in (9.7), and the operator L^*, which maps $\phi \in H_0^{-1/2}(\Sigma)$ to $w|_{\partial X}$, where w is the unique solution to the right equation in (9.7). We verify that both operators are in duality in the sense that

$$(L\phi, \psi)_\Sigma \equiv \int_\Sigma \psi L\phi \ d\sigma = \int_{\partial X} \phi L^* \psi \ d\sigma \equiv (\phi, L^*\psi)_{\partial X}.$$

Let us now define two operators G_Σ and G_Σ^* as follows. For any quantity f defined on $D \cup D^c$, we denote by $f^+(\mathbf{x})$ for $\mathbf{x} \in \Sigma$ the limit of $f(\mathbf{y})$ as $\mathbf{y} \to \mathbf{x}$ and $\mathbf{y} \in D^c$, and

by $f^-(\mathbf{x})$ the limit of $f(\mathbf{y})$ as $\mathbf{y} \to \mathbf{x}$ and $\mathbf{y} \in D$. Let v and w be the unique solutions to the following problems

$$\begin{array}{rcll} \nabla \cdot \gamma \nabla v &=& 0, & \text{in } X \backslash \Sigma \\ [v] &=& 0, & \text{on } \Sigma \\ [\mathbf{n} \cdot \gamma \nabla v] &=& 0 & \text{on } \Sigma \\ \mathbf{n} \cdot \gamma \nabla v &=& g & \text{on } \partial X \\ \displaystyle\int_\Sigma v \, d\sigma &=& 0 & \end{array} \qquad \begin{array}{rcll} \nabla \cdot \gamma \nabla w &=& 0, & \text{in } X \backslash \Sigma \\ [w] &=& \phi, & \text{on } \Sigma \\ [\mathbf{n} \cdot \gamma \nabla w] &=& 0 & \text{on } \Sigma \\ \mathbf{n} \cdot \gamma \nabla w &=& 0 & \text{on } \partial X \\ \displaystyle\int_{\partial X} w \, d\sigma &=& 0. & \end{array} \tag{9.8}$$

We define G_Σ as the operator mapping $g \in H_0^{-1/2}(\partial X)$ to $G_\Sigma g = \mathbf{n} \cdot \gamma \nabla v^+_{|\Sigma} \in H_0^{-1/2}(\Sigma)$ and the G_Σ^* as the operator mapping $\phi \in H_0^{1/2}(\Sigma)$ to $G_\Sigma^* \phi = w_{|\partial X} \in H_0^{1/2}(\partial X)$, where v and w are the unique solutions to the above equations (9.8).

Except for the normalization $\int_\Sigma v \, d\sigma = 0$, the equation for v is the same as (9.1) and thus admits a unique solution in $H^1(X)$, say. Moreover integrations by parts on D^c imply that

$$\int_\Sigma \mathbf{n} \cdot \gamma \nabla v^+ \, d\sigma = \int_{\partial X} g \, d\sigma = 0.$$

This justifies the well-posedness of the operator G_Σ as it is described above. The operator G_Σ^* is more delicate. We first obtain that for any smooth test function ψ,

$$\begin{array}{rcl} \displaystyle\int_{D^c} \gamma \nabla w \cdot \nabla \psi \, d\mathbf{x} + \int_\Sigma \mathbf{n} \cdot \gamma \nabla w \psi^+ \, d\sigma &=& 0 \\ \displaystyle\int_D \gamma \nabla w \cdot \nabla \psi \, d\mathbf{x} - \int_\Sigma \mathbf{n} \cdot \gamma \nabla w \psi^- \, d\sigma &=& 0, \end{array}$$

so that

$$\int_X \gamma \nabla w \cdot \nabla \psi \, d\mathbf{x} = \int_\Sigma (-\mathbf{n} \cdot \gamma \nabla w) \, [\psi] \, d\sigma. \tag{9.9}$$

It turns out that $\|\mathbf{n} \cdot \gamma \nabla w\|_{H_0^{-1/2}(\Sigma)}$ is bounded by the norm of $\gamma \nabla w$ in $H(\mathrm{div}, \mathrm{X})$ (see [23]). This and a Poincaré-type inequality shows that the above right-hand side with $\psi = w$ is bounded by $C\|\phi\|^2_{H_0^{1/2}(\Sigma)}$. Existence and uniqueness of the solution $w \in H^1(D) \otimes H^1(D^c)$ to (9.8) is then ensured by application of the Lax-Milgram theory. This also shows that the operator G_Σ^* as defined above is well-posed.

Integrations by parts in the equation for v in (9.8) by a test function φ yields

$$\begin{array}{rcl} \displaystyle\int_D \gamma \nabla u \cdot \nabla \varphi \, d\mathbf{x} - \int_\Sigma \mathbf{n} \cdot \gamma \nabla v \varphi^- d\sigma &=& 0 \\ \displaystyle\int_{D^c} \gamma \nabla v \cdot \nabla \varphi \, d\mathbf{x} + \int_\Sigma \mathbf{n} \cdot \gamma \nabla v \varphi^+ d\sigma &=& \displaystyle\int_{\partial X} g\varphi \, d\sigma, \end{array} \tag{9.10}$$

from which we deduce that

$$\int_X \gamma \nabla v \cdot \nabla \varphi = \int_{\partial X} g\varphi - \int_\Sigma (G_\Sigma g)[\varphi]. \tag{9.11}$$

That G_Σ and G_Σ^* are in duality in the sense that

$$\int_\Sigma G_\Sigma g\ \phi\ d\sigma = \int_{\partial X} g\ G_\Sigma^* \phi\ d\sigma, \tag{9.12}$$

follows from (9.11) with $\varphi = w$ and (9.9) with $\psi = v$ since $[v] = 0$.

We finally define F_Σ as the operator mapping $\phi \in H_0^{1/2}(\Sigma)$ to $F_\Sigma \phi = -\mathbf{n} \cdot \gamma \nabla w \in H_0^{-1/2}(\Sigma)$, where w is the solution to (9.8). Based on the above results, this is well-posed operator. Moreover, we deduce from (9.9) that

$$(F_\Sigma[w], [\psi])_\Sigma = \int_X \gamma \nabla w \cdot \nabla \psi\ d\mathbf{x} = ([w], F_\Sigma[\psi])_\Sigma, \tag{9.13}$$

so that $F_\Sigma = F_\Sigma^*$. Upon choosing $[w] = [\psi]$, we find that F_Σ is coercive on $H_0^{1/2}(\Sigma)$. This implies among other things that F_Σ is injective.

We now notice that

$$G_\Sigma^* = L^* F_\Sigma.$$

This follows from the uniqueness of the solution to the elliptic problem on D^c with conditions defined on $\partial D^c = \Sigma \cup \partial X$. By duality, this also implies that $G_\Sigma = F_\Sigma L$. The operators G_0 and F_0 are defined similarly except that γ is replaced by γ_0 in equations (9.8). Let us finally define the operator M, which maps $g \in H_0^{-1/2}(\partial X)$ to $u_{|\partial X} \in H_0^{1/2}(\partial X)$, where u is the solution to

$$\begin{aligned} \nabla \cdot \gamma \nabla u &= 0, && \text{in } D^c \\ \mathbf{n} \cdot \gamma \nabla u &= 0, && \text{on } \Sigma \\ \mathbf{n} \cdot \gamma \nabla u &= g, && \text{on } \partial X \\ \int_{\partial X} u\ d\sigma &= 0. \end{aligned} \tag{9.14}$$

Except for the normalization, the operator M is the same as the operator L (so that $L - M$ is proportional to identity) and is thus well-posed. We now verify from the linearity of the elliptic problems that

$$\Lambda_\Sigma = M - L^* G_\Sigma = M - L^* F_\Sigma L, \qquad \Lambda_0 = M - L^* G_0 = M - L^* F_0 L. \tag{9.15}$$

We thus deduce the main factorization result of this section, namely that

$$\Lambda_0 - \Lambda_\Sigma = L^* F L, \qquad F = F_\Sigma - F_0. \tag{9.16}$$

The above result would not be very useful if F did not have specific properties. We now show that F or $-F$ generates a coercive form on $H_0^{1/2}(\Sigma)$ and may be written as B^*B for some operator B^* *surjective*. Note that $F^* = F$ since both F_Σ and F_0 are self-adjoint.

We denote by w_Σ the solution w to (9.8) and by w_0 the solution to the same equation with γ replaced by γ_0. Upon multiplying the equation for w_Σ by w_0 and subtracting the equation for w_0 multiplied by w_Σ, we obtain since $\gamma = \gamma_0$ on D^c that

$$\begin{aligned} \int_D (\gamma - \gamma_0) \nabla w_0 \cdot \nabla w_\Sigma\, d\mathbf{x} &= \int_\Sigma (\mathbf{n} \cdot \gamma \nabla w_\Sigma w_0^- - \mathbf{n} \cdot \gamma \nabla w_0 w_\Sigma^-) d\sigma \\ 0 &= \int_\Sigma (\mathbf{n} \cdot \gamma \nabla w_\Sigma w_0^+ - \mathbf{n} \cdot \gamma \nabla w_0 w_\Sigma^+) d\sigma. \end{aligned}$$

Notice that both γ and ∇w_Σ jump across Σ but that $\mathbf{n}\cdot\gamma\nabla w_\Sigma$ does not. This yields that

$$\int_D (\gamma-\gamma_0)\nabla w_0\cdot\nabla w_\Sigma\,d\mathbf{x} = \int_\Sigma (F_\Sigma - F_0)\phi\,\phi\,d\sigma = \int_\Sigma F\phi\,\phi\,d\sigma. \tag{9.17}$$

Let us now introduce $\delta w = w_0 - w_\Sigma$. Upon multiplying $\nabla\cdot\gamma_0\nabla\delta w + \nabla\cdot(\gamma_0-\gamma)\nabla w_\Sigma = 0$ by δw and integrating by parts on D^c and D we deduce that

$$\int_X \gamma_0\nabla\delta w\cdot\nabla\delta w\,d\mathbf{x} + \int_D(\gamma-\gamma_0)\nabla w_\Sigma\cdot\nabla w_\Sigma\,d\mathbf{x} = \int_D(\gamma-\gamma_0)\nabla w_0\cdot\nabla w_\Sigma\,d\mathbf{x}.$$

By exchanging the roles of the indices Σ and 0 we also obtain

$$\int_X \gamma\nabla\delta w\cdot\nabla\delta w\,d\mathbf{x} + \int_D(\gamma_0-\gamma)\nabla w_0\cdot\nabla w_0\,d\mathbf{x} = \int_D(\gamma_0-\gamma)\nabla w_0\cdot\nabla w_\Sigma\,d\mathbf{x}.$$

Combining these results with (9.17) we deduce that

$$\begin{aligned}\int_\Sigma F\phi\,\phi\,d\sigma &= \int_X \gamma_0\nabla\delta w\cdot\nabla\delta w\,d\mathbf{x} + \int_D(\gamma-\gamma_0)\nabla w_\Sigma\cdot\nabla w_\Sigma\,d\mathbf{x}\\ \int_\Sigma -F\phi\,\phi\,d\sigma &= \int_X \gamma\nabla\delta w\cdot\nabla\delta w\,d\mathbf{x} + \int_D(\gamma_0-\gamma)\nabla w_0\cdot\nabla w_0\,d\mathbf{x}.\end{aligned} \tag{9.18}$$

Let us assume that (9.4) holds. Then F generates a coercive form on $H_0^{1/2}(\Sigma)$. Indeed, let us assume that the right-hand side of the first equality above is bounded. Then by a Poincaré-type inequality, we have $\delta w\in H^1(X)$ and $w_\Sigma|_D\in H^1(D)$ thanks to (9.4). This implies that $(\mathbf{n}\cdot\gamma\nabla w_\Sigma)_{|\Sigma}\in H^{-1/2}(\Sigma)$ and thus based on (9.8) that $w_\Sigma|_{D^c}\in H^1(D^c)$. This in turn implies that both w_Σ^+ and w_Σ^- belong to $H^{1/2}(\Sigma)$ so that their difference $\phi\in H_0^{1/2}(\Sigma)$. Thus, we have shown the existence of a positive constant C such that

$$\|\phi\|_{H_0^{1/2}(\Sigma)} \le C(F\phi,\phi)_\Sigma^{1/2}. \tag{9.19}$$

Exchanging the indices Σ and 0 also yields the existence of a constant C under hypothesis (9.5) such that

$$\|\phi\|_{H_0^{1/2}(\Sigma)} \le C(-F\phi,\phi)_\Sigma^{1/2}. \tag{9.20}$$

In what follows, we assume that (9.4) and (9.19) hold to fix notation. The final results are not modified when (9.5) and (9.20) hold instead.

The operator F is defined from $H_0^{1/2}(\Sigma)$ to $H_0^{-1/2}(\Sigma)$, which have not been identified. So writing $F = B^*B$ requires a little bit of work. Let $\mathcal{I}$ be the canonical isomorphism between $H_0^{-1/2}(\Sigma)$ and $H_0^{1/2}(\Sigma)$. Since it is positive definite we can decompose it as

$$\mathcal{I} = \mathcal{J}^*\mathcal{J}, \qquad \mathcal{J}:\ H_0^{-1/2}(\Sigma)\ \to\ L_0^2(\Sigma), \quad \mathcal{J}^*:\ L_0^2(\Sigma)\ \to\ H_0^{1/2}(\Sigma).$$

Both $\mathcal{J}$ and $\mathcal{J}^*$ are isometries as defined above. We can thus recast the coercivity of F as

$$(F\phi,\phi) = (F\mathcal{J}^*u,\mathcal{J}^*u) = (\mathcal{J}F\mathcal{J}^*u,u) \ge \alpha\|\phi\|^2_{H_0^{1/2}(\Sigma)} = \alpha\|u\|^2_{L_0^2(\Sigma)}.$$

So $\mathcal{J}F\mathcal{J}^*$ as a self-adjoint positive definite operator on $L^2(\Sigma)$ can be written as C^*C, where C and C^* are bounded operators from $L_0^2(\Sigma)$ to $L_0^2(\Sigma)$. Since

$$\|Cu\|^2_{L_0^2(\Sigma)} \ge \alpha\|u\|^2_{L_0^2(\Sigma)},$$

we deduce that C^* is surjective. We thus obtain that $F = B^*B$ where $B = C(\mathcal{J}^*)^{-1}$ maps $H_0^{1/2}(\Sigma)$ to $L_0^2(\Sigma)$ and its adjoint operator $B^* = \mathcal{J}^{-1}C^*$ maps $L_0^2(\Sigma)$ to $H_0^{1/2}(\Sigma)$. Since $\mathcal{J}$ is an isomorphism, we deduce that B^* is surjective.

From the above calculations we obtain that

$$\Lambda_0 - \Lambda_\Sigma = L^*FL = L^*B^*(L^*B^*)^* = A^*A, \qquad A = BL.$$

Since the Range of $(A^*A)^{1/2}$ for A acting on Hilbert spaces is equal to the Range of A^*, we deduce that

$$\mathcal{R}((\Lambda_0 - \Lambda_\Sigma)^{1/2}) = \mathcal{R}(L^*B^*) = \mathcal{R}(L^*) \tag{9.21}$$

since B^* is surjective. The latter is shown as follows. We always have that $\mathcal{R}(L^*B^*) \subset \mathcal{R}(L^*)$. Now for $y \in \mathcal{R}(L^*)$ there is x such that $y = L^*x$ and since B^* is surjective u such that $y = L^*B^*x$ so that $y \in \mathcal{R}(L^*B^*)$; whence $\mathcal{R}(L^*) \subset \mathcal{R}(L^*B^*)$.

When (9.5) and (9.20) hold instead of (9.4) and (9.19), we deduce that

$$\mathcal{R}((\Lambda_\Sigma - \Lambda_0)^{1/2}) = \mathcal{R}(L^*), \tag{9.22}$$

instead of (9.21). In both cases we have shown the following result.

Theorem 9.1.1 *Provided that* (9.4) *or* (9.5) *holds, the range of the operator* L^* *defined in* (9.7) *is determined by the Neumann-to-Dirichlet operator* Λ_Σ.

9.1.3 Reconstruction of Σ

The above theorem gives us a method to reconstruct Σ: we need to find a family of functions that belong to the Range of L^* when some parameter covers D and do not belong to it when the parameter covers D^c. Notice that the operator L^* does not depend on the domain D and thus only depends on the tensor γ_0 and the surface Σ. Consequently, the reconstruction of Σ will be independent of γ on γ_0, except for the existence of a positive definite tensor α_0 such that (9.4) or (9.5) holds.

Let us now introduce the family of functions $N(\cdot;\mathbf{y})$ indexed by the parameter $\mathbf{y} \in X$ solution of

$$\begin{aligned} \nabla \cdot \gamma_0 \nabla N(\cdot;\mathbf{y}) &= \delta(\cdot - \mathbf{y}), && \text{in } X \\ \mathbf{n} \cdot \gamma_0 \nabla N(\cdot;\mathbf{y}) &= 0 && \text{on } \partial X \\ \int_{\partial X} N(\cdot;\mathbf{y})\, d\sigma &= 0. \end{aligned} \tag{9.23}$$

We define the family of functions $g_{\mathbf{y}}(\mathbf{x}) = N(\mathbf{x};\mathbf{y})_{|\partial X}$ on ∂X. We then have

Theorem 9.1.2 *The function* $g_{\mathbf{y}}(\mathbf{x})$ *belongs to* $\mathcal{R}(L^*)$ *when* $\mathbf{y} \in D$ *and does not belong to* $\mathcal{R}(L^*)$ *when* $\mathbf{y} \in D^c$.

This theorem provides us with a constructive method to image $\Sigma = \partial D$. For each $\mathbf{y} \in X$, all we have to do is to solve (9.23) and verify whether the trace on ∂X belongs to the Range of $(\pm(\Lambda_0 - \Lambda_\Sigma))^{1/2}$, which can be evaluated boundary measurements. Only when the verification is positive can we deduce that $\mathbf{y} \in D$.

Proof. The proof of the theorem is as follows. When $\mathbf{y} \in D$, we have that $\mathbf{n} \cdot \gamma \nabla N(\mathbf{x};\mathbf{y})_{|\Sigma} \in H_0^{-1/2}(\Sigma)$ and $\nabla \cdot \gamma_0 \nabla N(\cdot;\mathbf{y}) = 0$ on D^c so that $g_{\mathbf{y}} \in \mathcal{R}(L^*)$. Let

us now assume that $\mathbf{y} \in D^c$ and $g_{\mathbf{y}}(\mathbf{x}) \in \mathcal{R}(L^*)$. Then there exists $\phi \in H_0^{-1/2}(\Sigma)$ such that $g_{\mathbf{y}} = L^*\phi = w_{|\partial X}$, where w is the solution to (9.7).Let $B(\mathbf{y};\varepsilon)$ be the ball of radius ε centered at $\mathbf{y}$ for ε sufficiently small. On $D^c\backslash\overline{B_\varepsilon}$, both w and $g_{\mathbf{y}}$ satisfy the same equation. By uniqueness of the solution to the Cauchy problem imposing Dirichlet data and vanishing Neumann data on ∂X, we deduce that $w = g_{\mathbf{y}}$ on $D^c\backslash\overline{B_\varepsilon}$. On $\omega_\varepsilon = B_{\varepsilon_0}\backslash\overline{B_\varepsilon}$ for some fixed $\varepsilon_0 > 0$, we verify that the $H^1(\omega_\varepsilon)$ norm of w remains bounded independently of ε, which is not the case for the fundamental solution $g_{\mathbf{y}}$; whence the contradiction implying that $g_{\mathbf{y}}$ is not in the Range of L^* when $\mathbf{y} \in D^c$. □

9.2 Reconstructing small inclusions

This second section concerns the reconstruction of small inclusions. We have seen that the reconstruction of diffusion or absorption coefficients in an elliptic equation resulted in a severely ill-posed problem. The previous chapter dealt with the issue by reconstructing the support of an inclusion instead of its detailed structure. Because the support of the inclusion remains an infinite dimensional object, the stability of the reconstruction is still a severely ill-posed problem. Here we further simplify the problem by assuming that the inclusions have small support. This introduces a small parameter allowing us to perform asymptotic expansions. We can then characterize the influence of the inclusion on the boundary measurements by successive terms in the expansion. The interest of such a procedure is the following. Since high-order terms in the expansion become small very quickly, the procedure tells us which parameters can be reconstruction from a given noise level in the measurements and which parameters cannot possibly be estimated. Moreover, these parameters come in small numbers. Each term in the asymptotic expansion is characterized by a finite number of parameters. This implies that by truncating the expansion, we are in the end looking at the reconstruction of only a finite number of parameters. Unlike the previous reconstructions, this becomes a well-posed problem since ill-posedness inevitably comes from the infinite dimensionality of the object we want to reconstruct, at least as long as the mapping from the object to be reconstructed to the noise-free measurements is one-to one (injective), which was always the case in the problems treated so far.

We consider in this chapter a mathematically very simple problem, namely the reconstruction of inclusions characterized by a variation in the absorption coefficient. We also restrict ourselves to the reconstruction from the leading term in the aforementioned asymptotic expansion. The interesting problem of variations of the diffusion coefficient is mathematically more difficult, although the main conclusions are in the end very similar. The presentation follows that in [7].

9.2.1 First-order effects

Let us consider the problem of optical tomography modeled by a diffusion equation on a domain X with current density $g(\mathbf{x})$ prescribed at the boundary ∂X. We assume here that the diffusion coefficient is known and to simplify, is set to $D \equiv 1$. Our main hypothesis on the absorption coefficient is that is is the superposition of a background absorption, to simplify the constant σ_0, and a finite number of fluctuations of arbitrary size $\sigma_m - \sigma_0$, with σ_m constant to simplify, but of small volume. Smallness of the volume

of the inclusions compared to the volume of the whole domain is characterized by the small parameter $\varepsilon \ll 1$. The diffusion equation with small absorption inclusions takes then the form

$$\begin{aligned} -\Delta u_\varepsilon(\mathbf{x}) + \sigma_\varepsilon(\mathbf{x}) u_\varepsilon(\mathbf{x}) &= 0, \qquad X \\ \frac{\partial u_\varepsilon}{\partial \boldsymbol{\nu}} &= g, \qquad \partial X, \end{aligned} \tag{9.24}$$

where absorption is given by

$$\sigma_\varepsilon(\mathbf{x}) = \sigma_0 + \sum_{m=1}^{M} \sigma_m \chi_{\mathbf{z}_m + \varepsilon B_m}(\mathbf{x}). \tag{9.25}$$

We have introduced here εB_m as the shape of the mth inclusion centered at $\mathbf{z}_m$, and $\chi_{\mathbf{z}_m+\varepsilon B_m}(\mathbf{x}) = 1$ if $\mathbf{x} - \mathbf{z}_m \in \varepsilon B_m$ and 0 otherwise. The inclusions are centered at $\mathbf{0}$ in the sense that

$$\int_{B_m} \mathbf{x} d\mathbf{x} = \mathbf{0} \qquad \text{for all } m, \tag{9.26}$$

and are assumed to be at a distance greater than $d > 0$, independent of ε, of each-other and of the boundary ∂X. The parameter ε is a measure of the diameter of the inclusions. In the three-dimensional setting, which we assume from now on, this implies that the volume of the inclusions is of order ε^3.

We want to derive an asymptotic expansion for u_ε in powers of ε and see which information on the inclusions we deduce from the first terms in the expansion. Let us first define the Green function of the corresponding homogeneous problem

$$\begin{aligned} -\Delta G(\mathbf{x};\mathbf{z}) + \sigma_0 G(\mathbf{x};\mathbf{z}) &= \delta(\mathbf{x}-\mathbf{z}), \qquad X \\ \frac{\partial G}{\partial \boldsymbol{\nu}}(\mathbf{x};\mathbf{z}) &= 0, \qquad \partial X, \end{aligned} \tag{9.27}$$

the homogeneous-domain solution $U(\mathbf{x})$ of

$$\begin{aligned} -\Delta U(\mathbf{x}) + \sigma_0 U(\mathbf{x}) &= 0, \qquad X \\ \frac{\partial U}{\partial \boldsymbol{\nu}}(\mathbf{x}) &= g(\mathbf{x}), \qquad \partial X. \end{aligned} \tag{9.28}$$

As $\varepsilon \to 0$, the volume of the inclusions tends to 0 and u_ε converges to U. To show this, we multiply (9.27) by u_ε and integrate by parts to obtain

$$u_\varepsilon(\mathbf{z}) = \int_{\partial X} g(\mathbf{x}) G(\mathbf{x};\mathbf{z}) d\sigma(\mathbf{x}) - \sum_{m=1}^{M} \int_{\mathbf{z}_m+\varepsilon B_m} \sigma_m G(\mathbf{x};\mathbf{z}) u_\varepsilon(\mathbf{x}) d\mathbf{x}.$$

Using the same procedure for $U(\mathbf{x})$, we obtain

$$u_\varepsilon(\mathbf{z}) = U(\mathbf{z}) - \sum_{m=1}^{M} \int_{\mathbf{z}_m+\varepsilon B_m} \sigma_m G(\mathbf{x};\mathbf{z}) u_\varepsilon(\mathbf{x}) d\mathbf{x}. \tag{9.29}$$

In three space dimensions, the Green function is given by

$$G(\mathbf{x};\mathbf{z}) = \frac{e^{-\sqrt{\sigma_0}|\mathbf{z}-\mathbf{x}|}}{4\pi|\mathbf{z}-\mathbf{x}|} + w(\mathbf{x};\mathbf{z}), \tag{9.30}$$

where $w(\mathbf{x};\mathbf{z})$ is a smooth function (because it solves (9.28) with smooth boundary conditions) provided that ∂X is smooth. For $\mathbf{z}$ at a distance greater than $d>0$ away from the inclusions $\mathbf{x}_m+\varepsilon B_m$, we then deduce from the L^∞ bound on u_ε (because g and ∂X are assumed to be sufficiently regular) that

$$u_\varepsilon(\mathbf{z}) = U(\mathbf{z}) + O(\varepsilon^3).$$

In the vicinity of the inclusions, we deduce from the relation

$$\int_{\mathbf{z}_m+\varepsilon B_m} G(\mathbf{x};\mathbf{z})d\mathbf{x} = O(\varepsilon^2), \qquad \mathbf{z}-\mathbf{z}_m \in \varepsilon B_m,$$

that $u_\varepsilon(\mathbf{z}) - U(\mathbf{z})$ is of order ε^2 when $\mathbf{z}$ is sufficiently close to an inclusion. This also shows that the operator

$$K_\varepsilon u_\varepsilon(\mathbf{z}) = -\sum_{m=1}^M \int_{\mathbf{z}_m+\varepsilon B_m} \sigma_m G(\mathbf{x};\mathbf{z}) u_\varepsilon(\mathbf{x})d\mathbf{x} \tag{9.31}$$

is a bounded linear operator in $\mathcal{L}(L^\infty(X))$ with a norm of order ε^2. This implies that for sufficiently small values of ε, we can write

$$u_\varepsilon(\mathbf{z}) = \sum_{k=0}^\infty K_\varepsilon^k U(\mathbf{z}). \tag{9.32}$$

The above series converges fast when ε is small. Notice however that the series does not converge as fast as ε^3, the volume of the inclusions, because of the singular behavior of the Green function $G(\mathbf{x};\mathbf{z})$ when $\mathbf{x}$ is close to $\mathbf{z}$.

Let us now use that

$$\begin{aligned} u_\varepsilon(\mathbf{z}) = U(\mathbf{z}) &- \sum_{m=1}^M \int_{\mathbf{z}_m+\varepsilon B_m} \sigma_m G(\mathbf{x};\mathbf{z}) U(\mathbf{x})d\mathbf{x} \\ &+ \sum_{m=1}^M \sum_{n=1}^M \int_{\mathbf{z}_m+\varepsilon B_m} \int_{\mathbf{z}_n+\varepsilon B_n} \sigma_m\sigma_n G(\mathbf{x};\mathbf{z}) G(\mathbf{y};\mathbf{x}) u_\varepsilon(\mathbf{y})d\mathbf{y}d\mathbf{x}. \end{aligned} \tag{9.33}$$

For the same reasons as above, the last term is of order ε^5, and expanding smooth solutions $U(\mathbf{x})$ and $G(\mathbf{x};\mathbf{z})$ inside inclusions of diameter ε, we obtain that

$$u_\varepsilon(\mathbf{x}) = U(\mathbf{x}) - \sum_{m=1}^M G(\mathbf{z};\mathbf{z}_m) C_m U(\mathbf{z}_m) + O(\varepsilon^5), \tag{9.34}$$

where C_m is given by

$$C_m = \varepsilon^3 |B_m| \sigma_m. \tag{9.35}$$

The reason why we obtain a correction term of order ε^5 in (9.34) comes from the fact that (9.26) holds so that the terms of order ε^4, proportional to $\mathbf{x}\cdot\boldsymbol{\nabla} U$ or $\mathbf{x}\cdot\boldsymbol{\nabla} G$, vanish.

9.2.2 Stability of the reconstruction

The above analysis tells us the following. Provided that our measurement errors are of order $O(\varepsilon^5)$, the only information that can possibly be retrieved on the inclusions is its location $\mathbf{z}_m$ and the product $C_m = \varepsilon^3 \sigma_m B_m$ of the absorption fluctuation with the volume of the inclusion. More refined information requires data with less noise. This information on the inclusion may look unimpressive. Yet assuming that the inclusions are sufficiently small so that the above asymptotic expansion makes sense, no other information can be obtained in a stable fashion from the data.

Notice that the problem to solve is now finite-dimensional. Indeed, each inclusion is represented by four real numbers, namely the three components of the position $\mathbf{z}_m$ and the product C_m. Assuming that only M inclusions are present, this leaves us with $4M$ parameters to reconstruct. The main advantage of reconstructing a finite number of parameters is that it is natural to expect stability of the reconstruction. We can even show stability of the reconstruction from boundary measurements corresponding to one current density $g(\mathbf{x})$ provided that the homogeneous solution $U(\mathbf{x})$ is uniformly positive inside the domain. Here is how it can be proved.

Let us assume that the boundary measurements have an accuracy of order $O(\varepsilon^5)$, consistent with

$$u_\varepsilon(\mathbf{z}) = U(\mathbf{z}) - \sum_{m=1}^{M} C_m (G(\mathbf{z}_m; \mathbf{z}) U(\mathbf{z}_m)) + O(\varepsilon^5). \tag{9.36}$$

We denote by u_ε and u'_ε the solution of two problems with absorption coefficients σ_ε and σ'_ε of the form (9.25). Using (9.36), we obtain that

$$u_\varepsilon(\mathbf{z}) - u'_\varepsilon(\mathbf{z}) = F(\mathbf{z}) + O(\varepsilon^5),$$

with

$$F(\mathbf{z}) = -\sum_{m=1}^{M} \Big(C_m (G(\mathbf{z}_m; \mathbf{z}) U(\mathbf{z}_m)) - C'_m (G(\mathbf{z}'_m; \mathbf{z}) U(\mathbf{z}'_m)) \Big). \tag{9.37}$$

Here we use $M = \max(M, M')$ with a small abuse of notation; we will see shortly that $M = M'$. The function $F(\mathbf{z})$ satisfies the homogeneous equation $-\Delta F + \sigma_0 F = 0$ on X except at the points $\mathbf{z}_m$ and $\mathbf{z}'_m$. Moreover, we have that $\frac{\partial F}{\partial \nu} = 0$ at ∂X. If $F = 0$ on ∂X, we deduce from the uniqueness of the Cauchy problem for the operator $-\Delta + \sigma_0$ that $F \equiv 0$ in X. As $\varepsilon \to 0$ and $u_\varepsilon - u'_\varepsilon \to 0$, we deduce that $F(\mathbf{z})$ becomes small not only at ∂X but also inside X (the continuation of F from ∂X to $X \backslash \{\mathbf{z}_m \cup \mathbf{z}'_m\}$ is independent of ε). However, the functions $G(\mathbf{z}_m; \mathbf{z}) U(\mathbf{z}_m)$ form an independent family. Each term must therefore be compensated by a term from the sum over the *prime* coefficients. We thus obtain that $M = M'$ and that

$$\Big| C_m (G(\mathbf{z}_m; \mathbf{z}) U(\mathbf{z}_m)) - C'_m (G(\mathbf{z}'_m; \mathbf{z}) U(\mathbf{z}'_m)) \Big| \;\leq\; C \| u_\varepsilon - u'_\varepsilon \|_{L^\infty(\partial X)} + O(\varepsilon^5).$$

The left-hand side can be recast as

$$(C_m - C'_m) G(\mathbf{z}_m; \mathbf{z}) U(\mathbf{z}_m) + C'_m (\mathbf{z}_m - \mathbf{z}'_m) \partial_{\mathbf{z}_m} (G(\bar{\mathbf{z}}_m; \mathbf{z}) U(\bar{\mathbf{z}}_m))$$

where $\bar{\mathbf{z}}_m = \theta \mathbf{z}_m + (1-\theta)\mathbf{z}'_m$ for some $\theta \in (0,1)$. Again these two functions are linearly independent so we deduce that

$$|C_m - C'_m| + |C'_m||\mathbf{z}_m - \mathbf{z}'_m| \leq C\|u_\varepsilon - u'_\varepsilon\|_{L^\infty(\partial X)} + O(\varepsilon^5).$$

Using (9.34) and (9.35), we then obtain assuming that $\|u_\varepsilon - u'_\varepsilon\|_{L^\infty(\partial X)} \approx \varepsilon^5$, that

$$\left|B_m\sigma_m - B'_m\sigma'_m\right| + |\mathbf{z}_m - \mathbf{z}'_m| \leq C\varepsilon^{-3}\|u_\varepsilon - u'_\varepsilon\|_{L^\infty(\partial X)} \approx \varepsilon^2. \tag{9.38}$$

Assuming that the accuracy of the measured data is compatible with the expansion (9.34), i.e. that the u_ε is known on ∂X up to an error term of order ε^5, we can then reconstruct the location $\mathbf{z}_m$ of the heterogeneities up to an error of order ε^2. The product of the volume of the inclusion and the absorption fluctuation is also known with the same accuracy.

Chapter 10

Hybrid Inverse Problems

An alternative to the prior assumptions we have made in the preceding two chapters is to acquire *different* measurements. Consider an elliptic equation and the reconstruction of the diffusion coefficient (the Calderón or inverse diffusion problem). The main difficulty is that the elliptic operator is a smoothing operator. As a consequence, by the time information is measured by the detectors at the domain's boundary, high frequencies have been exponentially attenuated and the inverse problem is severely ill-posed.

The cure to such an ill-posedness is to use well-posed inversions, for instance, inverse wave problems or inverse problems of integral geometry (such as the Radon transform). Since the solution of the diffusion equation does not solve a wave equation, we need to be sufficiently lucky that a physical phenomenon couples the wave-like and the elliptic-like phenomena. This way, we can obtain a *hybrid* imaging technique that couples the high contrast (but low-resolution) of the diffusion-like equation with the high resolution (but often low contrast) of the wave-like equation.

In this chapter, we consider two such physical couplings. The first coupling is the photo-acoustic effect whereas the second coupling is the ultrasound modulation coupling.

10.1 Quantitative Photo-acoustic tomography

Photoacoustic tomography (PAT) is a recent hybrid medical imaging modality that combines the high resolution of acoustic waves with the large contrast of optical waves. When a body is exposed to short pulse radiation, typically emitted in the near infra-red region in PAT, it absorbs energy and expands thermo-elastically by a very small amount; this is the photoacoustic effect. Such an expansion is sufficient to emit acoustic pulses, which travel back to the boundary of the domain of interest where they are measured by arrays of transducers.

The acoustic signals are modeled by the following wave equation

$$\frac{1}{c_s^2(x)}\frac{\partial^2 p}{\partial t^2} - \Delta p = \Gamma \frac{\partial \delta_0(t)}{\partial t} H(x), \tag{10.1}$$

with c_s the sound speed and Γ a coupling coefficient assumed to be constant and known. The acoustic pressure $p(t,x)$ is then measured on ∂X as a function of time. Finally, the amount of absorbed radiation $H(x)$ is given by

$$H(x) = \Gamma(x)\sigma(x)u(x), \tag{10.2}$$

where $\sigma(x)$ is the absorption coefficient, $\Gamma(x)$ is the Grüneisen coefficient quantifying the photo-acoustic effect, and $u(x)$ is the intensity of radiation at point x. Note that the wave equation above corresponds to stating that p solves an equation with vanishing sources when $t > 0$ and with initial conditions such that $p(t = 0, x) = H(x)$ and $\partial_t p(t = 0, x) = 0$.

A first step in PAT is therefore the reconstruction of an initial condition $H(x)$ in a wave equation from boundary measurements. Assuming that the sound speed c_s is constant and that the geometry of acquisition of $p(t, x)$ is simple, for instance with detectors located on a sphere containing the domain of interest X, then the reconstruction of $H(x)$ is a well-posed problem that admits explicit expressions. We do not consider further the first step of PAT. Let us simply note that when the sound speed is not constant, or when attenuation of waves are accounted for, the reconstruction of $H(x)$ can be significantly more complicated.

Once this first step is done, QPAT then consists of reconstructing $(D(x), \sigma(x), \Gamma(x))$ from knowledge of

$$H_j(x) = \Gamma(x)\sigma(x)u_j(x), \qquad 1 \leq j \leq J, \tag{10.3}$$

for $J \in \mathbb{N}^*$ illumination maps $g_j(x)$, where u_j is the solution to the steady-state equation

$$\begin{aligned} &-\nabla \cdot D(x)\nabla u_j + \sigma(x)u_j = 0, \qquad x \in X \subset \mathbb{R}^n, \\ &u_j = g_j \qquad x \in \partial X. \end{aligned} \tag{10.4}$$

Throughout the paper, we assume that all the coefficients are known at the domain's boundary ∂X. Our objective is to reconstruct them in X. Unfortunately, no matter how large J and the illuminations g_j may be chosen, we cannot reconstruct all of $(D(x), \sigma(x), \Gamma(x))$ from QPAT measurements of the form (10.3). However, when, say Γ is known (it is typically assumed to be constant), then we can reconstruct the other two coefficients $(D(x), \sigma(x))$ provided that illuminations are well-chosen. This is an interesting result. That we can reconstruct σ from $H(x)$ is expected. However, $D(x)$ appears in $H(x)$ only implicitly through $u(x)$. That it can be reconstructed is not clear a priori. The main results in QPAT are

- Two well chosen illuminations provide two independent relations $\chi = \chi(D, \sigma, \Gamma)$ and $q = q(D, \sigma, \Gamma)$ of the three coefficients (D, σ, Γ). This allows us to uniquely reconstruct two out of the three coefficients (D, σ, Γ) provided the third one is known.

- These two independent relations uniquely determine the measurements $H(x)$ for all other possible illuminations $g(x)$ on ∂X. In other words, independent of the number of illuminations and corresponding measurements, all that we can reconstruct about (D, σ, Γ) is (χ, q). This makes it impossible to reconstruct the three coefficients (D, σ, Γ) from QPAT data without additional prior information.

- For two well-chosen illuminations, the reconstruction of (χ, q) is Hölder-stable in appropriate norms, which means that an error of order ε in the data in an appropriate norm generates an error in the reconstruction of (χ, q) of order ε^κ in another appropriate norm for some $\kappa > 0$. (We say that a reconstruction is Lipschitz-stable when $\kappa = 1$.)

The last item provides the main advantage compared to the Calderón problem: because we have access to internal measurements, singularities no longer need to propagate to the domain's boundary. The reconstructions in QPAT are significantly more stable than for boundary measurements. However, only two out of three coefficients can be reconstructed from QPAT measurements. In practice, Γ is often assumed to be constant and known. In such a situation, the optical coefficients D and σ can be uniquely and stably reconstructed. Moreover, we have an explicit reconstruction procedure as we shall describe soon. However, such a procedure is guaranteed to work only when well-chosen illuminations g_j for $j = 1, 2$ are used. In dimension $n = 2$, a large class of illuminations works independent of the optical parameters. This is no longer the case in dimension $n \geq 3$. In the latter case, it is possible to show that there exist illuminations such that the above reconstruction can be performed in a stable manner. However, the description of these illuminations is rather implicit, which remains a problem in practice.

Here are now some mathematical assumptions on the coefficients and on two illuminations (boundary conditions) that we call *well-chosen*. Here and below, we denote by $W^{m,p}(X)$ the space of functions with derivatives of order less than or equal to m in $L^p(X)$.

(i) The coefficients (D, σ, Γ) are of class $W^{1,\infty}(X)$ and bounded above and below by positive constants. The coefficients (D, σ, Γ) are known on ∂X.

(ii) The illuminations g_1 and g_2 are positive functions on ∂X and we assume that they are the values (the restrictions) on ∂X of functions of class $C^3(\bar{X})$ (i.e., functions that are three times differentiable with continuous derivative of order 3 on $\bar{X} = X \cup \partial X$).

(iii) the vector field

$$\beta := H_1 \nabla H_2 - H_2 \nabla H_1 = H_1^2 \nabla \frac{H_2}{H_1} = H_1^2 \nabla \frac{u_2}{u_1} = -H_2^2 \nabla \frac{H_1}{H_2} \tag{10.5}$$

is a vector field in $W^{1,\infty}(X)$ such that

$$|\beta|(x) \geq \alpha_0 > 0, \qquad \text{a.e. } x \in X. \tag{10.6}$$

By standard regularity theory for elliptic equations and the maximum principle, the solutions to (10.4) are of class $W^{3,p}(X)$. Beyond the regularity assumptions on (D, σ, Γ), the domain X, and the boundary conditions g_1 and g_2 , the only real assumption we impose is thus (10.6). Let us make several remarks on this condition.

(i) In general, there is no guaranty that the gradient of $\frac{u_2}{u_1}$ does not vanish. Not all pairs of illuminations (g_1, g_2) are *well-chosen*.

(ii) That the vector field β does not vanish is a *sufficient* condition for the stability estimates presented below to be satisfied. It is not *necessary*. Stable (though probably less stable) reconstructions are still expected for a much larger choice of illuminations.

(iii) Guaranteeing (10.6) is relatively straightforward in dimension $n = 2$. It is much complicated in dimension $n \geq 3$. The only available methodology to ensure that (10.6) holds for a large class of conductivities is based on the method of complex geometric optics (CGO) solutions.

In dimension $n = 2$, we have:

Lemma 10.1.1 *Assume that $h = \frac{g_2}{g_1}$ on ∂X is an almost two-to-one function, i.e., a function that is a two-to-one map except possibly at its minimum and at its maximum. Then* (10.6) *is satisfied.*

Proof. Upon multiplying the equation for u_1 by u_2, the equation for u_2 by u_1, and subtracting both relations, we obtain

$$\begin{aligned} -\nabla \cdot (Du_1^2)\nabla \frac{u_2}{u_1} &= 0, \quad \text{in } X \\ \frac{u_2}{u_1} &= \frac{g_2}{g_1}, \quad \text{on } \partial X. \end{aligned} \tag{10.7}$$

This implies that $v := \frac{u_2}{u_1}$ satisfies an elliptic equation with a diffusion coefficient $\tilde{D} = Du_1^2$ bounded from above and below by positive constants. Note that $\beta = H_1^2 \nabla v$. Results in, e.g., [3, Theorem 1.2] show that ∇v cannot vanish inside X. By the maximum principle and the assumption on h, no critical point of v can occur on ∂X either. This implies that $|\nabla v| > 0$ and that we can find a constant such that (10.6) holds since H_1^2 is bounded from below by a positive constant and by continuity $|\nabla v|$ attains its (strictly positive) minimum in $\bar{X}$. □

Note that the above proof, as most proofs in this chapter, involve the analysis of solutions of (mostly elliptic) partial differential equations. Standard results in this very well developed field are not recalled here. We refer the (courageous) reader to [21, 22] for the details. In dimension $n \geq 3$, the above result on the (absence of) critical points of elliptic solutions no longer holds. However, by continuity, we may verify that (10.6) is satisfied for a large class of illuminations when D is close to a constant and σ is sufficiently small. For arbitrary coefficients (D, σ) in dimension $n \geq 3$, the only available proof that (10.6) is satisfied for an open set of illuminations is the one obtained in [16].

Uniqueness result. We first prove a result that provides uniqueness up to a specified transformation.

Theorem 10.1.2 *Assume that hypotheses (i)-(iii) hold. Then*

(a) $H_1(x)$ and $H_2(x)$ uniquely determine the whole measurement operator $\mathcal{H} : H^{\frac{1}{2}}(\partial X) \to H^1(X)$, which to g defined on ∂X associates $\mathcal{H}(g) = H$ in X defined by (10.3).

(b) The measurement operator $\mathcal{H}$ uniquely determines the two following functionals of (D, σ, Γ):

$$\chi(x) := \frac{\sqrt{D}}{\Gamma\sigma}(x), \qquad q(x) := -\Big(\frac{\Delta\sqrt{D}}{\sqrt{D}} + \frac{\sigma}{D}\Big)(x). \tag{10.8}$$

Here Δ is the Laplace operator.

(c) *Knowledge of the two functionals χ and q uniquely determines $H_1(x)$ and $H_2(x)$. In other words, the reconstruction of (D, σ, Γ) is unique up to transformations that leave (χ, q) invariant.*

Proof. Let us start with (a). As in the derivation of (10.7), we obtain

$$\begin{aligned} -\nabla \cdot (Du_1^2)\nabla \frac{H_2}{H_1} &= 0, \qquad \text{in } X \\ Du_1^2 &= D_{|\partial X} g_1^2, \quad \text{on } \partial X. \end{aligned} \tag{10.9}$$

This is a transport equation in conservative form for Du_1^2. More precisely, this is a transport equation $\nabla \cdot \rho\tilde{\beta} = 0$ for ρ with $\rho_{|\partial X} = 1$ and $\tilde{\beta} = \chi^2\beta = (Du_1^2)\nabla\frac{H_2}{H_1}$. Since $\tilde{\beta} \in W^{1,\infty}(X)$ and is divergence free, the above equation for ρ admits the unique solution $\rho \equiv 1$ since (10.6) holds. Indeed, we find that $\nabla\cdot(\rho-1)^2\tilde{\beta} = 0$ by application of the chain rule with $\rho_{|\partial X} - 1 = 0$ on ∂X. Upon multiplying the equation by $\frac{H_2}{H_1}$ and integrating by parts, we find

$$\int_X (\rho - 1)^2 \chi^2 H_1^2 \Big| \nabla \frac{H_2}{H_1} \Big|^2 dx = 0.$$

Using (10.6), we deduce that $\rho \equiv 1$. This proves that Du_1^2 is uniquely determined. Dividing by $H_1^2 = (\Gamma\sigma)^2 u_1^2$, this means that $\chi > 0$ is uniquely determined. Note that we do not need the full $W^{1,\infty}(X)$ regularity of β. However, we still need a condition of the form (10.6) to conclude that the solution to the transport equation is unique.

Let now g be an arbitrary boundary condition and let u be the solution to (10.4) and $H = \mathcal{H}g$ defined by (10.3). Replacing H_2 above by H yields

$$\begin{aligned} -\nabla \cdot \chi^2 H_1^2 \nabla \frac{H}{H_1} &= 0, \qquad \text{in } X \\ H &= \Gamma_{|\partial X}\sigma_{|\partial X} g, \quad \text{on } \partial X. \end{aligned} \tag{10.10}$$

This is a well-defined elliptic equation with a unique solution $H \in H^1(X)$ for $g \in H^{\frac{1}{2}}(\partial X)$. This proves that $\mathcal{H}$ is uniquely determined by (H_1, H_2).

Let us next prove (b). We have already seen that χ was determined by (H_1, H_2), which is clearly determined by $\mathcal{H}$. Moreover, define $v = \sqrt{D}u_1$, which is also uniquely determined based on the results in (a). Define

$$q = \frac{-\Delta v}{v} = -\frac{\Delta(\sqrt{D}u_1)}{\sqrt{D}u_1}.$$

Since u_1 is bounded from below, is sufficiently smooth, and solves (10.4), a routine calculation shows that q is given by (10.8).

Finally, we prove (c). Since q is known, we can solve

$$(\Delta + q)v_j = 0, \quad X, \qquad v_j = \sqrt{D_{|\partial X}}g_j \quad \partial X, \qquad j = 1, 2.$$

Because q is of the specific form (10.8) as a prescribed functional of (D, σ, Γ), it is known that $(\Delta + q)$ does not admit 0 as a (Dirichlet) eigenvalue, for otherwise, 0 would also be a (Dirichlet) eigenvalue of the elliptic operator

$$(-\nabla \cdot D\nabla + \sigma)\cdot = (-\sqrt{D}(\Delta + q)\sqrt{D})\cdot. \tag{10.11}$$

The latter calculation is the standard Liouville transformation allowing us to replace an elliptic operator by a Schrödinger operator. Thus v_j is uniquely determined for $j = 1, 2$. Now,

$$H_j = \Gamma\sigma u_j = \frac{\Gamma\sigma}{\sqrt{D}} v_j = \frac{v_j}{\chi}, \qquad j = 1, 2,$$

and is therefore uniquely determined by (χ, q). □

On the reconstruction of two coefficients. The above result shows that the unique reconstruction of (D, σ, Γ) is not possible even from knowledge of the full measurement operator $\mathcal{H}$ defined in Theorem 10.1.2. We therefore face this peculiar situation that two well-chosen illuminations uniquely determine the functionals (χ, q) but that acquiring additional measurements does not provide any new information, at least in the absence of noise in the data. However, if one coefficient in (D, σ, Γ) is known, then we have the following positive result that the other two coefficients are uniquely determined:

Corollary 10.1.3 *Under the hypotheses of the previous theorem, let* (χ, q) *in* (10.8) *be known. Then*

(a) If Γ *is known, then* (D, σ) *are uniquely determined.*

(b) If D *is known, then* (σ, Γ) *are uniquely determined.*

(c) If σ *is known, then* (D, Γ) *are uniquely determined.*

Proof. (a) is probably the most practical case as Γ is often assumed to be constant. Since Γ is known, then so is $\Gamma\chi = \sqrt{D}/\sigma$ so that we have the elliptic equation for $\sqrt{D}$:

$$(\Delta + q)\sqrt{D} + \frac{1}{\Gamma\chi} = 0, \quad X, \qquad \sqrt{D}_{|\partial X} = \sqrt{D_{|\partial X}}, \quad \partial X. \tag{10.12}$$

Again, because of the specific form of q, $(\Delta + q)$ is invertible and the above equation admits a unique solution. Once $\sqrt{D}$, hence D, is known, then so is $\sigma = \frac{\sqrt{D}}{\Gamma\chi}$.

If D is known in (b), then σ is known from q and Γ is known from χ.

Finally in (c), we obtain that from the expression for q that

$$\sqrt{D}(\Delta + q)\sqrt{D} + \sigma = 0 \quad X, \qquad \sqrt{D}_{|\partial X} = \sqrt{D_{|\partial X}}, \quad \partial X. \tag{10.13}$$

We need to prove a uniqueness result for the above nonlinear equation for $\sqrt{D}$. Let us assume that $\sqrt{D}$ and another solution $\tau\sqrt{D}$ for $0 < \tau(x)$ satisfy the above equation for σ fixed. We have

$$-\sqrt{D}(\Delta + q)\sqrt{D}\tau - \frac{\sigma}{\tau} = 0 \quad X.$$

Thanks to (10.11), this implies the following equation for τ:

$$-\nabla \cdot D\nabla\tau + \sigma(\tau - \frac{1}{\tau}) = 0, \quad X, \qquad \tau = 1, \quad \partial X.$$

Upon multiplying by $\tau - 1$ and integrating by parts, we find that

$$\int_X D|\nabla(\tau - 1)|^2 dx + \int_X \sigma|\tau - 1|^2 \frac{\tau + 1}{\tau} dx = 0.$$

Since $\tau > 0$, we deduce from the above that $\tau \equiv 1$ and that D is uniquely determined by q. We then retrieve Γ from knowledge of χ. □

Reconstruction formulas. Note that the above uniqueness results are *constructive*. In all cases, we need to solve the transport equation for χ:

$$-\nabla \cdot (\chi^2 \beta) = 0 \qquad \text{in } X, \qquad \chi_{|\partial X} \text{ known on } \partial X, \tag{10.14}$$

with β the vector field defined in (10.5). This uniquely defines $\chi > 0$. Then we find that

$$q(x) = -\frac{\Delta(H_1\chi)}{H_1\chi} = -\frac{\Delta(H_2\chi)}{H_2\chi}. \tag{10.15}$$

This provides explicit reconstructions for (χ, q). In case (b), no further equation needs to be solved. In cases (a) and (c), we need to solve an elliptic equation for $\sqrt{D}$, which is the linear equation (10.12) in (a) and the nonlinear equation (10.13) in (c). These are the steps that will be implemented in the sections on numerical simulations below.

Stability of the solution of the transport equation. Before presenting our numerical framework, we derive a stability result for the reconstruction of χ. A similar result was obtained in [16] by using the stability of the method of characteristics to solve ordinary differential equations. Here, we present a stability result that is directly obtained from the PDE (10.9) and is similar in spirit to estimates obtained in [3] and to the notion of renormalization property of transport equation [20]. Similar stability results can be obtained for q and then for (D, σ, Γ) depending on the reconstruction considered.

Theorem 10.1.4 *We assume that the hypotheses of Theorem 10.1.2 hold. Let $H = (H_1, H_2)$ be the measurements corresponding to the coefficients (D, σ, Γ) for which hypothesis (iii) holds. Let $\tilde{H} = (\tilde{H}_1, \tilde{H}_2)$ be the measurements corresponding to the same illuminations (g_1, g_2) with another set of coefficients $(\tilde{D}, \tilde{\sigma}, \tilde{\Gamma})$ such that (i) and (ii) still hold. Then we find that*

$$\|\chi - \tilde{\chi}\|_{L^p(X)} \leq C \|H - \tilde{H}\|^{\frac{1}{2}}_{(W^{1,\frac{p}{2}}(X))^2}, \qquad \text{for all } 2 \leq p < \infty. \tag{10.16}$$

Let us assume, moreover, that $\gamma(x)$ is of class $C^3(\bar{X})$. Then we have the estimate

$$\|\chi - \tilde{\chi}\|_{L^p(X)} \leq C \|H - \tilde{H}\|^{\frac{1}{3}}_{(L^{\frac{p}{2}}(X))^2}, \qquad \text{for all } 2 \leq p < \infty. \tag{10.17}$$

By interpolation [2], the latter result implies that

$$\|\chi - \tilde{\chi}\|_{L^\infty(X)} \leq C \|H - \tilde{H}\|^{\frac{p}{3(d+p)}}_{(L^{\frac{p}{2}}(X))^2}, \qquad \text{for all } 2 \leq p < \infty. \tag{10.18}$$

We may for instance choose $p = 4$ above to measure the noise level in the measurement H in the square integrable norm when noise is described by its power spectrum in the Fourier domain.

Proof. Define $\nu = \chi^2$ and $\tilde{\nu} = \tilde{\chi}^2$ with χ defined in (10.8) and β and $\tilde{\beta}$ as in (10.5). Then we find that

$$\nabla \cdot \frac{\nu - \tilde{\nu}}{\nu}(\nu\beta) + \nabla \cdot \tilde{\nu}(\beta - \tilde{\beta}) = 0.$$

Note that $\nu\beta = \chi^2 H_1^2 \nabla \frac{H_2}{H_1}$ is a divergence-free field. Let φ be a twice differentiable, non-negative, function from $\mathbb{R}$ to $\mathbb{R}$ with $\varphi(0) = \varphi'(0) = 0$. Then we find that

$$\nabla \cdot \varphi\Big(\frac{\nu - \tilde{\nu}}{\nu}\Big)(\nu\beta) + \varphi'\Big(\frac{\nu - \tilde{\nu}}{\nu}\Big)\nabla \cdot \tilde{\nu}(\beta - \tilde{\beta}) = 0.$$

Let us multiply this equation by a test function $\zeta \in H^1(X)$ and integrate by parts. Since $\nu = \nu'$ on ∂X, we find

$$\int_X \varphi\Big(\frac{\nu - \tilde{\nu}}{\nu}\Big)\nu\beta \cdot \nabla\zeta dx + \int_X \tilde{\nu}(\beta - \tilde{\beta})\nabla \cdot \Big[\zeta\varphi'\Big(\frac{\nu - \tilde{\nu}}{\nu}\Big)\Big]dx = 0.$$

Upon choosing $\zeta = \frac{H_2}{H_1}$, we find

$$\int_X \varphi\nu H_1^2 \Big|\nabla \frac{H_2}{H_1}\Big|^2 dx + \int_X \tilde{\nu}(\beta - \tilde{\beta}) \cdot \nabla \frac{H_2}{H_1}\varphi' dx + \int_X \tilde{\nu}(\beta - \tilde{\beta}) \cdot \nabla \frac{\nu - \tilde{\nu}}{\nu}\frac{H_2}{H_1}\varphi'' dx = 0.$$

Above, φ stands for $\varphi(\frac{\nu - \tilde{\nu}}{\nu})$ in all integrals. By assumption on the coefficients, $\nabla \frac{\nu - \tilde{\nu}}{\nu}$ is bounded a.e.. This is one of our main motivations for assuming that the optical coefficients are Lipschitz. The middle term is seen to be smaller than the third term and so we focus on the latter one. Upon taking $\varphi(x) = |x|^p$ for $p \geq 2$ and using assumption (iii), we find that

$$\|\nu - \tilde{\nu}\|^p_{L^p(X)} \leq C \int_X |\beta - \tilde{\beta}||\nu - \tilde{\nu}|^{p-2} dx.$$

By an application of the Hölder inequality, we deduce that

$$\|\nu - \tilde{\nu}\|_{L^p(X)} \leq C\|\beta - \tilde{\beta}\|^{\frac{1}{2}}_{L^{\frac{p}{2}}(X)}.$$

We next write $\beta - \tilde{\beta} = (H_1 - \tilde{H}_1)\nabla H_2 + \tilde{H}_1(\nabla(H_2 - \tilde{H}_2) - \dots$ and use the fact that the solutions to (10.4) and the coefficients are in $W^{1,\infty}(X)$ to conclude that (10.16) holds.

The other results are obtained by regularity theory and interpolation. Indeed from regularity results in [22] with coefficients in $W^{1,\infty}(X)$, we find that the solutions to (10.4) are of class $W^{3,q}(X)$ for all $1 \leq q < \infty$. Since the coefficient γ is of class $C^3(\bar{X})$, then the measurements H_j are of class $W^{3,q}(X)$ for all $1 \leq q < \infty$. Standard Sobolev estimates [22] show that

$$\|H_j - \tilde{H}_j\|_{W^{1,q}(X)} \leq C\|H_j - \tilde{H}_j\|^{\frac{2}{3}}_{L^q(X)}\|H_j - \tilde{H}_j\|^{\frac{1}{3}}_{W^{3,q}(X)}.$$

The last term is bounded by a constant, which gives (10.17) for $q = \frac{p}{2}$. Another interpolation result states that

$$\|\varphi\|_\infty \leq \|\nabla\varphi\|^\theta_\infty \|\varphi\|^{1-\theta}_p, \qquad \theta = \frac{d}{d+p}.$$

This provides the stability result in the uniform norm (10.18). □

On the reconstruction of one coefficient. We conclude our theoretical section by the reconstruction of one coefficient when the other two coefficients are known. This is significantly simpler than the reconstruction of two coefficients. In none of the cases do we need to solve a transport equation involving the vector field β. The latter was obtained by eliminating σ from the elliptic equation, which is no longer necessary when two coefficients are already known.

When only Γ is unknown, then we solve (10.4) for u and then construct $\Gamma = \dfrac{H}{\sigma u}$.

When only σ is unknown, then we solve

$$\begin{array}{rcll} -\nabla \cdot D\nabla u(\mathbf{x}) + \dfrac{H}{\Gamma} & = & 0, & \text{in } X \\ u(\mathbf{x}) & = & g(\mathbf{x}), & \text{on } \partial X \end{array}, \qquad \sigma = \frac{H}{\Gamma u}. \tag{10.19}$$

When only D is unknown, we obtain $u = \frac{H}{\sigma\Gamma}$ and then the above elliptic equation in (10.19) with $D_{|\partial X}$ known is a transport equation for D. As soon as ∇u is a sufficiently smooth, non-vanishing vector field, then D is uniquely determined by the above linear equation.

10.2 Ultrasound Modulation Tomography

The preceding section concerned the photo-acoustic effect. We now consider another physical mechanism that couples optical waves with ultrasound, namely the ultrasound modulation effect.

We consider the following elliptic equation

$$-\nabla \cdot \sigma(x)\nabla u = 0 \quad \text{in } X, \qquad u = g \quad \text{on } \partial X. \tag{10.20}$$

Here, σ is the unknown diffusion coefficient, which we assume is a real-valued, scalar, function defined on a domain $X \subset \mathbb{R}^n$ for $n = 2$ or $n = 3$. We assume that σ is bounded above and below by positive constants so that the above equation admits a unique solution. We also assume that σ is sufficiently smooth so that the solution to the above equation is continuously differentiable on $\bar{X}$, the closure of X [22]. We denote by ∂X the boundary of X and by $g(x)$ the imposed (sufficiently smooth) Dirichlet boundary conditions.

As we have seen already, the coefficient $\sigma(x)$ may model the electrical conductivity in the setting of electrical impedance tomography (EIT) or a diffusion coefficient of particles (photons) in the setting of optical tomography (OT). Both EIT and OT are modalities with high contrast, in the sense that $\sigma(x)$ takes different values in different tissues and allows one to discriminate between healthy and non-healthy tissues. In OT, high contrasts are mostly observed in the absorption coefficient, which is not modeled here [15].

A methodology to couple high contrast with high resolution consists of perturbing the diffusion coefficient acoustically. Let an acoustic signal propagate through the domain. We assume here that the sound speed is constant and that the acoustic signal is a plane wave of the form $p\cos(k \cdot x + \varphi)$ where p is the amplitude of the acoustic signal, k its wave-number and φ an additional phase. The acoustic signal modifies the properties of

the diffusion equation. We assume that such an effect is small and that the coefficient in (10.20) is modified as

$$\sigma_\varepsilon(x) = \sigma(x)(1 + \zeta\varepsilon\mathfrak{c}), \tag{10.21}$$

where we have defined $\mathfrak{c} = \mathfrak{c}(x) = \cos(k \cdot x + \varphi)$ and where $\varepsilon = p\Gamma$ is the product of the acoustic amplitude $p \in \mathbb{R}$ and a measure $\Gamma > 0$ of the coupling between the acoustic signal and the modulations of the constitutive parameter in (10.20). We assume that $\varepsilon \ll 1$ so that the influence of the acoustic signal on σ_ε admits an asymptotic expansion that we truncated at the second order as displayed in (10.21). The size of the terms in the expansion are physically characterized by ζ and depend on the specific application.

Let u and v be solutions of (10.20) with fixed boundary conditions g and h, respectively. When the acoustic field is turned on, the coefficients are modified as described in (10.21) and we denote by u_ε and v_ε the corresponding solutions. Note that $u_{-\varepsilon}$ is the solution obtained by changing the sign of p or equivalently by replacing φ by $\varphi + \pi$.

By the standard continuity of the solution to (10.20) with respect to changes in the coefficients and regular perturbation arguments, we find that $u_\varepsilon = u_0 + \varepsilon u_1 + O(\varepsilon^2)$. Let us multiply the equation for u_ε by $v_{-\varepsilon}$ and the equation for $v_{-\varepsilon}$ by u_ε, subtract the resulting equalities, and use standard integrations by parts. We obtain that

$$\int_X (\sigma_\varepsilon - \sigma_{-\varepsilon})\nabla u_\varepsilon \cdot \nabla v_{-\varepsilon} dx = \int_{\partial X} \sigma_{-\varepsilon}\frac{\partial v_{-\varepsilon}}{\partial \nu} u_\varepsilon - \sigma_\varepsilon \frac{\partial u_\varepsilon}{\partial \nu} v_{-\varepsilon} ds(x). \tag{10.22}$$

Here, $ds(x)$ is the standard surface measure on ∂X. We assume that $\sigma_\varepsilon \partial_\nu u_\varepsilon$ and $\sigma_\varepsilon \partial_\nu v_\varepsilon$ are measured on ∂X, at least on the support of $v_\varepsilon = h$ and $u_\varepsilon = g$, respectively, for all values ε of interest. Note that the above equation holds if the Dirichlet boundary conditions are replaced by Neumann boundary conditions. Let us define

$$J_\varepsilon := \frac{1}{2}\int_{\partial X} \sigma_{-\varepsilon}\frac{\partial v_{-\varepsilon}}{\partial \nu} u_\varepsilon - \sigma_\varepsilon \frac{\partial u_\varepsilon}{\partial \nu} v_{-\varepsilon} ds(x) \;=\; \varepsilon J_1 + \varepsilon^2 J_2 + O(\varepsilon^3). \tag{10.23}$$

We assume that the real valued functions $J_m = J_m(k, \varphi)$ are known (measured functions). Notice that such knowledge is based on the physical boundary measurement of the Cauchy data of the form $(u_\varepsilon, \sigma_\varepsilon \partial_\nu u_\varepsilon)$ and $(v_\varepsilon, \sigma_\varepsilon \partial_\nu v_\varepsilon)$ on ∂X.

Equating like powers of ε, we find that at leading order

$$\int_X \big[\zeta\sigma(x)\nabla u_0 \cdot \nabla v_0(x)\big] \cos(k \cdot x + \varphi) dx = J_1(k, \varphi). \tag{10.24}$$

This may be acquired for all $k \in \mathbb{R}^n$ and $\varphi = 0, \frac{\pi}{2}$, and hence provides the Fourier transform of

$$H[u_0, v_0](x) = \zeta\sigma(x)\nabla u_0 \cdot \nabla v_0(x). \tag{10.25}$$

Note that when $v_\varepsilon = u_\varepsilon$, then we find from the expression in (10.22) that $J_2 = 0$ in (10.23) so that the expression for J_1 may be obtained from available measurements in (10.23) with an accuracy of order $O(\varepsilon^2)$. Note also that

$$H[u_0, v_0](x) = \frac{1}{4}\big(H[u_0 + v_0, u_0 + v_0] - H[u_0 - v_0, u_0 - v_0]\big)$$

by polarization. In other words, the limiting measurements (for small ε) in (10.25) may also be obtained by considering expressions of the form (10.22) with $u_\varepsilon = v_\varepsilon$.

In the setting of optical tomography, the coefficient σ_ε in (10.21) takes the form

$$\sigma_\varepsilon(x) = \frac{\tilde{\sigma}_\varepsilon}{c_\varepsilon^{d-1}}(x),$$

where $\tilde{\sigma}_\varepsilon$ is the diffusion coefficient, c_ε is the light speed, and n is spatial dimension. When the pressure field is turned on, the location of the scatterers is modified by compression and dilation. Since the diffusion coefficient is inversely proportional to the scattering coefficient, we find that

$$\frac{1}{\sigma_\varepsilon(x)} = \frac{1}{\sigma(x)}\big(1 + \varepsilon\mathfrak{c}(x)\big).$$

Moreover, the pressure field changes the index of refraction (the speed) of light as follows

$$c_\varepsilon(x) = c(x)(1 + \gamma\varepsilon\mathfrak{c}(x)),$$

where γ is a constant (roughly equal to $\frac{1}{3}$ for water). This shows that

$$\zeta = -(1 + (d-1)\gamma). \tag{10.26}$$

In the setting of electrical impedance tomography, we simply assume that ζ models the coupling between the acoustic signal and the change in the electrical conductivity of the underlying material. The value of ζ thus depends on the application.

Assuming the validity of the above derivation, the objective of ultrasound modulated optical tomography (UMOT) and ultrasound modulated electrical impedance tomography (UMEIT) is to reconstruct the coefficient $\sigma(x)$ from knowledge of the interior functionals

$$H_{ij}(x) = \sigma(x)\nabla u_i(x) \cdot \nabla u_j(x), \qquad 1 \le i, j \le m, \tag{10.27}$$

where u_j is the solution to the equation

$$\begin{aligned} \nabla \cdot (\sigma \nabla u_i) &= 0 \quad X, \\ u_i &= g_i \quad \partial X, \qquad 1 \le i \le m, \end{aligned} \tag{10.28}$$

for appropriate choices of the boundary conditions g_i on ∂X. In practice, the ultrasound modulation effect is extremely weak because the coupling coefficient Γ is small. The above inverse problem is therefore more a theoretical problem than it is a practical problem at the moment. However, its mathematical analysis shows that UMOT and UMEIT are well posed problems unlike the (non ultrasound-modulated) OT and EIT problems.

Reformulation and Elimination. The main strategy we use to reconstruct σ from knowledge of $H = \{H_{ij}\}_{ij}$ is first to write equations for the quantities $S_i := \sqrt{\sigma}\nabla u_i$ that are independent of σ and then to show that σ is uniquely determined when S_i is known.

We define:

$$S_i := \sqrt{\sigma}\nabla u_i, \quad 1 \le i \le m, \qquad F := \nabla(\log\sqrt{\sigma}) = \frac{1}{2}\nabla \log\sigma. \tag{10.29}$$

Using the equations (10.28), the definitions (10.29) and the fact that $\sigma^{-\frac{1}{2}} S_i$ is a gradient, we obtain:

$$\nabla \cdot (\sqrt{\sigma} S_j) = 0 \quad \Leftrightarrow \quad \nabla \cdot S_j + F \cdot S_j = 0, \tag{10.30}$$

$$n = 2 : \quad \left[\nabla, \frac{1}{\sqrt{\sigma}} S_j\right] = 0 \quad \Leftrightarrow \quad [\nabla, S_j] - [F, S_j] = 0, \tag{10.31}$$

$$n = 3 : \quad \nabla \times \left(\frac{1}{\sqrt{\sigma}} S_j\right) = 0 \quad \Leftrightarrow \quad \nabla \times S_j - F \times S_j = 0, \tag{10.32}$$

where we have defined for $n = 2$ the product $[A, B] := A_x B_y - A_y B_x$ and $[\nabla, A] := \partial_x A_y - \partial_y A_x$ for smooth vector fields A and B, while for $n = 3$, $\times$ is the standard cross product.

We now wish to *eliminate* F from such equations and get a closed form equation for the vectors S_i with sources that only involve the known matrix H. Such an elimination requires that we find n vectors S_i that form a basis of $\mathbb{R}^n$ for $n = 2, 3$. In dimension $n = 2$, it is not difficult to find boundary conditions g_i guaranteeing the existence of such a basis for all $x \in X$. In dimension $n \geq 3$ and by appropriate choice of the illuminations g_i, we are able to construct such bases on subset of $\mathbb{R}^n$ that cover X using triplets of vectors that depend on the subset. We thus consider the above equations for S_i and F over an open subset $\Omega \subset X$, over which we make the assumption that

$$\inf_{x \in \Omega} \det(S_1(x), \ldots, S_n(x)) = c_0 > 0, \tag{10.33}$$

where the n vectors S_j for $1 \leq j \leq n$ are chosen among the m vectors considered in (10.27). Since the determinants are central in our derivations, we define

$$d(x) := \det(S_1(x), S_2(x)), \ n = 2 \quad \text{and} \quad D(x) := \det(S_1(x), S_2(x), S_3(x)), \ n = 3, \tag{10.34}$$

for all $x \in \Omega$. Note that $\det(S_1(x), \ldots, S_n(x)) = (\det H)^{\frac{1}{2}}(x) \geq c_0$ on Ω.

In two dimensions, the constraint (10.33) is satisfied over the whole domain $\Omega = X$ for a large class of boundary conditions g_i and we then choose $m = n = 2$ as shown in [5, Theorem 4]. For instance, we may choose $g_j(x_1, x_2) = x_j$ on ∂X and we are then guaranteed that $(x_1, x_2) \mapsto (u_1, u_2)$ is a diffeomorphism from X to its image so that the determinant is clearly signed, and by continuity (assuming that σ is sufficiently smooth so that gradients are continuous functions, which we also assume for the rest of the paper [22]), bounded from below by a positive constant. In three dimensions, there is no known guarantee that the determinant of three solutions with prescribed boundary conditions remains positive throughout the domain X independent of the conductivity σ. For boundary conditions of the form $g_j(x) = x_j$ on ∂X for $1 \leq j \leq 3$, it is known that the determinant can change signs for some conductivities [17].

Yet, (10.33) is an important constraint in the derivation of our results. By means of specific complex geometric optics solutions, it is possible to construct boundary conditions g_i such that (10.33) is valid on subsets of the domain X. More precisely, we construct g_i for $1 \leq i \leq 4$ so that $\det(S_1(x), S_2(x), S_3(x)) \geq c_0 > 0$ on $\cup_{k=1}^{M} \Omega_{2k}$ and $\det(S_1(x), S_2(x), S_4(x)) \geq c_0 > 0$ on $\cup_{k=1}^{M} \Omega_{2k-1}$ with all domains Ω_k simply connected, open and such that $X \subset \cup_{k=1}^{2M} \Omega_k$. In other words, we can choose $m = 4$ boundary conditions such that X is decomposed into a superposition of overlapping simply connected

subsets where the determinant of three out of the four vectors is positive. We refer the reader to [9] for the details.

Although our goal is ultimately to recover only the conductivity σ, our approach requires the reconstruction of the vector fields S_i, $1 \leq i \leq m$ and F. Since the data $H(x)$ give us $S(x)$ (the matrix with columns given by the vectors S_i) up to an $SO_n(\mathbb{R})$-valued function $R(x)$ (i.e., $R(x)$ is a rotation matrix), the unknowns are now the functions R and F, which are of dimension $\frac{n(n-1)}{2}$ and n, respectively. The local reconstruction of R and F proceeds as follows: we first derive divergence and curl equations for the column vectors R_j of the function $R : \Omega \mapsto SO_n(\mathbb{R})$ with right-hand-sides involving F, R_j and the data. Using structural properties of $SO_n(\mathbb{R})$ satisfied at every $x \in \Omega$, we then derive an equation for F of the form

$$F = \frac{1}{n}\left(\frac{1}{2}\nabla \log\det H + \sum_{1\leq i,j\leq n}((V_{ij}+V_{ji})\cdot R_i)R_j\right), \quad n = 2, 3, \tag{10.35}$$

where the vector fields V_{ij} depend only on the data. Plugging this equation back into the divergence and curl system closes the system for the R_j's. From this closed system, we derive gradient-type equations for the vectors R_j, where the right-hand sides either depend only on the data (case $n = 2$), or also depend polynomially on the R_j's (case $n = 3$). In either case, these gradient equations can be integrated along segments of the form $[x_0, x]$ for $x, x_0 \in \Omega$, parameterized by the curve

$$\gamma_{x_0,x} : [0, 1] \ni t \mapsto \gamma_{x_0,x}(t) = (1 - t)x_0 + tx \in \Omega, \tag{10.36}$$

in order to reconstruct locally the R_j's at x from knowledge of their value at x_0. Once the function R is locally reconstructed around x_0, we can reconstruct $\sigma(x)$ around x_0 from the knowledge of $\sigma(x_0)$ and integrating (10.35) along segments $[x_0, x]$. This local construction can be made global by patching local reconstructions together. Again, we refer the reader to [9] for the details.

Equations for rotations in dimensions $n = 2$ and $n = 3$. Knowledge of the matrix $H(x)$ together with the determinant condition (10.33) allow us to reconstruct the matrix-valued function $S(x) := [S_1(x)|\dots|S_n(x)]$ up to an $SO_n(\mathbb{R})$-valued function $R(x) = [R_1(x)|\dots|R_n(x)]$. This fact can be seen for instance by noticing that the orientation-preserving Gram-Schmidt procedure that creates the orthonormal R_j's from the S_i's with $\det R \det S > 0$ only involves coefficients that depend on the inner products $S_i \cdot S_j = H_{ij}$, which are indeed known. Note that $SH^{-\frac{1}{2}}(x)$ is also a rotation-valued function on Ω.

We take $T(x) = \{t_{ij}(x)\}_{1\leq i,j\leq n}$ a matrix-valued function on Ω that satisfies the properties

$$T^T(x)T(x) = H^{-1}(x) \quad \text{and} \quad 0 < \det T(x) = (\det H(x))^{-\frac{1}{2}}, \quad x \in \Omega. \tag{10.37}$$

We also define $T^{-1} = \{t^{ij}\}_{1\leq i,j\leq n}$ and the vector fields

$$V_{ij} := \nabla(t_{ik})t^{kj}, \quad \text{i.e.} \quad V^l_{ij} := \partial_l(t_{ik})t^{kj}, \quad 1 \leq i, j, l \leq n. \tag{10.38}$$

Here and below, repeated indices are summed over.

Two examples for T are the symmetric $T = H^{-\frac{1}{2}}$ and the lower-triangular T obtained by the Gram Schmidt procedure

$$T = \{t_{ij}\}_{1\le i,j\le 3} = \begin{bmatrix} H_{11}^{-\frac{1}{2}} & 0 & 0 \\ -H_{12}H_{11}^{-\frac{1}{2}}d^{-1} & H_{11}^{\frac{1}{2}}d^{-1} & 0 \\ (H_{12}H_{23} - H_{22}H_{13})(dD)^{-1} & (H_{12}H_{13} - H_{11}H_{23})(dD)^{-1} & dD^{-1} \end{bmatrix}, \tag{10.39}$$

$$\text{with} \quad d := (H_{11}H_{22} - H_{12}^2)^{\frac{1}{2}} \quad \text{and} \quad D = (\det H)^{\frac{1}{2}}.$$

The vector fields V_{ij} defined in (10.38) then take the following expression

$$\{V_{ij}\}_{1\le i,j\le 3} = \begin{bmatrix} \nabla \log t_{11} & 0 & 0 \\ \frac{t_{22}}{t_{11}}\nabla\frac{t_{21}}{t_{22}} & \nabla \log t_{22} & 0 \\ \frac{t_{33}}{t_{11}}\nabla\frac{t_{31}}{t_{33}} - \frac{t_{21}t_{33}}{t_{11}t_{22}}\nabla\frac{t_{32}}{t_{33}} & \frac{t_{33}}{t_{22}}\nabla\frac{t_{32}}{t_{33}} & \nabla \log t_{33} \end{bmatrix}. \tag{10.40}$$

When $n = 2$, T and V_{ij} are given by the top-left 2×2 blocs of (10.39) and (10.40), respectively.

Denote $\widetilde{T}$ the matrix constructed from $\widetilde{H}$ in the same manner as T. We impose the existence of a constant $C_T > 0$ such that

$$\|T - \widetilde{T}\|_{W^{1,\infty}(\Omega)} \le C_T \|H - \widetilde{H}\|_{W^{1,\infty}(\Omega)}, \tag{10.41}$$

here C_T only depends on H and $\widetilde{H}$ and is completely determined by the way we construct T and $\widetilde{T}$. It is relatively straightforward to check that the above relation holds for T chosen by the Gram Schmidt procedure.

From S and T, let us now build the matrix R by defining

$$R(x) := S(x)T(x)^T, \quad \text{i.e. } R_j(x) = t_{ij}(x)S_j(x), \quad 1 \le i \le n, \quad x \in \Omega. \tag{10.42}$$

From conditions (10.37) and the fact that $S(x)^T S(x) = H(x)$ for every $x \in \Omega$, the matrix R (10.42) satisfies $R^T(x)R(x) = I_n$ for all $x \in \Omega$, as well as $\det R(x) = 1$, thus $R(x) \in SO_n(\mathbb{R})$ for all $x \in \Omega$. Moreover, plugging relation (10.42) into equation (10.30) and using the vector calculus identity $\nabla \cdot (fV) = \nabla f \cdot V + f\nabla \cdot V$, we obtain

$$\nabla \cdot R_i = \nabla \cdot (t_{ij}S_j) = (\nabla t_{ij}) \cdot S_j + t_{ij}\nabla \cdot S_j = (\nabla t_{ij})t^{jk} \cdot R_k - t_{ij}F \cdot S_j.$$

Thus the R_i's satisfy the following divergence equation:

$$\nabla \cdot R_i = V_{ik} \cdot R_k - F \cdot R_i, \quad 1 \le i \le n. \tag{10.43}$$

In a similar manner, one can derive the following curl-type equations for the R_i's:

$$n = 2 \,:\, [\nabla, R_i] = [V_{ik}, R_k] + [F, R_i], \quad i = 1, 2, \tag{10.44}$$

$$n = 3 \,:\, \nabla \times R_i = V_{ik} \times R_k + F \times R_i, \quad i = 1, 2, 3. \tag{10.45}$$

We now show that the redundancies in the systems (10.43)-(10.44) when $n = 2$ and (10.43)-(10.45) when $n = 3$ allow us to derive formula (10.35) for F in terms of the R_i's and the data. We prove this formula for $n = 2$ and refer the reader to [9] for $n = 3$.

Elimination of source term and proof of formula (10.35) **in dimension** $n = 2$**.** First notice that because $R \in SO_2(\mathbb{R})$, we have the relations

$$JR_i = \varepsilon_{ij} R_j, \qquad (i,j) \in I_2 := \{(1,2),(2,1)\}, \tag{10.46}$$

where $\varepsilon_{12} = -\varepsilon_{21} = 1$ and we have defined $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. In particular, this implies that

$$\nabla \cdot R_i = [\nabla, JR_i] = \varepsilon_{ij}[\nabla, R_j], \qquad (i,j) \in I_2,$$

as well as the relations, for any vector field A,

$$[A, R_i] = JA \cdot R_i = -A \cdot JR_i = -\varepsilon_{ij} A \cdot R_j, \qquad (i,j) \in I_2.$$

Together with the system of equations (10.43)-(10.44), these relations allow us to get the components of F in the basis (R_1, R_2):

$$\begin{aligned} F \cdot R_i &= -\nabla \cdot R_i + V_{ii} \cdot R_i + V_{ij} \cdot R_j \\ &= -\varepsilon_{ij}[\nabla, R_j] + V_{ii} \cdot R_i + V_{ij} \cdot R_j \\ &= -\varepsilon_{ij}([F, R_j] + [V_{ji}, R_i] + [V_{jj}, R_j]) + V_{ii} \cdot R_i + V_{ij} \cdot R_j \\ &= -F \cdot R_i + V_{ji} \cdot R_j - V_{jj} \cdot R_i + V_{ii} \cdot R_i + V_{ij} \cdot R_j, \end{aligned}$$

and hence the formula

$$2F \cdot R_i = -(V_{ii} + V_{jj}) \cdot R_i + 2V_{ii} \cdot R_i + (V_{ij} + V_{ji}) \cdot R_j, \tag{10.47}$$

for $(i,j) \in I_2$. We then obtain equation (10.35) by plugging (10.47) into the relation $F = (F \cdot R_1)R_1 + (F \cdot R_2)R_2$, and using the following identity

$$-V_{11} - V_{22} = \nabla \log d. \tag{10.48}$$

Indeed, we first notice that for any $k = 1 \dots n$, the function $T(x)$ satisfies locally the (tautological) relation

$$\partial_k T(x) = V^k(x) T(x),$$

where we have defined $V^k := \{\partial_k(t_{il}) t^{lj}\}_{i,j} = (\partial_k T) T^{-1}$. Therefore by Liouville's formula, we obtain that

$$\partial_k(\log \det T) = \text{trace } (V^k) = \sum_{i=1}^{n} V^k_{ii}.$$

Finally noting that $\det T = d^{-1}$ when $n = 2$, the identity (10.48) is proved component by component. A similar proof holds in dimensions $n \geq 3$ though we do not present it here.

Equation for the rotation matrix R in dimension $n = 2$. Plugging formula (10.35) into (10.43) and (10.44) provides a closed system of equations for the function R. It remains to show that R is then uniquely and stably determined by such equations. Since $SO_2(\mathbb{R})$ is a one-dimensional manifold, $R = [R_1|R_2]$ is described by a $\mathbb{S}^1$-valued function $\theta(x)$ such that, at each point, $R_1(\theta) = (\cos\theta, \sin\theta)^T$ and $R_2(\theta) = JR_1(\theta)$. We wish to derive an equation for $\nabla\theta$. Plugging the expression (10.35) of F into (10.43), we arrive at

$$\nabla \cdot R_i = \frac{1}{2}\left[-(\nabla \log d) \cdot R_i + (V_{ij} - V_{ji}) \cdot R_j\right], \quad (i,j) \in I_2. \tag{10.49}$$

Let us now derive a differential equation for $R = [R_1|R_2]$. The relation $R^T R = I_2$ implies that $R^T \partial_i R \in \mathcal{A}_2(\mathbb{R})$, the space of two-dimensional anti-symmetric matrices. For $i = 1, 2$, i.e., it can be written in the form $R^T \partial_i R = \alpha_i J$, where $\alpha_i = R_2 \cdot \partial_i R_1$. Defining $\vec{\alpha} := \alpha_i \mathbf{e}_i$, it is clear that $\vec{\alpha} \cdot Z = R_2 \cdot [(Z \cdot \nabla)R_1]$ for any vector field Z. For the sequel, we need the following vector calculus identity, which holds for any smooth vector field A

$$\nabla |A|^2 = 2(A \cdot \nabla)A - 2[\nabla, A]JA. \tag{10.50}$$

We decompose $\vec{\alpha}$ in the basis (R_1, R_2) and use equations (10.43)

$$\begin{aligned}\vec{\alpha} &= (\alpha \cdot R_1)R_1 + (\alpha \cdot R_2)R_2 = (R_2 \cdot [(R_1 \cdot \nabla)R_1])R_1 + (R_2 \cdot [(R_2 \cdot \nabla)R_1])R_2 \\ &= (R_2 \cdot [(R_1 \cdot \nabla)R_1])R_1 - (R_1 \cdot [(R_2 \cdot \nabla)R_2])R_2.\end{aligned}$$

We now use (10.50), which, in this case, becomes $(R_i \cdot \nabla)R_i = [\nabla, R_i]JR_i = -(\nabla \cdot R_j)R_j$ for $(i,j) \in I_2$, and then equation (10.49) to obtain that

$$\begin{aligned}\vec{\alpha} &= -(\nabla \cdot R_2)R_1 + (\nabla \cdot R_1)R_2 \\ &= -\frac{1}{2}(-(V_{12} - V_{21}) \cdot R_1 - (\nabla \log d) \cdot R_2)R_1 + \frac{1}{2}(-(\nabla \log d) \cdot R_1 + (V_{12} - V_{21}) \cdot R_2)R_2 \\ &= \frac{1}{2}[V_{12} - V_{21} - (R_2 \otimes R_1 - R_1 \otimes R_2)\nabla \log d] = \frac{1}{2}[V_{12} - V_{21} - J\nabla \log d].\end{aligned}$$

Finally, given the above parameterization $R(\theta)$, one check using the chain rule that $R^T \partial_i R = (\partial_i \theta)J,\ i = 1, 2$. Using the preceding calculations, we obtain that

$$\nabla\theta = \vec{\alpha} = \frac{1}{2}[V_{12} - V_{21} - J\nabla \log d]. \tag{10.51}$$

Since the right-hand side is known, we obtain a (redundant) equation for θ, and hence for R. The situation is more complicated in $n \geq 3$ but a similar result still emerges. Let us assume that σ and ∇u_j are known at a point x_m. Then R and θ are also known at the point x_m. Therefore, we can reconstruct $\theta(x)$ at every $x \in X$ by integrating (10.51) along the segment $[x_m, x]$ parameterized by $\gamma_{x_m,x}$:

$$\begin{aligned}\theta(x) &= \theta(x_m) + \int_0^1 \dot{\gamma}_{x_m,x}(t) \cdot \nabla\theta(\gamma_{x_m,x}(t))\ dt \\ &= \theta(x_m) + \frac{1}{2}(x - x_m) \cdot \int_0^1 (V_{12} - V_{21} - J\nabla \log d)(\gamma_{x_m,x}(t))\ dt.\end{aligned}$$

Once θ is recovered throughout X, one then reconstructs $\sigma(x)$ for all $x \in X$ from the knowledge of $\sigma(x_0)$ for some $x_0 \in X$ and integrating equation (10.35) along the segment $[x_0, x]$, that is

$$\begin{aligned}\log \sigma(x) &= \log \sigma(x_0) + 2 \int_0^1 \dot{\gamma}_{x_0,x}(t) \cdot F(\gamma_{x_0,x}(t))\, dt \\ &= \log \sigma(x_0) + (x - x_0) \cdot \int_0^1 (\nabla \log d + ((V_{ij} + V_{ji}) \cdot R_i) R_j)(\gamma_{x_0,x}(t))\, dt.\end{aligned}$$

From this, we (relatively easily) prove the following result. As we mentioned earlier, we can set $\Omega = X$ when $n = 2$ by choosing illuminations $\mathbf{g} = (g_1, g_2)$ such that (10.33) is satisfied throughout X. Then the above reconstruction procedure yields unique and stable reconstructions, as described in the following theorem.

Theorem 10.2.1 (2D global uniqueness and stability) *Let $(H, \widetilde{H})$ be two data sets with coefficients in $W^{1,\infty}(X)$ corresponding to the same illumination $\mathbf{g} = (g_1, g_2)$ and with determinants satisfying the condition*

$$\inf_{x \in X} (d, \tilde{d}) \geq c_0 > 0, \qquad \text{with} \quad d^2(x) = \det H(x), \quad \tilde{d}^2(x) = \det \tilde{H}(x). \tag{10.52}$$

Let σ and $\tilde{\sigma}$ be the corresponding conductivities and assume that $\sigma(x_0) = \tilde{\sigma}(x_0)$ for some $x_0 \in X$. Then, we have the following stability estimate

$$\| \log \sigma - \log \tilde{\sigma} \|_{W^{1,\infty}(X)} \leq C \| H - \widetilde{H} \|_{W^{1,\infty}(X)}. \tag{10.53}$$

This result also ensures uniqueness, since (10.53) *implies $H = \widetilde{H} \implies \sigma = \tilde{\sigma}$.*

Note that since σ is bounded above and below a priori, the above expression also provides a bound for $\sigma - \tilde{\sigma}$ and not only the logarithms.

10.3 Remarks on hybrid inverse problems

We have briefly presented two hybrid inverse problems above based on the photo-acoustic effect and the ultrasound modulation effect. What characterizes these hybrid inverse problems is that after a preliminary step (involving and inverse wave equation in photoacoustics and an inverse Fourier transform in ultrasound modulation) we obtain an inverse problem with internal functionals of the unknown parameters.

These internal functionals have an immediate advantage: singularities of the unknown coefficients no longer need to be propagated to the boundary of the domain by an elliptic operator that severely damps high frequencies. The main reason for ill-posedness of invertible operators, namely the smoothing property of such operators, is therefore no longer an issue. However, injectivity of the measurement operator is not guaranteed. We have seen in QPAT that only two out of three coefficients could be reconstructed in QPAT. QPAT data acquired at one frequency are thus not sufficient to reconstruct three coefficients. With appropriate prior information about the dependency of coefficients with respect to a frequency parameter, then injectivity of the measurement operator can be restored [13]. But again, this requires prior information that one

may not wish to make. The alternative is then to come up with *other* prior models that restore injectivity or to combine QPAT measurements with additional measurements. Hybrid inverse problems face the same shortcomings as any other inverse problem.

Once injectivity is guaranteed, then stability of the reconstructions are guaranteed in principle by the fact that singularities no longer need to propagate. We have seen a few such stability results. Note that these results typically require a certain degree of smoothness of the unknown coefficients. This is to a large extent undesirable. The reason why we have recourse to hybrid inverse problems is to obtain high resolution. The reason we typically need high resolution is because coefficients may vary rapidly and we wish to quantify such variations. It would therefore be useful to understand how stability estimates degrade when the coefficients are not smooth. That said, numerical experiments conducted in e.g. [12, 13, 14] show that reconstructions based on algorithms similar to those presented above do allow us to obtain very accurate reconstructions even for highly discontinuous coefficients, and this even in the presence of relatively significant noise.

Bibliography

[1] M. J. Ablowitz and A. S. Fokas, *Complex Variables*, Cambridge University press, 2000.

[2] R. A. Adams, *Sobolev Spaces*, Academic Press, NY, San Francisco, London, 1975.

[3] G. Alessandrini, *An identification problem for an elliptic equation in two variables*, Ann. Mat. Pura Appl., 145 (1986), pp. 265–296.

[4] ——, *Singular solutions of elliptic equations and the determination of conductivity by boundary measurements*, J. Differ. Equ., 84 (1990), pp. 252–273.

[5] G. Alessandrini and V. Nesi, *Univalent e^{σ}-harmonic mappings*, Arch. Rat. Mech. Anal., 158 (2001), pp. 155–171.

[6] E. V. Arbuzov, A. L. Bukhgeim, and S. G. Kazantsev, *Two-dimensional tomography problems and the theory of $A-$ analytic functions*, Sib. Adv. Math., 8 (1998), pp. 1–20.

[7] G. Bal, *Optical tomography for small volume absorbing inclusions*, Inverse Problems, 19 (2003), pp. 371–386.

[8] ——, *Inverse transport theory and applications*, Inverse Problems, 25 (2009), p. 053001.

[9] G. Bal, E. Bonnetier, F. Monard, and F. Triki, *Inverse diffusion from knowledge of power densities*, Inverse Problems and Imaging, 7(2) (2013), pp. 353–375.

[10] G. Bal and A. Jollivet, *Stability estimates in stationary inverse transport*, Inverse Probl. Imaging, 2(4) (2008), pp. 427–454.

[11] ——, *Stability for time-dependent inverse transport*, SIAM J. Math. Anal., 42(2) (2010), pp. 679–700.

[12] G. Bal and K. Ren, *Multi-source quantitative PAT in diffusive regime*, Inverse Problems, 27(7) (2011), p. 075003.

[13] ——, *On multi-spectral quantitative photoacoustic tomography*, Inverse Problems, 28 (2012), p. 025010.

[14] G. Bal, K. Ren, G. Uhlmann, and T. Zhou, *Quantitative thermo-acoustics and related problems*, Inverse Problems, 27(5) (2011), p. 055007.

[15] G. Bal and J. C. Schotland, *Inverse Scattering and Acousto-Optics Imaging*, Phys. Rev. Letters, 104 (2010), p. 043902.

[16] G. Bal and G. Uhlmann, *Inverse diffusion theory for photoacoustics*, Inverse Problems, 26(8) (2010), p. 085010.

[17] M. Briane, G. W. Milton, and V. Nesi, *Change of sign of the corrector's determinant for homogenization in three-dimensional conductivity*, Arch. Ration. Mech. Anal., 173(1) (2004), pp. 133–150.

[18] M. Choulli and P. Stefanov, *An inverse boundary value problem for the stationary transport equation*, Osaka J. Math., 36 (1999), pp. 87–104.

[19] D. L. Colton and R. Kress, *Inverse acoustic and electromagnetic scattering theory*, Springer Verlag, Berlin, 1998.

[20] R. J. DiPerna and P.-L. Lions, *On the cauchy problem for boltzmann equations: global existence and weak stability*, Ann. of Math. (2), 130 (1989), pp. 321–366.

[21] L. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics Vol.19, AMS, 1998.

[22] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin, 1977.

[23] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations*, Springer, Berlin, 1986.

[24] T. Goldstein and S. Osher, *The split Bregman method for L1 regularized problems*, SIAM Journal on Imaging Sciences, 2 (2009), pp. 323–343.

[25] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, Springer-Verlag, New York, first ed., 1996.

[26] R. Kohn and M. Vogelius, *Determining conductivity by boundary measurements*, Comm. Pure Appl. Math., 37(3) (1984), pp. 289–298.

[27] F. Natterer, *The mathematics of computerized tomography*, Wiley, New York, 1986.

[28] F. Natterer and F. Wübbeling, *Mathematical Methods in Image Reconstruction*, SIAM monographs on Mathematical Modeling and Computation, Philadelphia, 2001.

[29] R. G. Novikov, *An inversion formula for the attenuated X-ray transformation*, Ark. Math., 40 (2002), pp. 145–167 (Rapport de Recherche 00/05–3, Université de Nantes, Laboratoire de Mathématiques).

[30] J. Radon, *Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten*, Berichte Sächsische Akademie der Wissenschaften, Leipzig, Math.-Phys. Kl., 69 (1917), pp. 262–267.

[31] O. Scherzer, *Handbook of Mathematical Methods in Imaging*, Springer Verlag, New York, 2011.

[32] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen, *Variational Methods in Imaging*, Springer Verlag, New York, 2009.

[33] J. Sylvester and G. Uhlmann, *A global uniqueness theorem for an inverse boundary value problem*, Ann. of Math., 125(1) (1987), pp. 153–169.

[34] M. E. Taylor, *Partial Differential Equations I*, Springer Verlag, New York, 1997.

[35] O. Tretiak and C. Metz, *The exponential Radon transform*, SIAM J. Appl. Math., 39 (1980), pp. 341–354.

[36] J.-N. Wang, *Stability estimates of an inverse problem for the stationary transport equation*, Ann. Inst. Henri Poincaré, 70 (1999), pp. 473–495.