

Analysis and X-ray tomography

Joonas Ilmavirta

Version 3

These are lecture notes for the course “MATS4300 Analysis and X-ray tomography” given at the University of Jyväskylä in Fall 2017. This version is the last one to be released during the course. A final version with bigger changes will be available some time after the course for future use.

Many of the topics covered in these notes are possible for a bachelor’s or master’s thesis. Related topics are also possible at a higher level. Contact the instructor or other inverse problems researchers for details.

Contents

1	Introduction	2
2	The Fourier series	7
3	X-ray tomography on a torus	12
4	Injectivity via angular Fourier series	18
5	Abel transforms	23
6	The normal operator	29
7	The geometry of Euclidean geodesics	35
8	The sphere bundle in two dimensions	40
9	X-ray tomography and the transport equation	46
10	X-ray tomography of vector fields	50
11	The Fourier transform	55
12	Partial data	59
13	Outlook	63
A	More exercises	68

1 Introduction

1.1 Direct and inverse problems

Consider a physical system whose behavior depends on some parameters. Here are some examples:

1. X-ray images depend on how the object attenuates X-rays (described by an attenuation coefficient depending on position).
2. The way how boundary current (current flux density) depends on boundary voltage of an electrically conducting object depends on the (position-dependent) conductivity.
3. The spectrum of oscillations of a drum depends on the shape of the drum.

The direct problem asks to determine the behavior, given the parameters:

1. Given the attenuation coefficient, find the attenuation of any X-ray.
2. Given the conductivity, find how the boundary current depends on boundary voltage.
3. Given the drum shape, find the spectrum.

The inverse problem asks the opposite:

1. Given the attenuation data for all lines, find the attenuation coefficient everywhere.
2. Given how the boundary current depends on boundary voltage, find the conductivity everywhere inside.
3. Given the spectrum, find the shape.

These inverse problems are theoretical problems in physics. We are interested in the mathematical formulations of these problems, particularly the first one. Solving the mathematical problem is a necessary step in solving the physical problem, but there are many more steps to take. We will ignore numerical implementation, data acquisition, and other practical considerations, and focus on the underlying mathematics.

It is not at all unusual that a physical problem becomes a mathematical problem after some analysis. This is done in a number of courses in physics, and physicists are well acquainted with solving mathematical problems arising from physics. The issue with these three inverse problems is that the underlying mathematical problems are hard:

1. Given the integral of a continuous (or other) function $\mathbb{R}^n \rightarrow \mathbb{R}$ over each line, reconstruct the function.
2. Let $\Omega \subset \mathbb{R}^n$ be a nice domain and $\gamma: \Omega \rightarrow (0, \infty)$ with $\log(\gamma) \in L^\infty$. Given $\{(u|_{\partial\Omega}, \gamma\nu \cdot \nabla u|_{\partial\Omega}); u \in H^1(\Omega), \nabla \cdot (\gamma \nabla u) = 0\}$, find γ .
3. Given the Dirichlet spectrum of the Laplace operator on a domain $\Omega \subset \mathbb{R}^n$, find the domain.

Here and henceforth “nice” is not a precise term, but used when we want to avoid a precise technical definition for the sake of clarity.

The first one is the simplest, also because it is linear. The second one is harder and there are some big open problems related to it, but it is still relatively well understood. Our understanding of the third problem is very limited.

To give specific examples, the first problem has been solved for compactly supported distributions, the second one for $n = 2$ (and $n \geq 3$ if γ is Lipschitz), and the third one very partially (there are some counterexamples and rigidity results and very few full uniqueness results).

Let us then see how the mathematical and physical versions of the first problem are related. In the following, feel free to assume any regularity assumptions on f .

Exercise 1. Consider a ray of light travelling on the real axis in the positive direction. Let the intensity at $x \in \mathbb{R}$ be $I(x)$. If the attenuation function is $f: \mathbb{R} \rightarrow \mathbb{R}$ (a sufficiently regular positive function), then I satisfies the Beer–Lambert law

$$I'(x) = -f(x)I(x). \quad (1)$$

Solve this differential equation. Show that if $I(0) \neq 0$ (if the intensity was zero, there would not be any real measurement), then the knowledge of $I(0)$ and $I(L)$ determines $\int_0^L f(x) dx$. \bigcirc

In physics, the attenuation coefficient is often denoted by μ . Since it is the most important function on this course, it will be most convenient to follow the mathematical convention and call it f .

Exercise 2. Consider a bounded domain (some physical object) $\Omega \subset \mathbb{R}^3$. Suppose the attenuation is described by a continuous function $f: \mathbb{R}^3 \rightarrow [0, \infty)$ with $f = 0$ in $\mathbb{R}^3 \setminus \Omega$. Consider a line segment $\gamma: [0, L] \rightarrow \mathbb{R}^3$, $\gamma(t) = x_0 + tv$, and suppose that $\gamma(0)$ and $\gamma(L)$ are both outside Ω . Suppose that we fire an X-ray beam along γ and measure the initial and final intensity. Argue that such a measurement determines the integral of f over γ . \bigcirc

1.2 Goals

All the mathematical inverse problems above are of the following form: Consider a function $F: X \rightarrow Y$. Given $F(x)$, find x .

The direct problem is finding the function F (and proving it is well defined), and this function is called the forward operator. The function F can be complicated. Let us see what the sets X and Y are in the three examples above:

1. $\{\text{continuous compactly supported functions supported in } \bar{\Omega} \subset \mathbb{R}^n\}$
 $\rightarrow \{\text{real-valued functions on the set of lines}\}$
2. $\{\gamma: \Omega \rightarrow (0, \infty); \log(\gamma) \in L^\infty\}$
 $\rightarrow \mathcal{P}(H^{1/2}(\partial\Omega) \times H^{-1/2}(\partial\Omega))$
3. $\{\text{smooth bounded domains in } \mathbb{R}^n\}$
 $\rightarrow \{\text{multisets of positive real numbers}\}$

Once one understands the forward operator, one can start studying the corresponding inverse problem. So, what does it exactly mean to find x ? There are several kinds of goals:

- Uniqueness: Show that if $F(x) = F(x')$, then $x = x'$.
- Reconstruction: Give a formula or other method to reconstruct x , given $F(x)$. That is, find a left inverse function $G: Y \rightarrow X$ so that $G \circ F = \text{id}_X$.
- Stability: Show that if $F(x) \approx F(x')$, then $x \approx x'$. Equip the spaces X and Y with suitable norms or topologies, and prove that the left inverse G is continuous.

A left inverse is what we use to process the data. We have measured $F(x)$, and we compute $G(F(x))$ to find x . There is no need for a two-sided inverse, and there can be several ways G to analyze the data.

Ideally, we want a stable reconstruction, so that the left inverse G is continuous. This has nothing to do with continuity of the unknown function; the operators F and G can map between any kinds of function spaces.

The study of any mathematical inverse problem starts with uniqueness, and that is what we shall focus on in this course. That is, our sole goal is to prove that a certain function F is injective. Some uniqueness proofs immediately give a formula for G .

One important aspect we will ignore is range characterizations. This is about finding what kind of data can really arise from real measurements — finding the set $F(X) \subset Y$.

1.3 The X-ray transform

The forward operator in the X-ray tomography problem is known as the X-ray transform. There are several different notations out there. We will denote it by \mathcal{I} . It maps functions on \mathbb{R}^n into functions on the set of all lines.

In general, it is defined so that if f is a function in \mathbb{R}^n and γ is a line in \mathbb{R}^n , then $\mathcal{I}f(\gamma)$ is the integral of f over γ . This definition can be extended to various classes of functions, or even distributions.

Let us see one example of a definition precisely.

Exercise 3. Let $B \subset \mathbb{R}^n$ be the unit ball, and let C_B denote the space of continuous functions $f: \mathbb{R}^n \rightarrow \mathbb{R}$ with $f(x) = 0$ for $x \notin B$. Show that if the space C_B is equipped with the norm $\|f\| = \sup_B |f|$, then it is a Banach space. ○

Exercise 4. Let us parametrize all lines in \mathbb{R}^n with $x \in \mathbb{R}^n$ and $v \in S^{n-1}$. Explain why $\mathcal{I}: C_B \rightarrow C_b(\mathbb{R}^n \times S^{n-1})$ given by

$$\mathcal{I}f(x, v) = \int_{-\infty}^{\infty} f(x + tv) dt \quad (2)$$

is well defined, linear, and continuous, when $C_b(\mathbb{R}^n \times S^{n-1})$, the space of continuous and bounded functions $\mathbb{R}^n \times S^{n-1} \rightarrow \mathbb{R}$, is also equipped with the supremum norm. (It turns out that this \mathcal{I} is injective but it does not have a continuous left inverse.) ○

We will not pursue optimal regularity in this course. Our interest will be in ideas and tools, not proving theorems with sharp assumptions. A reader with suitable experience in analysis is invited to consider lower regularity versions of the results presented here.

In particular, we want to show that the X-ray transform defined in the previous exercise is injective. If we need to make additional assumptions like differentiability, we will. Since the operator is linear, we need to show that $\mathcal{I}f = 0$ implies $f = 0$. We will prove this result in a number of different ways and review the necessary tools. This is the whole plan for this course.

Exercise 5. Physically, there is a constraint on the attenuation function f . Namely, the attenuation must be non-negative: $f \geq 0$. Recall the Beer–Lambert law and explain why this is physically reasonable. ○

Exercise 6. Prove that if $f \in C_B$, $f \geq 0$, and $\mathcal{I}f = 0$ (the integral is zero over all lines), then $f = 0$. This is far easier to prove than injectivity of the X-ray transform \mathcal{I} . Does the desired uniqueness result for non-negative attenuation functions follow from this observation? ○

Non-continuous attenuation functions are physically relevant. We restrict our attention to continuous functions for technical convenience. In some exercises we will consider non-continuous functions, but they will be integrable over each geodesic.

In one dimension the problem is hopeless. Therefore we make the standing assumption that the dimension n is at least 2 unless otherwise mentioned.

1

1.4 The Radon transform and parametrizations

In the X-ray transform a function is integrated over all lines. In the Radon transform a function is integrated over all hyperplanes. In the plane these two transforms coincide, but in higher dimensions they do not.

Let us give a more detailed description of the Radon transform. Let H be the set of all hyperplanes in \mathbb{R}^n . Then the Radon transform of a, say, compactly supported continuous function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is a function $Rf: H \rightarrow \mathbb{R}$ given by $Rf(h) = \int_h f d\mathcal{H}^{n-1}$. The integral over the hyperplane h is of course taken with respect to the Hausdorff measure of dimension $n-1$. This is the same thing as identifying the hyperplane (isometrically) with \mathbb{R}^{n-1} and using the usual Lebesgue measure.

Exercise 7. Explain how one can calculate the Radon transform of a function, given its X-ray transform. Then explain how injectivity of the Radon transform implies injectivity of the X-ray transform. \bigcirc

Whichever transform we study, we need to describe the lines or hyperplanes somehow. There are various options:

- Consider the abstract set of all lines in \mathbb{R}^n .
- Parametrize a line with a point $x \in \mathbb{R}^n$ and a direction $v \in S^{n-1}$. The line is $x + v\mathbb{R}$.
- Parametrize a line in \mathbb{R}^2 with the closest point to the origin. (This only fails to parametrize the lines through origin.)

This is not all. One can also use fan beam coordinates, parallel beam coordinates, or identify a line with a direction and the boundary point of entrance.

When the parametrization of lines is redundant, the X-ray transform should have the same value with different parameters representing the same

¹Another exercise belongs here. See the appendix.

line. This is a simple example of a (partial) range characterization. We will not try to characterize $\mathcal{I}(C_B) \subset C_b(\mathbb{R}^n \times S^{n-1})$, for example.

Hyperplanes in \mathbb{R}^n can also be parametrized by the closest point to the origin (with difficulties at the origin), just like one can do with lines in the plane. This is common in the analysis of the Radon transform.

Exercise 8. Consider the characteristic function of a ball centered at the origin with unit radius. Find the X-ray transform using each of the ways listed above to describe the lines. ○

Exercise 9. Let us use the second parametrization given above, characterizing lines as $x + v\mathbb{R}$. The X-ray transform of a certain function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is

$$\mathcal{I}f(x, v) = \begin{cases} \sqrt{2 + |v \cdot x|^2 - |x|^2} & \text{when } 2 + |v \cdot x|^2 - |x|^2 \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

What is the function f ? ○

Exercise 10. The most typical X-ray imaging method is computerized tomography (CT), where a three-dimensional image (of the attenuation function) is reconstructed slice by slice.

If we can show that the X-ray transform is injective in two dimensions, then it follows that it will also be injective in higher dimensions. Explain why this is so. ○

Exercise 11. Do you have any questions or comments regarding section 1? ○

2 The Fourier series

2.1 Introduction

Consider the function series

$$f(x) = b_0 + \sum_{k=1}^{\infty} (b_k \cos(kx) + c_k \sin(kx)). \quad (4)$$

Whether or not the series converges and in which sense depends on the sequences of coefficients $(b_k)_{k=0}^{\infty}$ and $(c_k)_{k=1}^{\infty}$. It is quite obvious that if the series defines a reasonable function, then it will be periodic with period 2π .

The surprise is that every 2π -periodic function can be written as a series like this, and that the coefficient sequences are unique. The regularity of the function and the mode of convergence depends on how fast (if at all) $b_k, c_k \rightarrow 0$ as $k \rightarrow \infty$.

Having two coefficient series like above is quite awkward for a number of reasons. It is far more convenient to study the series

$$f(x) = \sum_{k \in \mathbb{Z}} a_k e^{ikx} \quad (5)$$

with complex coefficients a_k . Even if the function f is real-valued, complex coefficients are needed, so the whole theory is best built over \mathbb{C} . The definition of the X-ray transform can be easily extended from real functions to complex ones. To do so, replace \mathbb{R} by \mathbb{C} in the definition.

Exercise 12. Compare the series in (4) and (5). Write either $e^{it} = \cos(t) + i \sin(t)$ or $\cos(t) = \frac{1}{2}(e^{it} + e^{-it})$ and $\sin(t) = \frac{1}{2i}(e^{it} - e^{-it})$, and compare the two representations term by term. (No need to justify yet why it is enough to compare the terms.) Express each coefficient a_k in terms of b_k and c_k , and vice versa. ○

Exercise 13. Define the equivalence relation \sim on \mathbb{R} by declaring $x \sim y$ whenever $\frac{1}{2\pi}(x - y) \in \mathbb{Z}$. Explain briefly why this is an equivalence relation. We define the quotient $\mathbb{R}/2\pi\mathbb{Z}$ as the set of equivalence classes. Explain how functions $\mathbb{R}/2\pi\mathbb{Z} \rightarrow \mathbb{C}$ correspond uniquely to 2π -periodic functions $\mathbb{R} \rightarrow \mathbb{C}$. ○

In fact, more is true than implied by the previous exercise. The quotient $\mathbb{R}/2\pi\mathbb{Z}$ inherits a lot of structure from \mathbb{R} : topology, the structure of a smooth manifold, measure, various function spaces...

We will take much of Fourier analysis as a given fact. More details can be found in a book or course focusing on Fourier analysis. We will review some key results needed to successfully and understandingly apply Fourier tools to X-ray tomography.

2.2 Fourier transform and inverse Fourier transform on a circle

Consider the space $L^2(\mathbb{R}/2\pi\mathbb{Z})$ of measurable 2π -periodic functions $f: \mathbb{R} \rightarrow \mathbb{C}$ that satisfy

$$\int_0^{2\pi} |f(x)|^2 dx < \infty. \quad (6)$$

It is a complex Hilbert space with the inner product

$$\langle f, g \rangle = \int_0^{2\pi} \overline{f(x)} g(x) dx. \quad (7)$$

We defined $L^2(\mathbb{R}/2\pi\mathbb{Z})$ to be a space of functions $\mathbb{R} \rightarrow \mathbb{C}$, not $\mathbb{R}/2\pi\mathbb{Z} \rightarrow \mathbb{C}$. However, due to periodicity we can regard the functions in this space as functions on the quotient $\mathbb{R}/2\pi\mathbb{Z}$.

Exercise 14. Recall the space $L^2(0, 2\pi)$ of square integrable Lebesgue measurable functions $(0, 2\pi) \rightarrow \mathbb{C}$. This is a Hilbert space, and it is naturally isomorphic to $L^2(\mathbb{R}/2\pi\mathbb{Z})$. Give the natural isomorphisms in both directions. Are they isometric? (The fact that they are isomorphic follows from the fact that they are both separable infinite-dimensional complex Hilbert spaces, but there is something far simpler and more natural here.) \bigcirc

Let us denote by $\ell^2(\mathbb{Z})$ the space of “sequences” (functions) $a: \mathbb{Z} \rightarrow \mathbb{C}$ with $\sum_{k \in \mathbb{Z}} |a_k|^2 < \infty$. This, too, is a Hilbert space. For convenience, we equip it with the norm

$$\|a\|^2 = 2\pi \sum_{k \in \mathbb{Z}} |a_k|^2 \quad (8)$$

and the corresponding inner product. (The inner product is left implicit, and the reader is encouraged to figure out what the inner product should be. One can of course use the polar formula to find the inner product from the norm, but in a simple case like this one one can see the correct inner product by eye.)

Definition 2.1. The Fourier transform of a function or distribution expressed as the Fourier series (5) takes the function f into the sequence $(a_k)_{k \in \mathbb{Z}}$ of Fourier coefficients. The inverse Fourier transform takes the sequence back to the function or distribution. In symbols, $\mathcal{F}f = a$ and $\mathcal{F}^{-1}a = f$.

The definition above is purposely vague. It describes the overall idea of the Fourier transform and its inverse in the present context. The same definition can be used for a large number of different function spaces. Observe that the Fourier transform and its inverse are linear operators. The Fourier transform of functions on the whole line is a different animal, and we shall greet it later.

A central result in Fourier analysis is that the Fourier transform is well-defined and the inverse exists. Even more is true:

Theorem 2.2. *The Fourier transform on $\mathbb{R}/2\pi\mathbb{Z}$ is a unitary isometry $\mathcal{F}: L^2(\mathbb{R}/2\pi\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z})$, given by*

$$(\mathcal{F}f)(k) = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx, \quad (9)$$

which is well defined as a Lebesgue integral. The inverse Fourier transform $\mathcal{F}^{-1}: \ell^2(\mathbb{Z}) \rightarrow L^2(\mathbb{R}/2\pi\mathbb{Z})$ is also unitary and isometric, and is given by

$$(\mathcal{F}^{-1}a)(x) = \sum_{k \in \mathbb{Z}} a_k e^{ikx}, \quad (10)$$

where the series of functions converges in $L^2(\mathbb{R}/2\pi\mathbb{Z})$.

This theorem will not be proven on this course. The theorem can be rephrased as the functions $x \mapsto \frac{1}{\sqrt{2\pi}} e^{ikx}$, $k \in \mathbb{Z}$, being an orthonormal Hilbert basis for $L^2(\mathbb{R}/2\pi\mathbb{Z})$. In general, a Hilbert space is isometric to the ℓ^2 space over the index set of a Hilbert basis.

Every L^2 function $f: \mathbb{R}/2\pi\mathbb{Z} \rightarrow \mathbb{C}$ can be written uniquely as a series

$$f(x) = \sum_{k \in \mathbb{Z}} \mathcal{F}f(k) e^{ikx}. \quad (11)$$

This series is the Fourier series. Sometimes $\mathcal{F}f$ is denoted by \hat{f} .

The elements of a sequence are typically denoted as a_k instead of $a(k)$, but in Fourier analysis it is customary and convenient to write $\mathcal{F}f(k)$ or $\hat{f}(k)$ instead of $\mathcal{F}f_k$ or \hat{f}_k .

Exercise 15. Recall definitions from an earlier course or some other source. What does it mean in formulas (involving sums and integrals) that the Fourier transform is isometric and unitary? \bigcirc

One may wonder why the Fourier transform of a 2π -periodic function $\mathbb{R} \rightarrow \mathbb{C}$ is a function on \mathbb{Z} , not on \mathbb{R} . This has nothing to do with the specific problem, it is a mathematical property. A function cannot be 2π -periodic unless all frequencies are integers. To make this statement more rigorous, one can show that the Fourier transform (in the sense of whole \mathbb{R} , not $\mathbb{T}^1 = \mathbb{R}/2\pi\mathbb{Z}$) of a periodic function is a distribution supported on the lattice $2\pi\mathbb{Z}$. The same is true in higher dimensions as well. Another way to see this will come in section 11 when we discuss the Fourier transform in greater generality. The fact that only discrete frequencies are possible is not obvious at first. It is a key result in Fourier analysis that is seldom stated explicitly.

2.3 Multidimensional Fourier series

In the previous section we considered Fourier series in one dimension. The theory is very, very similar in higher dimensions.

In higher dimensions, one studies functions $\mathbb{R}^n \rightarrow \mathbb{C}$ which are 2π -periodic in all n real variables. Notice that the space of such functions is not rotation

invariant; the coordinate axes give n preferred directions. These preferred directions are perhaps more apparent in the lattice

$$2\pi\mathbb{Z}^n = \{x \in \mathbb{R}^n; x_i/2\pi \in \mathbb{Z} \text{ for all } i\}. \quad (12)$$

As we did above for $n = 1$, we may quotient the space \mathbb{R}^n with the lattice $2\pi\mathbb{Z}^n$ to form $\mathbb{R}^n/2\pi\mathbb{Z}^n$.

Exercise 16. What is the equivalence relation in \mathbb{R}^n corresponding to the lattice? What are the equivalence classes? \bigcirc

The quotient space (not a quotient *vector* space) $\mathbb{R}/2\pi\mathbb{Z}$ is homeomorphic (in fact isometric) to the circle S^1 . The quotient space $\mathbb{R}^n/2\pi\mathbb{Z}^n$ is the same as $(\mathbb{R}/2\pi\mathbb{Z})^n$ or $(S^1)^n$, but not S^n . We have n coordinates, each considered modulo 2π . The topological space $\mathbb{R}^n/2\pi\mathbb{Z}^n$ is the n -dimensional torus. The most famous torus is the two-dimensional one, and the one-dimensional torus is often called simply the circle.

For the differential geometrically oriented: The lattice acts isometrically on the space \mathbb{R}^n , so the quotient inherits the (Euclidean) Riemannian metric. The torus with this metric is locally isometric to \mathbb{R}^n , and is called the flat torus.

Exercise 17. The Euclidean space \mathbb{R}^n is an additive group and $2\pi\mathbb{Z}^n$ is a subgroup. Why is the quotient group $\mathbb{R}^n/2\pi\mathbb{Z}^n$ well defined? (That is, why is the subgroup normal?) How does this quotient group correspond to the quotient space $\mathbb{R}^n/2\pi\mathbb{Z}^n$ described above? Describe the group operation. \bigcirc

A function $\mathbb{R}^n/2\pi\mathbb{Z}^n \rightarrow \mathbb{C}$ — or, equivalently, a function on \mathbb{R}^n with period 2π in each variable — is written as a Fourier series as follows:

$$f(x) = \sum_{k \in \mathbb{Z}^n} a_k e^{ik \cdot x}. \quad (13)$$

The spaces $L^2(\mathbb{R}^n/2\pi\mathbb{Z}^n)$ and $\ell^2(\mathbb{Z}^n)$ are defined analogously to the one-dimensional case. The norm on the latter space is

$$\|a\|^2 = (2\pi)^n \sum_{k \in \mathbb{Z}^n} |a_k|^2. \quad (14)$$

Using these spaces, we have the following generalization of theorem 2.2:

Theorem 2.3. *The Fourier transform on the torus \mathbb{T}^n is a unitary isometry $\mathcal{F}: L^2(\mathbb{R}^n/2\pi\mathbb{Z}^n) \rightarrow \ell^2(\mathbb{Z}^n)$, given by*

$$(\mathcal{F}f)(k) = \frac{1}{(2\pi)^n} \int_{[0,2\pi]^n} f(x) e^{-ik \cdot x} dx, \quad (15)$$

which is well defined as a Lebesgue integral. The inverse Fourier transform $\mathcal{F}^{-1}: \ell^2(\mathbb{Z}^n) \rightarrow L^2(\mathbb{R}^n/2\pi\mathbb{Z}^n)$ is also unitary and isometric, and is given by

$$(\mathcal{F}^{-1}a)(x) = \sum_{k \in \mathbb{Z}^n} a_k e^{ikx}, \quad (16)$$

where the series of functions converges in $L^2(\mathbb{R}^n/2\pi\mathbb{Z}^n)$.

Exercise 18. Let us denote $e_k(x) = e^{ik \cdot x}$. For any $k \in \mathbb{Z}^n$ we have $e_k \in L^2(\mathbb{R}^n/2\pi\mathbb{Z}^n)$. Using the given inner product, prove that

$$\langle e_k, e_m \rangle = c \delta_{km}, \quad (17)$$

where δ_{km} is the Kronecker delta and c is a constant. What is the constant? Do not appeal to theorem 2.3, but calculate by hand. ○

2

The general idea is to show that $\mathcal{I}f = 0 \implies \mathcal{F}f = 0 \implies f = 0$. That is, the X-ray transform of f is easier to connect to the Fourier transform $\mathcal{F}f$ than f itself.

The Fourier series can be defined on other spaces, for example non-flat tori, using the eigenfunctions of the Laplace operator. It will be considerably more clumsy and it will not work so nicely together with the X-ray transform. Fourier analysis tends to be most convenient when one has enough symmetry.

3

Exercise 19. Do you have any questions or comments regarding section 2? ○

3 X-ray tomography on a torus

In this section we will give our first injectivity proof based on Fourier series on \mathbb{T}^n .

3.1 Geodesics on a torus

The analogue of a straight line in differential geometry is a geodesic. Similarly to the problem we set out to study, one can ask whether a function on a manifold is determined by its integrals over geodesics. This is an active field of study, but beyond the scope of this course.

²Another exercise belongs here. See the appendix.

³See section A.14 for more exercises.

However, we will study this problem now on the flat torus $\mathbb{T}^n = \mathbb{R}^n / 2\pi\mathbb{Z}^n$. The reason is that this provides one of the simplest proofs of the injectivity of the X-ray transform in a bounded Euclidean domain.

Geodesics, like any curves, can be regarded as subsets of the space or as functions from an interval to the space. As a set, a geodesic in \mathbb{R}^n is simply a straight line. As a function, a geodesic can be described as $\gamma: [0, 1] \rightarrow \mathbb{R}^n$, $\gamma(x) = x + tv$. The velocity $v \in \mathbb{R}^n$ can be any non-zero vector; it will be convenient not to assume unit speed in this section. A geodesic is a curve with constant velocity.

Here we chose to parametrize the geodesic by $[0, 1]$, and we have therefore described a geodesic between two points (x and $x + v$). Another option is to replace the interval with \mathbb{R} ; this leads to what is called a maximal geodesic. In our Euclidean X-ray tomography problem we consider the integrals of an unknown function over all geodesics through a given domain. It is irrelevant whether the geodesics are maximal or between two points, as long as the two points are outside (or at the boundary of) the domain.

Let us then turn to geodesics on a torus. Let $q: \mathbb{R}^n \rightarrow \mathbb{T}^n$ be the quotient map that takes a point to its equivalence class. One can write it as $q(x) = x + 2\pi\mathbb{Z}^n \subset \mathbb{R}^n$. This formula is seldom very useful in practice, but perhaps it helps get a hold of the idea.

A geodesic between two points on the torus \mathbb{T}^n is simple to describe: we may compose a geodesic on \mathbb{R}^n with the quotient map. We take $x \in \mathbb{R}^n$ and $v \in \mathbb{R}^n \setminus 0$ and define $\gamma: [0, 1] \rightarrow \mathbb{T}^n$ by $\gamma(t) = q(x + tv)$. We will be interested in maximal geodesics that do not terminate in either direction, which corresponds to replacing $[0, 1]$ above by \mathbb{R} .

On a torus, there is an interesting new class of maximal geodesics: closed geodesics, also known as periodic geodesics. The simplest example of a periodic geodesic is

$$\mathbb{R} \rightarrow \mathbb{T}^n, \quad t \mapsto (2\pi t, 0, \dots, 0), \quad (18)$$

which has period 1. Geodesics $\mathbb{R} \rightarrow \mathbb{T}^n$ with period 1 can be naturally identified with geodesics $\gamma: [0, 1] \rightarrow \mathbb{T}^n$ for which $\gamma(0) = \gamma(1)$.

The most convenient way to describe geodesics for our purposes is to take two parameters $x \in \mathbb{T}^n$ and $v \in \mathbb{R}^n$, and let the corresponding geodesic $[0, 1] \rightarrow \mathbb{T}^n$ be

$$\gamma(t) = q(x' + tv), \quad (19)$$

where $x' \in \mathbb{R}^n$ is any point so that $q(x') = x$. Equivalently, we may take

$$\gamma(t) = x + q(tv), \quad (20)$$

where “+” is the addition on \mathbb{T}^n — which is naturally an abelian group. We

shall write this geodesic simply as

$$\gamma(t) = x + tv \in \mathbb{T}^n, \quad (21)$$

where the quotient is left implicit.

All geodesics on a torus are of this form. This is because the quotient map $q: \mathbb{R}^n \rightarrow \mathbb{T}^n$ is a local isometry and isometries preserve geodesics. It is crucial that the torus is flat. If one uses another metric (such as the donut embedded in \mathbb{R}^3), the geodesics will be different and there will be less symmetry.

Exercise 20. Consider the geodesic described in (19) and (20) above. Show that the endpoints coincide if and only if $v \in 2\pi\mathbb{Z}^n$. \bigcirc

Exercise 21. Explain how a geodesic with velocity $v \in 2\pi\mathbb{Z}^n$ can be regarded as a function $\mathbb{R}/\mathbb{Z} \rightarrow \mathbb{T}^n$. \bigcirc

In the X-ray transform on a torus we will only integrate over periodic geodesics. The reason for this is two-fold. First, periodic geodesics are convenient and, as it turns out, sufficient. Second, the integrals are ill-defined over a non-periodic geodesic. There is a way to renormalize the integral, but it is rather awkward.

By exercise 21 a periodic geodesic is a function $\mathbb{R}/\mathbb{Z} \rightarrow \mathbb{T}^n$, and it is easy to integrate a continuous function over the compact set \mathbb{R}/\mathbb{Z} . However, when there is no periodicity, one would have to integrate over all of \mathbb{R} , and the resulting integral typically does not exist (as a finite number).

3.2 Injectivity from a torus to a Euclidean space

For any $v \in 2\pi\mathbb{Z}^n \setminus 0$, $x \in \mathbb{T}^n$ and $f \in C(\mathbb{T}^n)$, we write

$$\mathcal{I}_v f(x) = \int_0^1 f(x + tv) dt. \quad (22)$$

If v is fixed, this defines an operator

$$\mathcal{I}_v: C(\mathbb{T}^n) \rightarrow C(\mathbb{T}^n). \quad (23)$$

For us the key property is that \mathcal{I}_v is linear, but it does indeed map continuous functions to continuous functions. It has other properties as well: $\|\mathcal{I}_v\| = 1$ and $\mathcal{I}_v^2 = \mathcal{I}_v$. It is also a symmetric operator $L^2(\mathbb{T}^n) \rightarrow L^2(\mathbb{T}^n)$.

We call the family of operators \mathcal{I}_v with $v \in 2\pi\mathbb{Z}^n \setminus 0$ the X-ray transform on the torus \mathbb{T}^n . This point of view is convenient here. In the usual view \mathcal{I} is not a symmetric operator and that leads us to consider its normal operator later in this course.

Now consider a function $f \in C_B \subset C(\mathbb{R}^n)$. The function f is supported in the closed unit ball \bar{B} , so we can extend it periodically to a function \tilde{f} on \mathbb{R}^n so that $\tilde{f} = f$ on $(-\pi, \pi)^n$. Observe that $\bar{B} \subset (-\pi, \pi)^n$.

Exercise 22. Give a formula for \tilde{f} in terms of f . ○

Since it is periodic, the function \tilde{f} can be regarded as a function on the torus \mathbb{T}^n .

Lemma 3.1. *The X-ray transform of $\tilde{f} \in C(\mathbb{T}^n)$ is uniquely determined by the X-ray transform of $f \in C_B$.*

Proof. The idea is simple, but writing it down is awkward. We will do it anyway.

Take any $v \in 2\pi\mathbb{Z}^n \setminus 0$. We need to show that $\mathcal{I}_v \tilde{f}$ can be expressed in terms of $\mathcal{I}f$. Recall that $\mathcal{I}f$ is a function defined on the set of all lines in \mathbb{R}^n .

The restriction $q|_{[-\pi, \pi]^n} : [-\pi, \pi]^n \rightarrow \mathbb{T}^n$ is a bijection. Let us denote its inverse by ι . It satisfies $q \circ \iota = \text{id}_{\mathbb{T}^n}$.

Since f is supported in \bar{B} , we may consider it to be a function defined on $[-\pi, \pi]^n$. Then we have $f = \tilde{f} \circ q$ and $\tilde{f} = f \circ \iota$.

Consider any $x \in \mathbb{T}^n$ and $v \in 2\pi\mathbb{Z}^n \setminus 0$. Let $0 \leq t_0 < t_1 < \dots < t_m < 1$ be the points where at least one coordinate of $x + t_j v$ is $-\pi$. If there are infinitely many such times, we set $t_0 = 0$ and $m = 0$. Moving x in the direction of v does not affect $\mathcal{I}_v \tilde{f}(x)$ (exercise), so we may freely shift x so that $t_0 = 0$. For clarity, let us call this new point x' ; it is obtained as $x' = x + sv$ for suitable s . In fact, s can be chosen to be the original t_0 . For convenience, we write $x_{m+1} = 1$. (Exercise: Write this paragraph less confusingly.)

We have

$$\begin{aligned}
 \mathcal{I}_v \tilde{f}(x) &= \mathcal{I}_v \tilde{f}(x') \\
 &= \int_0^1 \tilde{f}(x' + tv) dt \\
 &= \int_0^1 f(\iota(x' + tv)) dt \\
 &= \sum_{j=0}^m \int_{t_j}^{t_{j+1}} f(\iota(x' + tv)) dt.
 \end{aligned} \tag{24}$$

Now, each $\int_{t_j}^{t_{j+1}} f(\iota(x' + tv)) dt$ is an integral of f over a straight line joining two boundary points of the cube $[-\pi, \pi]^n$. This is, by definition, determined by the X-ray transform $\mathcal{I}f$. Therefore $\mathcal{I}_v \tilde{f}$ can be written in terms of $\mathcal{I}f$ — although there is no pretty formula — and the proof is complete. □

Exercise 23. Describe the function $\iota \circ q : \mathbb{R}^n \rightarrow [-\pi, \pi]^n$ in words, formulas, pictures, or a combination thereof. ○

Exercise 24. Explain why $\mathcal{I}_v g(x) = \mathcal{I}_v g(x + sv)$ for any $s \in \mathbb{R}$, $g \in C(\mathbb{T}^n)$, and $v \in 2\pi\mathbb{Z}^n$. ○

Bonus exercise 1. Find a clearer way to rewrite the marked paragraph in the proof above. ○

The conclusion of the lemma is important:

Exercise 25. Suppose we know that $\mathcal{I}_v g = 0$ for all $v \in 2\pi\mathbb{Z}^n \setminus 0$ implies that the function $g \in C(\mathbb{T}^n)$ has to vanish identically. Show that if $\mathcal{I}f = 0$ for some $f \in C_B$, then $f = 0$. ○

In other words, injectivity of the X-ray transform in the Euclidean space follows from an injectivity result on the torus. This is our first solution of the inverse problem of X-ray tomography. The missing step is proving the desired result on the torus.

3.3 Interplay between the X-ray and Fourier transforms on a torus

For any fixed $v \in 2\pi\mathbb{Z}^n \setminus 0$ and $f \in C(\mathbb{T}^n)$, the X-ray transform $\mathcal{I}_v f$ is a continuous function on the torus \mathbb{T}^n . Therefore it makes sense to calculate its Fourier transform.

Lemma 3.2. *Let $v \in 2\pi\mathbb{Z}^n \setminus 0$ and $f \in C(\mathbb{T}^n)$. Then for every $k \in \mathbb{Z}^n$*

$$\mathcal{F}(\mathcal{I}_v f)(k) = \begin{cases} \mathcal{F}f(k) & \text{when } k \cdot v = 0 \\ 0 & \text{otherwise.} \end{cases} \quad (25)$$

Proof. The proof is a mere calculation:

$$\begin{aligned} \mathcal{F}(\mathcal{I}_v f)(k) &= \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} e^{-ik \cdot x} \mathcal{I}_v f(x) \, dx \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} e^{-ik \cdot x} \int_0^1 f(x + tv) \, dt \, dx \\ &\stackrel{\text{a}}{=} \frac{1}{(2\pi)^n} \int_0^1 \int_{\mathbb{T}^n} e^{-ik \cdot x} f(x + tv) \, dx \, dt \\ &\stackrel{\text{b}}{=} \frac{1}{(2\pi)^n} \int_0^1 \int_{\mathbb{T}^n} e^{-ik \cdot (y - tv)} f(y) \, dy \, dt \\ &\stackrel{\text{c}}{=} \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} e^{-ik \cdot y} f(y) \, dy \times \int_0^1 e^{i(k \cdot v)t} \, dt \\ &\stackrel{\text{d}}{=} \mathcal{F}f(k) \times \begin{cases} 1 & \text{when } k \cdot v = 0 \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (26)$$

It only remains to justify the steps. □

Exercise 26. Explain the steps a–d in (26). ○

We will next show that the X-ray transform is injective. Bear in mind that the X-ray transform is understood as a family of operators. Here injectivity means “collective injectivity”; the individual operators are not injective.

Theorem 3.3. *Let $f \in C(\mathbb{T}^n)$. If $\mathcal{I}_v f = 0$ for all $v \in 2\pi\mathbb{Z}^n \setminus 0$, then $f = 0$.*

Proof. Since the Fourier transform is bijective by theorem 2.3, it suffices to show that the Fourier series of f vanishes. To that end, take any $k \in \mathbb{Z}^n$. There is some $v \in 2\pi\mathbb{Z}^n \setminus 0$ so that $k \cdot v = 0$ (exercise). By lemma 3.2 we have $\mathcal{F}\mathcal{I}_v f(k) = \mathcal{F}f(k)$. Since $\mathcal{I}_v f = 0$ by assumption, we get $\mathcal{F}f(k) = 0$ for all $k \in \mathbb{Z}^n$. □

Exercise 27. Show that for any $k \in \mathbb{Z}^n$ there exists $w \in \mathbb{Z}^n \setminus 0$ so that $k \cdot w = 0$. ○

4

Exercise 28. We have excluded $v = 0$ from our discussion. Why is this reasonable, considering the original problem? What is the operator \mathcal{I}_0 ? ○

As a corollary, we get the following injectivity result:

Theorem 3.4. *Suppose $f \in C_B$ integrates to zero over all lines through B . Then $f = 0$.*

Proof. This follows from lemma 3.1, exercise 25, and theorem 3.3. □

Exercise 29. Summarize in your own words the proof of injectivity of the X-ray transform given in this section. ○

Injectivity in a larger ball and therefore in the whole space $C_c(\mathbb{R}^n)$ follows by a scaling argument.

It is worth noting that in this proof we did not use X-rays in all directions. Only the directions in $2\pi\mathbb{Z}^n \setminus 0$ were used. If one projects this set radially to the unit sphere S^{n-1} , one gets a countable dense set.

Exercise 30. Do you have any questions or comments regarding section 3? ○

⁴More exercises belong here. See the appendix.

4 Injectivity via angular Fourier series

In this section and the next section we will give our second injectivity proof based on Fourier series with respect to the angular variable in polar coordinates.

4.1 Angular Fourier series

In this section we will give a new way to prove injectivity of the X-ray transform. This is the one found by Allan Cormack, who together with the electrical engineer Godfrey Hounsfield was awarded the Nobel Prize in Physiology or Medicine for the development of computer assisted tomography in 1979. However, Cormack was not the first one to solve the mathematical inverse problem; it had been done in 1917 by Johann Radon, but without an idea to apply it to tomography. Radon's inversion method will not be considered in this course.

We study the problem in two dimensions. It is most convenient to consider the problem in the punctured closed disc

$$\bar{D}^* = \{x \in \mathbb{R}^2; 0 < |x| \leq 1\}.$$

Recall exercise 10 concerning the two-dimensional case.

Our aim is to reconstruct a continuous function $f: \bar{D}^* \rightarrow \mathbb{C}$ from its integrals over all lines through \bar{D}^* . We will not use the lines that pass through the origin. That is, we throw away some data. Avoiding the origin simply makes the use of polar coordinates more convenient and does not make the result any weaker.

Our original problem was to reconstruct a function in the whole disc, but it turns out to be convenient to throw away some data. This is not unusual in inverse problems. It is often best to look at a convenient subset of the data. However, the results are often stated for all of the data for clarity.

We will use polar coordinates $r \in (0, 1]$ and $\theta \in \mathbb{R}/2\pi\mathbb{Z}$ on \bar{D}^* . For any fixed r , the function $f(r, \cdot)$ is a continuous function $\mathbb{R}/2\pi\mathbb{Z} \rightarrow \mathbb{C}$. We expand it in a Fourier series. Now the coefficients of the Fourier series depend on the variable r . We have

$$f(r, \theta) = \sum_{k \in \mathbb{Z}} a_k(r) e^{ik\theta}. \quad (27)$$

The Fourier coefficients may be calculated as

$$a_k(r) = \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\theta} f(r, \theta) d\theta. \quad (28)$$

From this expression one can see that each $a_k: (0, 1] \rightarrow \mathbb{C}$ is continuous. The only difference to the usual Fourier series on the circle is the appearance of the parameter r .

We will write $f_k(r, \theta) = a_k(r)e^{ik\theta}$, so that the Fourier series becomes simply

$$f(r, \theta) = \sum_{k \in \mathbb{Z}} f_k(r, \theta). \quad (29)$$

We will not study the details of this series too deeply, but we remark that the terms are L^2 -orthogonal and the usual L^2 theory of Fourier series applies with some modifications due to the presence of r . It will suffice for us that $f = 0$ if and only if $f_k = 0$ for all $k \in \mathbb{Z}$.

Exercise 31. Suppose $f: \bar{D}^* \rightarrow \mathbb{C}$ is continuous. Show that the following are equivalent:

- (a) $f = 0$
- (b) $f_k = 0$ for all $k \in \mathbb{Z}$
- (c) $a_k = 0$ for all $k \in \mathbb{Z}$

Theorem 2.2 will be of use. In fact, the whole angular Fourier series makes sense because of this theorem. ○

In higher dimensions the functions $e^{ik\theta}$ need to be replaced with spherical harmonics. This is one of the reasons why it is convenient to restrict to dimension two.

One can study the angular Fourier series in the whole plane if one wants. As long as the function is continuous or L^2 (or whatever space one might be working with), one can apply the one-dimensional Fourier series circle by circle.

4.2 The X-ray transform in polar coordinates

For any point $x \in \bar{D}^*$, let L_x be the line segment connecting boundary points of the unit disc so that x is the closest point to the origin on L_x . If $|x| = 1$, the line will degenerate into a point. This is a convenient way to parametrize all lines through the closed unit disc that do not meet the origin.

For a continuous function $f: \bar{D}^* \rightarrow \mathbb{C}$, we define $\mathcal{I}f(x)$ to be the integral of f over L_x . Again, we use polar coordinates, so that the X-ray transform of f is a function $\mathcal{I}f(r, \theta)$. It will be useful to write this as a Fourier series in the variable θ .

For $\theta \in \mathbb{R}/2\pi\mathbb{Z}$, denote $v_\theta = (\cos(\theta), \sin(\theta))$. For $r > 0$ and $\theta \in \mathbb{R}/2\pi\mathbb{Z}$, the corresponding line can be written as

$$L_{r,\theta} = \{x \in \mathbb{R}^2; x \cdot v_\theta = r\}. \quad (30)$$

As mentioned above, this covers all the lines that do not meet the origin. If we use “extended polar coordinates” where $r \geq 0$, then we can indeed parametrize all lines. In some sense, this corresponds to replacing the origin with “directed origins”, which is a compactification of the punctured disc. In fact, one can even let the radius r to be any real number; this would lead to a global two-fold parametrization of all the lines.

4.3 Rotations and diagonalizability

Fix any $\varphi \in \mathbb{R}$. Let us define the rotation operator \mathcal{R}_φ on functions defined on \bar{D}^* so that $(\mathcal{R}_\varphi f)(r, \theta) = f(r, \theta + \varphi)$. It is clear that \mathcal{R}_φ maps continuous functions to continuous functions.

For a continuous $f: \bar{D}^* \rightarrow \mathbb{C}$, both f and $\mathcal{I}f$ are functions on \bar{D}^* . This allows us to make sense of the function $\mathcal{R}_\varphi \mathcal{I}f$. The interplay between rotations and the X-ray transform is important.

Exercise 32. Take any $\varphi \in \mathbb{R}$ and a continuous $f: \bar{D}^* \rightarrow \mathbb{C}$. Explain why $\mathcal{I}\mathcal{R}_\varphi f = \mathcal{R}_\varphi \mathcal{I}f$. \circ

The fact that rotations commute with the X-ray transform will bring additional structure.

Lemma 4.1. *Let $f(x; \varphi)$ be a continuous function defined on $\bar{D}^* \times [0, 2\pi]$. Let γ be any line through \bar{D}^* that does not meet the origin. Then*

$$\int_0^{2\pi} (\mathcal{I}f(\cdot; \varphi))(\gamma) d\varphi = \mathcal{I}F(\gamma), \quad (31)$$

where $F(x) = \int_0^{2\pi} f(x; \varphi) d\varphi$.

Exercise 33. Prove the lemma. \circ

Lemma 4.2. *Let $f: \bar{D}^* \rightarrow \mathbb{C}$ be a continuous function. Then*

$$\frac{1}{2\pi} \int_0^{2\pi} e^{-ik\theta} \mathcal{I}f(r, \theta) d\theta = \mathcal{I}f_k(r, 0) \quad (32)$$

for all $k \in \mathbb{Z}$.

The angle 0 might seem weird at first. It is best to regard the right-hand side as the X-ray transform of the one-dimensional function a_k . Introducing a non-zero angle is possible in the formula above, but it gives no additional information.

Proof of lemma 4.2. In the integrals below limits are occasionally shifted from $(0, 2\pi)$ due to changes of variables. Since the relevant functions are 2π -periodic, we do not need to change the interval of integration.

Fix any $k \in \mathbb{Z}$. First, we observe that

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\theta} \mathcal{R}_\theta f(r, \varphi) \, d\theta &= \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\theta} f(r, \varphi + \theta) \, d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} e^{-ik(\omega - \varphi)} f(r, \omega) \, d\omega \\ &= e^{ik\varphi} a_k(r) \\ &= f_k(r, \varphi). \end{aligned} \tag{33}$$

Using the definitions, exercise 32, lemma 4.1, and equation (33), we get

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\theta} \mathcal{I}f(r, \theta) \, d\theta &= \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\theta} \mathcal{R}_\theta \mathcal{I}f(r, 0) \, d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\theta} \mathcal{I} \mathcal{R}_\theta f(r, 0) \, d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} \mathcal{I}(e^{-ik\theta} \mathcal{R}_\theta f)(r, 0) \, d\theta \\ &= \mathcal{I} \left(\frac{1}{2\pi} \int_0^{2\pi} (e^{-ik\theta} \mathcal{R}_\theta f) \, d\theta \right) (r, 0) \\ &= \mathcal{I}f_k(r, 0). \end{aligned} \tag{34}$$

This concludes the proof. □

Exercise 34. The function $f(r, \theta)$ was written as a Fourier series $f = \sum_{k \in \mathbb{Z}} f_k$ in (29). Similarly, $g(r, \theta) = \mathcal{I}f(r, \theta)$ can be written as a Fourier series $g = \sum_{k \in \mathbb{Z}} g_k$. Give a formula for the function g_k in terms of $\mathcal{I}f$.

Explain why g_k only depends on f_k but not on any other f_m for $m \neq k$. ○

Our goal, as always, is to show that if $\mathcal{I}f = 0$, then $f = 0$. By exercise 34 it follows from the assumption that $\mathcal{I}f_k = 0$ for every $k \in \mathbb{Z}$. We will then fix any k and show that $\mathcal{I}f_k = 0$ implies $f_k = 0$. This problem is essentially one-dimensional, since f_k corresponds to the continuous function

$a_k: (0, 1] \rightarrow \mathbb{C}$. The aim of the next section is to solve this family of one-dimensional problems. After that we know that $f_k = 0$ for each k , and so $f = 0$.

4.4 Remarks on symmetries and functional analysis

The Fourier series of the X-ray transform depends in a rather simple way on the Fourier series of the original function. The k th Fourier component of the X-ray transform only depends on the k th Fourier component of the function. This is not a coincidence.

The X-ray transform is an operator that takes a function on \bar{D}^* to another function on \bar{D}^* . It commutes with the rotation operator \mathcal{R}_θ for any θ , so it also commutes with the derivative ∂_θ with respect to the angular coordinate. Of course, the derivative operator does not map continuous functions to continuous functions, so it should be defined on a different space or be treated as an unbounded operator, but we ignore this technicality. Now, the derivative ∂_θ and the X-ray transform \mathcal{I} commute.

At least physicists are likely to remember that if two symmetric matrices commute, they are simultaneously diagonalizable. Similar results hold for infinite-dimensional spaces. The symmetric situation is an example of a broader phenomenon. The operator ∂_θ is antihermitean, but the integral transform does not need any adjointness properties.

In our specific case this suggests that if we write the whole function space as a direct sum of eigenspaces of ∂_θ , then the X-ray transform is block diagonal. This is indeed what happens.

For example, in $L^2(D)$, the eigenspace of ∂_θ with eigenvalue ik , $k \in \mathbb{Z}$, is

$$H_k = \{f \in L^2(D); f(r, \theta) = a(r)e^{ik\theta} \text{ for some function } a\}. \quad (35)$$

One can do similar things over other function spaces. Our result in this section shows (apart from regularity assumptions), that $\mathcal{I}(H_k) \subset H_k$. It is also somewhat easy to see that $H_k \perp H_m$ when $k \neq m$. Moreover, if the operator has suitable continuity properties in L^2 — and the X-ray transform has — then one has a very convenient theory in a Hilbert space.

In higher dimensions H_k can be defined similarly, with the exponential functions $e^{ik\theta}$ replaced by spherical harmonics.

Exercise 35. Sometimes people mistakenly claim that if two matrices $A, B \in \mathbb{R}^{n \times n}$ commute (that is to say, $AB = BA$), then they are simultaneously diagonalizable (that is, there is an invertible U so that both UAU^{-1} and UBU^{-1} are diagonal). Show that this is false. \bigcirc

Exercise 36. Let the two matrices $A, B \in \mathbb{R}^{n \times n}$ be symmetric. For simplicity, you may additionally assume that all eigenvalues have multiplicity one. (The result will be true without this assumption.) Prove that if $AB = BA$, then there is an orthogonal matrix U so that UAU^T and UBU^T are both diagonal. You may assume it known that for a single real symmetric matrix such a U exists. \bigcirc

Passing from rotation symmetry (\mathcal{R}_φ) to angular derivatives (∂_θ) was a useful trick. One may ask how one might find the derivative operator, given the rotations. A formal calculation gives

$$\partial_\theta = \left. \frac{d}{d\varphi} R_\varphi \right|_{\varphi=0}. \quad (36)$$

The derivative does not exist as a limit of the difference quotient in L^2 , but it does exist in C^∞ , for example. Passing from a full symmetry to a differential symmetry is an example of passing from a Lie group to its Lie algebra. If something commutes with the Lie group, then it commutes with the Lie algebra, and the Lie algebra of a symmetry group can often be realized as differential operators.

⁵

Another thing worth pointing out is that rotation symmetry was crucially important, but Euclidean geometry was not. Similar arguments work in other rotation symmetric situations.

Exercise 37. Do you have any questions or comments regarding section 4? \bigcirc

5 Abel transforms

5.1 The block diagonal structure of the X-ray transform in polar coordinates

As discovered in the previous section, the X-ray transform has a peculiar block diagonal structure. Now it remains to find the operators on the diagonal. We shall not use the block diagonal structure in any formal way, but it is an underlying idea the reader should be aware of.

Consider the function $f_k(r, \theta) = e^{ik\theta} a_k(r)$, where $a_k: (0, 1] \rightarrow \mathbb{C}$ is continuous. We want to find an explicit formula for the X-ray transform of f_k .

⁵Another exercise belongs here. See the appendix.

To this end, consider the line $L_{s,\varphi}$ whose closest point to the origin is (s, φ) . We may assume $0 < s < 1$. Using unit length parametrization with zero parameter at the midpoint, we can write this line as the curve

$$\gamma: [-\sqrt{1-s^2}, \sqrt{1-s^2}] \rightarrow \bar{D}^* \quad (37)$$

with

$$\gamma(t) = (\sqrt{s^2 + t^2}, \varphi + \arctan(t/s)). \quad (38)$$

Exercise 38. Justify this formula geometrically. ○

We will split the interval in half and change the variable of integration from arc length $t \in (0, \sqrt{1-s^2})$ to radius $r = \sqrt{s^2 + t^2} \in (s, 1)$.

Now we can simply calculate:

$$\begin{aligned} \mathcal{I}f_k(s, \varphi) &= \int_{-\sqrt{1-s^2}}^{\sqrt{1-s^2}} f_k(\sqrt{s^2 + t^2}, \varphi + \arctan(t/s)) dt \\ &= \sum_{\pm} \int_0^{\sqrt{1-s^2}} f_k(\sqrt{s^2 + t^2}, \varphi \pm \arctan(t/s)) dt \\ &= \sum_{\pm} \int_s^1 f_k(r, \varphi \pm \arccos(s/r)) \frac{dt}{dr} dr \\ &= \sum_{\pm} \int_s^1 a_k(r) e^{ik\varphi \pm ik \arccos(s/r)} \frac{1}{\sqrt{1 - (s/r)^2}} dr. \end{aligned} \quad (39)$$

Two steps need justification, and they are left as exercises.

Exercise 39. Explain why $\arctan(t/s) = \arccos(s/r)$. ○

Exercise 40. Why is the Jacobian $\frac{dt}{dr}$ equal to $1/\sqrt{1 - (s/r)^2}$ as indicated above? ○

Our change of variable was in fact singular. But the singularity is integrable and our calculation is still valid, but to be pedantic, one may want to consider the integral with $t \in (\varepsilon, \sqrt{1-s^2})$ first and then let $\varepsilon \rightarrow 0$.

To proceed with the calculation, we must do some trigonometric manipulations.

Exercise 41. Show that $\sum_{\pm} e^{ik\varphi \pm ik \arccos(s/r)} = 2e^{ik\varphi} \cos(k \arccos(s/r))$. ○

It turns out that for $k \in \mathbb{Z}$ and $x \in [-1, 1]$, we have $\cos(k \arccos(x)) = T_{|k|}(x)$, where T_k is the k th Chebyshev polynomial of the first kind. For convenience, we will use the notation T_k instead of $T_{|k|}$ even when $k < 0$. It follows from this cosine property of the Chebyshev polynomials that $\max_{x \in [0,1]} T_k(x) = 1$ for any $k \in \mathbb{Z}$. This family of polynomials can be

defined recursively for $k \in \mathbb{N}$ by $T_0(x) = 1$, $T_1(x) = x$, and $T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x)$. Once one establishes this recursion relation, it indeed follows that the function T_k is a polynomial.

⁶

Now we can proceed from (39) to

$$\mathcal{I}f_k(s, \varphi) = 2e^{ik\varphi} \int_s^1 h(r) \frac{T_k(s/r)}{\sqrt{1 - (s/r)^2}} dr. \quad (40)$$

Based on the last section (exercise 34 implies that Fourier transform of the k th Fourier component of f contains only the k th Fourier component), we expected to pull out the factor $e^{ik\varphi}$, but the exact structure of the rest might be a bit of a surprise.

5.2 Abel transforms

Fix any $k \in \mathbb{Z}$. For a continuous function $h: (0, 1] \rightarrow \mathbb{C}$ we define a new continuous function $\mathcal{A}_k h: (0, 1] \rightarrow \mathbb{C}$ by

$$(\mathcal{A}_k h)(s) = 2 \int_s^1 a_k(r) \frac{T_k(s/r)}{\sqrt{1 - (s/r)^2}} dr. \quad (41)$$

Exercise 42. The integral above is actually only defined for $s \in (0, 1)$. Show that $\lim_{s \rightarrow 1} \mathcal{A}_k h(s) = 0$, so that it makes sense to let $\mathcal{A}_k h(1) = 0$ regardless of the value $h(1)$. \square

We call \mathcal{A}_k the k th generalized Abel transform. The reason for this name is that for $k = 0$ we have $T_0 \equiv 1$ and \mathcal{A}_0 is (one form of) the Abel transform. These are all integral transforms that take one function on the interval and turn it into another function on the interval by means of an integral formula.

We have thus found that if

$$f(r, \theta) = \sum_{k \in \mathbb{Z}} e^{ik\theta} a_k(r), \quad (42)$$

then

$$\mathcal{I}f(r, \theta) = \sum_{k \in \mathbb{Z}} e^{ik\theta} \mathcal{A}_k a_k(r). \quad (43)$$

This means that the Abel transforms are the operators on our block diagonal.

We want to show that if $\mathcal{A}_k f_k = 0$, then $f_k = 0$. That is, we want to show that the generalized Abel transform $\mathcal{A}_k: C((0, 1]) \rightarrow C((0, 1])$ is an injection. (We will not need or prove that \mathcal{A}_k maps continuous functions to continuous functions.)

⁶Another exercise belongs here. See the appendix.

Lemma 5.1. *The generalized Abel transform $\mathcal{A}_k: C((0, 1]) \rightarrow C((0, 1])$ is an injection. Moreover, $h \in C((0, 1])$ can be calculated from $\mathcal{A}_k h$ via*

$$h(r) = -\frac{1}{\pi} \frac{d}{dr} \int_r^1 \mathcal{A}_k h(s) \frac{T_k(s/r)}{s\sqrt{(s/r)^2 - 1}} ds. \quad (44)$$

Proof. It suffices to prove (44). We start by examining the integral:

$$\begin{aligned} J(r) &:= \int_r^1 \mathcal{A}_k h(s) \frac{T_k(s/r)}{s\sqrt{(s/r)^2 - 1}} ds \\ &= \int_r^1 \left(2 \int_s^1 h(t) \frac{T_k(s/t)}{\sqrt{1 - (s/t)^2}} dt \right) \frac{T_k(s/r)}{s\sqrt{(s/r)^2 - 1}} ds \\ &= 2 \int_r^1 \int_s^1 h(t) \frac{T_k(s/t)}{\sqrt{1 - (s/t)^2}} \frac{T_k(s/r)}{s\sqrt{(s/r)^2 - 1}} dt ds \\ &= 2 \int_r^1 \int_r^t h(t) \frac{T_k(s/t)}{\sqrt{1 - (s/t)^2}} \frac{T_k(s/r)}{s\sqrt{(s/r)^2 - 1}} ds dt \\ &= 2 \int_r^1 h(t) K_k(r, t) dt, \end{aligned} \quad (45)$$

where

$$K_k(r, t) = \int_r^t \frac{T_k(s/t)}{\sqrt{1 - (s/t)^2}} \frac{T_k(s/r)}{s\sqrt{(s/r)^2 - 1}} ds. \quad (46)$$

Now, somewhat magically,

$$K_k(r, t) = \frac{\pi}{2} \quad \text{whenever } 0 < r < t \text{ and } k \in \mathbb{Z}. \quad (47)$$

Some cases will be treated in the exercises. Therefore

$$J(r) = \pi \int_r^1 h(t) dt. \quad (48)$$

The desired result now follows by the fundamental theorem of calculus. \square

Exercise 43. Explain why the limits work as they do when we applied Fubini's theorem in (45). \bigcirc

Exercise 44. Prove that for every $\lambda > 0$ we have $K_k(\lambda r, \lambda t) = K_k(r, t)$. You can use this to make simplifying assumptions in subsequent calculations if you want to. \bigcirc

Exercise 45. Make the change of variable $s^2 = \frac{1}{2}[(t^2 + r^2) + y(t^2 - r^2)]$ to (46) and simplify the resulting expression. It is wise to regard the measure as ds/s . You can leave the Chebyshev polynomials untouched. \bigcirc

Exercise 46. Calculate by hand

$$A = \int_{-1}^1 \frac{a+y}{\sqrt{1-y^2}} dy \quad (49)$$

and

$$B = \int_{-1}^1 \frac{1}{(a+y)\sqrt{1-y^2}} dy, \quad (50)$$

where $a > 1$ is a real parameter. It may be convenient to differentiate $\arctan\left(\frac{1+ay}{\sqrt{(a^2-1)(1-y^2)}}\right)$. ○

Exercise 47. Making use of $T_0(x) = 1$, $T_1(x) = x$, and the previous exercises, calculate $K_0(r, t)$ and $K_1(r, t)$. ○

Bonus exercise 2. Prove the recurrence relation $K_{k+2}(r, t) = K_k(r, t)$. This together with the previous results shows (47). ○

One can define an operator \mathcal{B}_k via

$$\mathcal{B}_k h(r) = -\frac{1}{\pi} \int_r^1 h(s) \frac{T_k(s/r)}{s\sqrt{(s/r)^2 - 1}} ds. \quad (51)$$

It seems that \mathcal{B}_k is very similar in nature to \mathcal{A}_k , so one expects it to have similar mapping properties. Since $\mathcal{B}_k \mathcal{A}_k h(r) = \int_r^1 h(r) dr$, this means that both \mathcal{A}_k and \mathcal{B}_k are “integrals of order $\frac{1}{2}$ ”. In fact, the X-ray transform does integrate by order $\frac{1}{2}$, but we will not give this statement a precise meaning in this course.

Having Chebyshev polynomials in \mathcal{A}_k is not important at all for injectivity. It does help with finding an explicit inversion formula, but similar injectivity results are true in far more generality. The important things are the limits of integration and the kind of singularity at the lower limit.

5.3 Injectivity of the X-ray transform

We have now collected the needed tools, and it remains to declare the result.

Theorem 5.2. *A continuous function $\bar{D}^* \rightarrow \mathbb{C}$ is uniquely determined by its integrals over all straight lines.*

Exercise 48. Summarize the proof of the theorem in your own words. Refer to the key steps (equations, lemmas, exercises, or other). ○

Observe that no regularity assumption was made at the origin. Singularities at the origin do not matter.

7

⁷See section A.14 for more exercises.

5.4 Helgason's support theorem

In fact, even more is true than theorem 5.2.

Proposition 5.3. *Let $R \in (0, 1)$. If a continuous function $f: \bar{D} \rightarrow \mathbb{C}$ integrates to zero over all lines with distance $> R$ to the origin, then $f(x) = 0$ when $|x| > R$.*

Proof. Since $\mathcal{I}f(r, \theta)$ only depends on $f(s, \varphi)$ for $s \geq r$, it follows that $\mathcal{A}_k a_k(s) = 0$ for all $s > R$. The inversion formula for the generalized Abel transform \mathcal{A}_k is also valid for this case: If $h: (0, 1] \rightarrow \mathbb{C}$ is continuous and $\mathcal{A}_k h(r) = 0$ for all $r \in (R, 1]$, then $h(s) = 0$ for $s \in (R, 1]$. Therefore $a_k(s) = 0$ for all $s > R$ and $k \in \mathbb{Z}$, and so $f(s, \varphi) = 0$ for all $s > R$ and $\varphi \in \mathbb{R}/2\pi\mathbb{Z}$. \square

We may consider the disc $\bar{D}(0, R)$ to be an obstacle. A sufficiently nice function is uniquely determined outside the obstacle by its integrals over all lines that avoid the obstacle. Of course, nothing can be said about the function inside the obstacle from this data.

Results of this kind are often called support theorems for the X-ray transform. From a more physical point of view, this is a matter of exterior tomography — there are actual physical obstacles in the real world that one cannot fire X-rays through.

One of the most famous support theorems is due to Sigurður Helgason. We present a variant of the two-dimensional version.

Theorem 5.4 (Helgason's support theorem in the plane). *Let $K \subset \mathbb{R}^2$ be a compact and convex set. Suppose $f \in C_c(\mathbb{R}^2)$ integrates to zero over all lines $L \subset \mathbb{R}^2$ for which $L \cap K = \emptyset$. Then $f|_{\mathbb{R}^2 \setminus K} = 0$.*

Exercise 49. Argue that a compact and convex planar set is the intersection of all closed discs containing it. Then prove theorem 5.4 using proposition 5.3. \circ

Exercise 50. Explain why Helgason's support theorem (often) fails if the compact set K is not convex. Also, what does the support theorem say if $K = \emptyset$? \circ

If K is not compact, the support theorem can fail. For example, if K is a closed halfplane, then the data only contains integrals parallel to ∂K , which is certainly insufficient.

If a set is “almost convex”, then Helgason's support theorem can still work. For example, if K is the union of a closed ball and a point, the theorem is still valid as stated. This is because, in some sense, a single point is removable — the missing lines can be approximated by existing ones.

Exercise 51. Do you have any questions or comments regarding section 5?

○

6 The normal operator

In this section we will give our third injectivity proof based on the normal operator of \mathcal{I} .

6.1 Why care about a normal operator

The normal operator of a bounded linear operator $A: E \rightarrow F$ between complex or real Hilbert spaces is $A^*A: E \rightarrow E$, where $A^*: F \rightarrow E$ is the adjoint of A . Let us work over \mathbb{C} . The adjoint is defined to be the operator that satisfies

$$\langle y, Ax \rangle_F = \langle A^*y, x \rangle_E \quad (52)$$

for all $x \in E$ and $y \in F$.

Exercise 52. Using this definition, show that the adjoint A^* is unique if it exists.

○

8

Existence of the adjoint follows from the Riesz representation theorem which characterizes the dual of a Hilbert space. Namely, for any y , the mapping $x \mapsto \langle y, Ax \rangle_F$ is in E^* , and by the representation theorem there is $z \in E$ so that $\langle y, Ax \rangle_F = \langle z, x \rangle_E$. It is easy to check that this z has to depend linearly on y . This gives rise to a linear operator A^* which maps y to z . It also turns out that A^* is bounded.

Exercise 53. Show that $\|A^*\| = \|A\|$ in the operator norm.

○

It is very convenient to work with self-adjoint operators. An operator A is called self-adjoint if $A^* = A$. For the operator A to be self-adjoint, we must have $E = F$, but this is not always the case. Therefore it is convenient to replace A with its normal operator. Self-adjointness in itself is convenient, but the normal operator tends to be nicer than the original operator.

Exercise 54. Using the definition of an adjoint given above, show that the normal operator of any bounded linear operator between Hilbert spaces is self-adjoint.

○

In our case A is the X-ray transform. Now E is a function space over \mathbb{R}^n and F is a function space over the set of all lines in \mathbb{R}^n . There is no natural

⁸Another exercise belongs here. See the appendix.

way to identify the two spaces, so we will study the normal operator of the X-ray transform. Our goal is then to show that $\mathcal{I}^*\mathcal{I}$ is injective, from which it follows that \mathcal{I} is injective; see exercise 61.

6.2 Measures on spheres and sets of lines

The sphere S^{n-1} has a canonical measure. We will not define it, but we will give some descriptions, some of which count as definitions for the reader with suitable knowledge of measure theory or differential geometry. We will give a similar treatment to the set of lines soon.

In fact, a regular Borel measure is uniquely determined by the integrals of functions in C_c , so our descriptions do secretly constitute a definition of a measure.

The sphere inherits a metric from \mathbb{R}^n . The metric allows us to define Hausdorff measures of any dimension, and the natural one has dimension $n - 1$.

The sphere S^{n-1} is also a smooth manifold of dimension $n - 1$. It inherits a Riemannian metric from \mathbb{R}^n , and the Riemannian metric induces a Riemannian volume form. This leads to the same measure as the Hausdorff approach. The Riemannian metric gives rise to a metric (distance along great circles in this case), and this is the same Hausdorff measure as the Euclidean (chordal) metric.

We will denote the metric on the sphere by S . The most important property for us is a spherical Fubini's theorem. For $f \in C_c(\mathbb{R}^n)$, we have

$$\int_{\mathbb{R}^n} f(x) \, dx = \int_0^\infty \int_{S^{n-1}} f(r\omega) r^{n-1} \, dS(\omega) \, dr. \quad (53)$$

This property could also be used as a definition of the measure S .

Let us denote the set of all straight lines in \mathbb{R}^n by Γ . The lines themselves are easy to visualize, but the set of lines is a somewhat less intuitive geometrical object. There is a natural structure of a Riemannian manifold on Γ , and that gives rise to other structures as well: topology, metric, measure, ... The manifold structure is a bit tricky and unnecessary for us, but we will need to understand the structure and measure of Γ .

We will write lines as $x + v\mathbb{R} = \{x + vt; t \in \mathbb{R}\} \subset \mathbb{R}^n$ for $x \in \mathbb{R}^n$ and $v \in S^{n-1}$. This parametrization is redundant — each line is counted several times — but in the set Γ every line is only included once.

Exercise 55. Let $x_1, x_2 \in \mathbb{R}^n$ and $v_1, v_2 \in S^{n-1}$. When is $x_1 + v_1\mathbb{R} = x_2 + v_2\mathbb{R}$? ○

We will give some more details on the structure of Γ later in connection with sphere bundles. For now we rely on intuition and acknowledge that the space $C_c(\Gamma)$ of continuous and compactly supported functions $\Gamma \rightarrow \mathbb{C}$ is not rigorously defined. We point out that although the lines themselves are not compact, there are non-trivial compact sets in the space of lines.

Let us then describe the measure μ on Γ . For any $v \in S^{n-1}$, we denote by $v^\perp := \{x \in \mathbb{R}^n; x \cdot v = 0\}$ the orthogonal complement of the space spanned by v . The space v^\perp can be identified with \mathbb{R}^{n-1} , and we denote the measure there by \mathcal{H}^{n-1} (the $(n-1)$ -dimensional Hausdorff measure). The measure μ is defined so that the integral of $g \in C_c(\Gamma)$ is

$$\int_{\Gamma} g(\gamma) d\mu(\gamma) = \int_{S^{n-1}} \int_{v^\perp} g(x + v\mathbb{R}) d\mathcal{H}^{n-1}(x) dS(v). \quad (54)$$

In this representation the same line appears twice in the integral — in both orientations. The same formula can be used for the space of oriented lines as well. The double counting could be removed by replacing S^{n-1} with its antipodal quotient (the real projective space of dimension $n-1$), but multiple counting of finite order is not an issue for our purposes. More precisely, one can define an equivalence relation \sim on S^{n-1} identifying antipodal points; the antipodal quotient is S^n/\sim .

Exercise 56. For a vector $a \in \mathbb{R}^n$, define the translation operator $\varphi_a: \Gamma \rightarrow \Gamma$ by $\varphi_a(\gamma) = a + \gamma$. Show that for $g \in C_c(\Gamma)$ we have $\int_{\Gamma} g d\mu = \int_{\Gamma} g \circ \varphi_a d\mu$ for any a . ○

Exercise 57. The previous exercise shows that the measure μ is translation invariant. It is also rotation invariant. What does this property mean? Write the statement in terms of the integral of an arbitrary function like above. Then prove the statement. ○

In the definition above we chose to take base points of lines in direction v in the hyperplane v^\perp . If the hyperplanes were chosen differently, the measure would still be translation invariant, but rotation invariance requires a good choice. If we were to use a single fixed hyperplane, it would fail to parametrize all the required lines when v is contained in it. However, this is a zero measure error. If the fixed hyperplane is given as w^\perp for some fixed w , one would need to multiply the Hausdorff measure with $|w \cdot v|$. Using v^\perp induces the inconvenience of changing the space of integration depending on v , but the geometrical picture is far clearer.

6.3 The formal adjoint of the X-ray transform

Now we are ready to find the formal adjoint of the X-ray transform. The adjoint \mathcal{I}^* is a convenient operator turning (by composition) the X-ray trans-

form into an operator from a function space to itself. To find a convenient operator, it suffices to find the formal normal operator.

Our Hilbert spaces are $L^2(\mathbb{R}^n)$ and $L^2(\Gamma)$. In the definition of the adjoint, we will not use all L^2 functions — in fact, it is not important whether the X-ray transform is continuous or well defined $L^2(\mathbb{R}^n) \rightarrow L^2(\Gamma)$. Instead, we will only use functions in $C_c(\mathbb{R}^n)$ and $C_c(\Gamma)$ to find the adjoint. The whole point is to find an operator that ends up behaving nicely, and it does not matter how fishy the method to find the operator is.

We want to use the L^2 inner product, but the adjoint as we defined it does not make sense since the X-ray transform is not continuous $L^2(\mathbb{R}^n) \rightarrow L^2(\Gamma)$. This is why we need a formal adjoint, found by using only the nicer subset C_c of L^2 over both \mathbb{R}^n and Γ .

Exercise 58. Let $p \in [1, \infty)$. Let f be the characteristic function of the ball $B(0, R) \subset \mathbb{R}^n$. Show that $\|f\|_{L^p(\mathbb{R}^n)} = aR^{n/p}$ and $\|\mathcal{I}f\|_{L^p(\Gamma)} = bR^{(n+p-1)/p}$ for some constants a and b depending on the exponent p and the dimension n .

Therefore the X-ray transform is not continuous $L^p(\mathbb{R}^n) \rightarrow L^p(\Gamma)$ for any $p \in (1, \infty)$. (It is quite easy to see that the X-ray transform is also discontinuous for $p = \infty$ but is continuous for $p = 1$.) \circ

Let $f \in C_c(\mathbb{R}^n)$ and $g \in C_c(\Gamma)$. Then

$$\begin{aligned}
 \langle f, \mathcal{I}^*g \rangle &= \langle \mathcal{I}f, g \rangle \\
 &= \int_{S^{n-1}} \int_{v^\perp} \overline{\mathcal{I}f(x + v\mathbb{R})} g(x + v\mathbb{R}) \, d\mathcal{H}^{n-1}(x) \, dS(v) \\
 &= \int_{S^{n-1}} \int_{v^\perp} \int_{\mathbb{R}} \overline{f(x + tv)} g(x + v\mathbb{R}) \, dt \, d\mathcal{H}^{n-1}(x) \, dS(v) \\
 &\stackrel{\text{a}}{=} \int_{S^{n-1}} \int_{v^\perp} \int_{\mathbb{R}} \overline{f(x + tv)} g(x + tv + v\mathbb{R}) \, dt \, d\mathcal{H}^{n-1}(x) \, dS(v) \\
 &\stackrel{\text{b}}{=} \int_{S^{n-1}} \int_{\mathbb{R}^n} \overline{f(y)} g(y + v\mathbb{R}) \, dy \, dS(v) \\
 &= \int_{\mathbb{R}^n} \overline{f(y)} \left(\int_{S^{n-1}} g(y + v\mathbb{R}) \, dS(v) \right) \, dy.
 \end{aligned} \tag{55}$$

Exercise 59. Explain the steps a and b in (55). \circ

Here we used Fubini's theorem on the direct sum (product) $v\mathbb{R} \oplus v^\perp = \mathbb{R}^n$. This calculation indicates that the formal adjoint is

$$\mathcal{I}^*g(x) = \int_{S^{n-1}} g(x + v\mathbb{R}) \, dS(v). \tag{56}$$

The formal adjoint of the X-ray transform is also known as the backprojection operator.

There is a certain kind of duality between points and lines. It might be more illuminating to describe the situation in words:

- For $f \in C_c(\mathbb{R}^n)$ and $\gamma \in \Gamma$, the X-ray transform $\mathcal{I}f(\gamma)$ is the integral of $f(x)$ over all x for which $x \in \gamma$.
- For $g \in C_c(\Gamma)$ and $x \in \mathbb{R}^n$, the backprojection $\mathcal{I}^*g(x)$ is the integral of $g(\gamma)$ over all γ for which $x \in \gamma$.

Now that we have found the adjoint, it remains to find the normal operator.

Exercise 60. Show that for $f \in C_c(\mathbb{R}^n)$ we have

$$\mathcal{I}^*\mathcal{I}f(x) = 2 \int_{\mathbb{R}^n} f(x+y) |y|^{1-n} dy. \quad (57)$$

This is the (formal) normal operator that we have been looking for. ○

6.4 Convolutions and Riesz potentials

Now, we ought to show that the normal operator $\mathcal{I}^*\mathcal{I}: C_c(\mathbb{R}^n) \rightarrow C(\mathbb{R}^n)$ defined by (57) is injective.

Exercise 61. Consider a function $F: X \rightarrow Y$ between any two sets. Show that there is a left inverse $F_L^{-1}: Y \rightarrow X$ so that $F_L^{-1} \circ F = \text{id}_X$ if and only if F is injective. (Similarly, invertibility from the right is equivalent with surjectivity, but we do not need this side. In fact, this equivalence for right inverses is equivalent with the axiom of choice, but for left inverses it is not. One-sided inverse functions are typically not unique.)

Suppose we have a left inverse A for $\mathcal{I}^*\mathcal{I}$. What is a left inverse of \mathcal{I} ? ○

The convolution of two functions $f, h: \mathbb{R}^n \rightarrow \mathbb{C}$ is the function $f*h: \mathbb{R}^n \rightarrow \mathbb{C}$ defined by

$$f * h(x) = \int_{\mathbb{R}^n} f(x-y)h(y) dy \quad (58)$$

whenever this integral makes sense.

Exercise 62. The normal operator is a convolution: $\mathcal{I}^*\mathcal{I}f = f * h$. What is h ? ○

For $\alpha \in (0, n)$, the Riesz potential I_α is an integral operator defined by

$$I_\alpha f = f * h_\alpha, \quad (59)$$

where

$$h_\alpha(x) = c_\alpha^{-1} |x|^{\alpha-n} \quad (60)$$

and c_α is a constant. The Riesz representation theorem and the Riesz potential are named after two different people. They were brothers.

Theorem 6.1. *The Riesz potential $I_\alpha: C_c(\mathbb{R}^n) \rightarrow C(\mathbb{R}^n)$ is an injection for every $\alpha \in (0, n)$.*

The proof of this property of Riesz potentials is beyond the scope of this course. However, it has a corollary which is perfectly within our scope:

Theorem 6.2. *The X-ray transform is injective on $C_c(\mathbb{R}^n)$.*

Exercise 63. Prove theorem 6.2 using the results and ideas obtained in this section. \bigcirc

6.5 Remarks

The normal operator depends on the choice of the target space and the inner product on it. If we parametrized lines redundantly with $\mathbb{R}^n \times S^{n-1}$, the adjoint of the X-ray transform would take $g \in C_c(\mathbb{R}^n \times S^{n-1})$ into the function

$$\mathcal{I}^*g(x) = \int_{S^{n-1}} \int_{\mathbb{R}} g(x + tv, v) dt dS(v). \quad (61)$$

This is very similar to what we found before, but there is an additional integral over \mathbb{R} .

If we now try to compute the normal operator, we find that

$$\mathcal{I}^*\mathcal{I}f(x) = \int_{S^{n-1}} \int_{\mathbb{R}} \int_{\mathbb{R}} f(x + tv + sv) ds dt dS(v). \quad (62)$$

This differs from our earlier normal operator by the factor $\int_{\mathbb{R}} dt$ which is famously infinite. This is why redundancy in the parametrization of geodesics is problematic. We had two-fold redundancy, so we ended up with a factor 2 in our normal operator. It is easier to divide by 2 than by ∞ .

Let us then see what the effect of changing inner products is. Let $P \in \mathbb{R}^{n \times n}$ and $Q \in \mathbb{R}^{m \times m}$ be symmetric and positive definite. Equip \mathbb{R}^n with the inner product

$$\langle x, y \rangle_P = x^T P y \quad (63)$$

and similarly \mathbb{R}^m with $\langle \cdot, \cdot \rangle_Q$.

Let $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear operator (matrix). Let $B: \mathbb{R}^m \rightarrow \mathbb{R}^n$ be the adjoint with respect to these inner products. That is, suppose

$$\langle x, Ay \rangle_Q = \langle Bx, y \rangle_P \quad (64)$$

for all $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$.

⁹Another exercise belongs here. See the appendix.

The conclusion is that changing the inner product can introduce an operator between \mathcal{I}^* and \mathcal{I} , where \mathcal{I}^* is understood as the L^2 -adjoint. This is popular in X-ray tomography. The corresponding inversion method is known as filtered backprojection.

Exercise 64. Do you have any questions or comments regarding section 6? ○

7 The geometry of Euclidean geodesics

In this section and the next two sections we will give our fourth injectivity proof based on analysis on the sphere bundle.

7.1 The sphere bundle

As silly as it sounds, we will now study the geometry of straight lines in the Euclidean space \mathbb{R}^n . These same geometrical ideas will remain valid and applicable on Riemannian manifolds. We will restrict our attention to Euclidean spaces, but some differential geometric ideas will be involved, and we will have to consider a certain non-Euclidean space.

In this section we will study straight lines as curves parametrized by arc length. This is most conveniently done on the sphere bundle

$$S\mathbb{R}^n = \mathbb{R}^n \times S^{n-1} \tag{65}$$

of the Euclidean space \mathbb{R}^n . This is a bundle over \mathbb{R}^n and comes with the natural projection $\pi: S\mathbb{R}^n \rightarrow \mathbb{R}^n$ to the first component. In the Euclidean setting the bundle is simply a product of two spaces and the bundle is trivial. In a non-Euclidean situation the bundle structure becomes more complicated.

A point $(x, v) \in S\mathbb{R}^n$ describes a point and a velocity at that point. The velocity variable v takes values in the fiber S^{n-1} of the bundle. Parametrization by unit speed instead of arbitrary speed is very convenient, as it makes the fibers compact. The usefulness becomes apparent when we integrate over a sphere bundle later on.

There are two kinds of directions on $S\mathbb{R}^n$. Directions in the \mathbb{R}^n component are called horizontal and those in S^{n-1} are called vertical. This terminology will reappear in the next section when we consider horizontal and vertical derivatives of a function on the sphere bundle. This choice of words correspond to the canonical way of drawing the base \mathbb{R}^n of the bundle horizontally and the fibers S^{n-1} vertically.

7.2 The geodesic flow

Straight lines are geodesics. There is a dynamical system associated with geodesics, and we will examine it next.

Definition 7.1. A continuous time dynamical system on a set Z is a function $\varphi: [0, \infty) \times Z \rightarrow Z$ which satisfies $\varphi_0(z) = z$ and $\varphi_s(\varphi_t(z)) = \varphi_{s+t}(z)$ for all $z \in Z$ and $s, t \in [0, \infty)$.

A dynamical system describes the time evolution of a point in the phase space Z . Every $z \in Z$ has a unique trajectory or integral curve $t \mapsto \varphi_t(z)$. It is also possible to view a dynamical system more algebraically, as the action of the additive monoid $[0, \infty)$ on the set Z .

Exercise 65. Which of the following are dynamical systems on \mathbb{R} ?

- (a) $\varphi_t(z) = z + 3t$.
- (b) $\varphi_t(z) = 4$.
- (c) $\varphi_t(z) = z^t$.
- (d) $\varphi_t(z) = z - t^2$.
- (e) $\varphi_t(z) = e^{-2t}z$.
- (f) $\varphi_t(z) = z + tz$.

Explain briefly. ○

The geodesic flow is a dynamical system on the sphere bundle. It is simply given by

$$\varphi_t(x, v) = (x + tv, v). \quad (66)$$

The geodesic flow could be equally well defined on any bundle $\mathbb{R}^n \times A$ for $A \subset \mathbb{R}^n$ with the same formula. The most natural choices are $A = \mathbb{R}^n$ (all velocities possible) and $A = S^{n-1}$ (unit speed geodesics). This is what we meant above by saying that unit speed geodesics make fibers compact.

Exercise 66. Why cannot the geodesic flow on \mathbb{R}^n be a dynamical system on \mathbb{R}^n ? How does moving to the sphere bundle help? ○

As in the case of geodesics, dynamical systems are often studied on manifolds. Then there is a vector field W on the manifold Z so that for any initial point $z \in Z$ the function $f(t) = \varphi_t(z)$ solves the differential equation $f'(t) = W(f(t))$. Such a vector field is called the generator of the flow. Whenever φ is smooth enough, the generator exists and can be computed by differentiating the flow with respect to t — that is, $W(z) = \partial_t \varphi_t(z)|_{t=0}$. One

can also impose much more structure on a flow but it will not be necessary for us here. We only remark that the geodesic flow can be seen as a contact flow or a Hamiltonian flow.

Exercise 67. On the real line \mathbb{R} a vector field can be considered to be just a function $\mathbb{R} \rightarrow \mathbb{R}$. Go back to the dynamical systems of exercise 65. What are their generators? \bigcirc

The generator of the geodesic flow is called the geodesic vector field, and it is denoted by X . It is typical in differential geometry to identify a vector field with the associated differential operator. For example, a vector field $w: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is identified with the differential operator $f \mapsto w \cdot \nabla f$ which maps scalar functions to scalar functions. For us the geodesic vector field is just a differential operator, but we still call it a vector field to follow standard terminology.

The differential operator corresponding to the generator of the flow is the derivative along the flow. Consider a function $u: S\mathbb{R}^n \rightarrow \mathbb{R}$. The geodesic vector field is defined to be

$$Xu(x, v) = \partial_t u(\varphi_t(x, v))|_{t=0}. \quad (67)$$

Exercise 68. Using the definition of the geodesic flow, find a formula for the geodesic vector field. For a function $u: \mathbb{R}^n \times S^{n-1} \rightarrow \mathbb{R}$, let us denote the gradient with respect to the first component by $\nabla_x u$. \bigcirc

If we were to write X as a vector field instead of a differential operator, it would be $X(x, v) = (v, 0)$. The second component is the zero vector field on S^{n-1} . The fact that the second component vanishes means that X is horizontal.

Exercise 69. Suppose $u(x, v) = x \cdot v$. What is $Xu(x, v)$? \bigcirc

A geodesic on \mathbb{R}^n is the projection of a trajectory of the geodesic flow. Trajectories are of the form $t \mapsto \varphi_t(x, v) = (x + tv, v)$, and geodesics are of the form $t \mapsto \pi(\varphi_t(x, v)) = x + tv$.

This process can also be reversed. If $\gamma: \mathbb{R} \rightarrow \mathbb{R}^n$ is a differentiable unit speed curve, we can define its lift $\tilde{\gamma}: \mathbb{R} \rightarrow S\mathbb{R}^n$ by $\tilde{\gamma}(t) = (\gamma(t), \dot{\gamma}(t))$. The lift of a geodesic is a trajectory of the geodesic flow.

7.3 The manifold of geodesics

Previously we discussed the set Γ of all straight lines in \mathbb{R}^n . We can describe its structure a little more now.

The geodesic flow φ_t makes sense for any $t \in \mathbb{R}$ with exactly the same formula. This gives rise to an equivalence relation on the sphere bundle,

where (x, v) is considered equivalent to (x', v') if and only if $\varphi_t(x, v) = (x', v')$ for some $t \in \mathbb{R}$. We can form the quotient space of the sphere bundle with this relation, and we denote it by $S\mathbb{R}^n/\varphi$.

This quotient space has the structure of a smooth Riemannian manifold. In general, quotient manifolds are somewhat ill-behaved, but this particular quotient does make sense. For geodesics on a general manifold this is no longer the case. However, the quotient is always sensible as a topological space, but the resulting space can be wild.

This $S\mathbb{R}^n/\varphi$ is the space of all oriented lines. We previously denoted it by Γ . To get the set of all unoriented lines, one has to take another quotient to identify opposite orientations. This second quotient is well behaved.

Bonus exercise 3. Consider the geodesic flow on the sphere bundle of the torus \mathbb{T}^2 . The flow is well defined since the manifold is geodesically complete. The sphere bundle $S\mathbb{T}^2 = \mathbb{T}^2 \times S^1$ is a topological space and the quotient with respect to any equivalence relation can be given the quotient topology. Show that the topological quotient space $S\mathbb{T}^2/\varphi$ is not Hausdorff. Does the same apply in \mathbb{T}^n in any dimension n ? \bigcirc

7.4 Bounded domains and integral functions

It will be convenient to consider geodesics on bounded sets. For a bounded set $\Omega \subset \mathbb{R}^n$ we define the unit sphere bundle $S\Omega$ as $\Omega \times S^{n-1}$. This causes a technical inconvenience for the geodesic flow on $S\Omega$, since it is not defined for all times.

One can define a local dynamical system on a space Z so that for each $z \in Z$ the flow $\varphi_t(z)$ is defined for times t in some interval $J_z \subset \mathbb{R}$. For a simple example, consider $Z = [0, 1)$ and $\varphi_t(z) = z + t$ whenever it is well defined. In this case $J_z = [-z, 1 - z)$. For $z \in (0, 1)$ the interval J_z contains a neighborhood of $0 \in \mathbb{R}$ and the dynamical system behaves locally just as well as the usual kind of a dynamical system, but one just cannot go arbitrarily far in time. Things are a little more complicated at $z = 0$, where one can only follow the flow to the positive direction. The geodesic flow on a bounded domain is essentially of this type, with only one direction available at $\partial(S\Omega)$. At a tangential boundary point the flow is stuck; it cannot move in either direction.

Now, let $\Omega \subset \mathbb{R}^n$ be a smooth and strictly convex domain and $\bar{\Omega}$ its closure.

For $(x, v) \in S\Omega$, let $\tau(x, v) = \max\{t > 0; \varphi_t(x, v) \in S\bar{\Omega}\}$. This is the time it takes for a geodesic starting at (x, v) to escape Ω .

The boundary of the sphere bundle $S\Omega$ is $\partial(S\Omega) = \partial\Omega \times S^{n-1}$. Observe

that $\partial(S\Omega) = S\bar{\Omega} \setminus S\Omega$. For $x \in \partial\Omega$, let $\nu(x)$ denote the outer unit normal vector to $\partial\Omega$. A point $(x, v) \in \partial(S\Omega)$ is called an inward boundary point if $v \cdot \nu(x) < 0$. Similarly, the outward part of the boundary consists of points in $\partial(S\Omega)$ where the inner product is positive and the tangential part of the points where it is zero. Let us denote the inward boundary by $\partial_{\text{in}}(S\Omega) \subset \partial(S\Omega)$.

Exercise 70. The definition of $\tau(x, v)$ can be naturally extended to all $(x, v) \in S\bar{\Omega}$. What is the definition at an inward boundary point? What should τ be defined to be at other boundary points? \bigcirc

Exercise 71. The domain $\Omega \subset \mathbb{R}^n$ is called smooth if there is a smooth boundary defining function $\rho: \mathbb{R}^n \rightarrow \mathbb{R}$ so that $\Omega = \{x \in \mathbb{R}^n; \rho(x) > 0\}$ and $\nabla\rho \neq 0$ at $\partial\Omega$. By smoothness of $\partial\Omega$ we refer to the smoothness of the boundary defining function ρ . (This is equivalent with $\partial\Omega$ being locally a graph of the required smoothness.) In addition, we may assume that $|\nabla\rho(x)| = 1$ for all $x \in \partial\Omega$. The the outer unit normal is given by $\nu(x) = -\nabla\rho(x)$.

Consider a point $(x, v) \in S\Omega$ for which $\varphi_{\tau(x, v)}(x, v)$ points outward (is not tangential). Use the implicit function theorem to show that the function τ is C^1 in a neighborhood of (x, v) . \bigcirc

We define the integral function $u^f: S\bar{\Omega} \rightarrow \mathbb{R}$ of a function $f: S\bar{\Omega} \rightarrow \mathbb{R}$ as

$$u^f(x, v) = \int_0^{\tau(x, v)} f(\varphi_t(x, v)) dt \quad (68)$$

whenever this integral makes sense. In words, $u^f(x, v)$ is the integral of f over the lift of the geodesic starting at the point x in the direction v . This kind of integral function will play a big role in our next proof of injectivity of the X-ray transform.

Exercise 72. Let Ω be the unit ball and $f \equiv 1$ the constant function on $\bar{\Omega}$. Find a formula for $u^f: S\bar{\Omega} \rightarrow \mathbb{R}$. As you will notice, the resulting function has differentiability issues at the tangential part of the boundary. \bigcirc

This integral function satisfies a fundamental theorem of calculus:

Exercise 73. Prove that $Xu^f = -f$ in $S\Omega$ for $f \in C(S\bar{\Omega})$. \bigcirc

Previously we defined the X-ray transform of a scalar function $f: \mathbb{R}^n \rightarrow \mathbb{R}$. With the help of the previous exercise we can in fact define the X-ray transform of a compactly supported continuous function $f: S\mathbb{R}^n \rightarrow \mathbb{R}$. This leads to so-called tensor tomography.

Exercise 74. Do you have any questions or comments regarding section 7? \bigcirc

8 The sphere bundle in two dimensions

8.1 Horizontal and vertical vector fields

To simplify matters, we choose $n = 2$ for this and the next section. Now Ω is a smooth and convex planar domain. The unit sphere bundle $S\Omega = \Omega \times S^1$ could also be called the circle bundle.

When convenient, we may consider a function f on Ω to be a function on $S\Omega$ which just happens to be independent of v . Formally, this amounts to replacing f with π^*f . The pullback is defined as $\pi^*f = f \circ \pi$.

The sphere bundle is three-dimensional. One special direction is given by the geodesic vector field. There is a second horizontal direction and one vertical direction, too, and we will study derivatives in these directions next. We will define the horizontal vector field and the vertical vector field. As in the case of the geodesic vector field, these will be differential operators.

Points $(x, v) \in S\bar{\Omega}$ can be written as (x, v_θ) for $x \in \bar{\Omega}$ and $\theta \in \mathbb{R}/2\pi\mathbb{Z}$, where $v_\theta = (\cos(\theta), \sin(\theta))$. The vertical vector field is simply differentiation with respect to θ . That is, for $u: S\bar{\Omega} \rightarrow \mathbb{R}$ we define $Vu: S\bar{\Omega} \rightarrow \mathbb{R}$ by

$$Vu(x, v_\theta) = \partial_\theta u(x, v_\theta). \quad (69)$$

Exercise 75. Let Ω be the unit disc and write x in polar coordinates. Write the function u^f of exercise 72 in these coordinates (one radius and two angles). Calculate Vu^f . (The derivative will blow at the tangential part of the boundary of the sphere bundle. This is the typical place for regularity issues.)

○

The geodesic vector field may be written as $Xu = v_\theta \cdot \nabla_x u = \cos(\theta)\partial_{x_1} u + \sin(\theta)\partial_{x_2} u$. Let v^\perp denote the rotation of v by $\frac{\pi}{2}$ clockwise. In terms of angles, $v_\theta^\perp = v_{\theta-\pi/2}$.

We define the horizontal vector field X_\perp so that

$$\begin{aligned} X_\perp u(x, v_\theta) &= v_\theta^\perp \cdot \nabla_x u(x, v_\theta) \\ &= \sin(\theta)\partial_{x_1} u(x, v_\theta) - \cos(\theta)\partial_{x_2} u(x, v_\theta). \end{aligned} \quad (70)$$

Exercise 76. Earlier we used v^\perp to denote something else. How are these two v^\perp s related in the two-dimensional setting?

○

Bonus exercise 4. For $f \in C^2(\Omega)$, let π^*f be the pullback over the projection $\pi: S\Omega \rightarrow \Omega$. Show that f is harmonic if and only if $(X^2 + X_\perp^2)\pi^*f = 0$ in $S\Omega$.

○

The geodesic vector field X at (x, v) is the derivative with respect to x in the direction of v . The horizontal vector field X_\perp is the derivative with

respect to x in the direction orthogonal to v . The vertical vector field is the derivative with respect to the direction v .

As vector fields (as opposed to differential operators), these three vector fields are orthogonal and have unit length. They are an orthonormal basis to the tangent spaces of the sphere bundle. This happens on any two-dimensional Riemannian manifold when the sphere bundle is equipped with the so-called Sasaki metric. In higher dimensions the Sasaki metric is trickier, and its definition needs to be taken to a separate course.

In higher dimensions there are still natural horizontal and vertical derivatives, but they are no longer vector fields. To avoid technicalities, we stick to dimension two.

Exercise 77. Show that if $f \in C(\bar{\Omega})$, then $VXu^f = 0$ in $S\Omega$. Here $f \in C(\bar{\Omega})$ is identified with $\pi^*f \in C(S\bar{\Omega})$. \bigcirc

8.2 Commutators

To calculate with differential operators, we need a couple of basic tools. We need to be able to integrate by parts and change the order of differentiation. Integration by parts comes in the next section, and now we will study what happens when the order of differentiation changes. In our situation the order of differentiation does matter, but it only matters to a lower order, so to say. The effect of changing the order is captured by commutators.

The commutator of two linear operators A and B is $[A, B] = AB - BA$. For example, consider the following two operators on functions on the real line:

$$\begin{aligned} Af(x) &= f'(x), \\ Bf(x) &= h(x)f(x), \end{aligned} \tag{71}$$

where h is a sufficiently smooth function. Then

$$\begin{aligned} [A, B]f(x) &= (ABf)(x) - (BAf)(x) \\ &= \frac{d}{dx}(h(x)f(x)) - h(x)f'(x) \\ &= h'(x)f(x). \end{aligned} \tag{72}$$

Exercise 78. Let us call A a first order differential operator on the real line if it is of the form

$$Af(x) = h(x)f'(x) + g(x)f(x) \tag{73}$$

for some smooth functions h and g . Show that the commutator of two first order differential operators is a first order differential operator. \bigcirc

In general, the product of differential operators of orders k and m is a differential operator of order $k+m$, and the commutator has order $k+m-1$. The leading order derivative may vanish, in which case the order is actually lower.

Exercise 79. Show that $[X, V] = X_\perp$. ○

Exercise 80. Show that $[V, X_\perp] = X$. ○

Exercise 81. Show that $[X, X_\perp] = 0$. ○

The case of a two-dimensional Riemannian manifold is surprisingly similar. The first two commutators above stay intact, and $[X, X_\perp]$ contains V and the curvature. In higher dimensions the formulas are a little trickier, but still the same in spirit.

In fact, the only thing we need to know about commutators in the proof is the following lemma. Observe that X_\perp does not appear in the claim, but it is useful for the proof.

Lemma 8.1. *Our vector fields satisfy*

$$[XV, VX] = -X^2. \quad (74)$$

Proof. We simply use our commutator formulas and calculate:

$$\begin{aligned} [XV, VX] &= XVVX - VXXV \\ &= (VX + X_\perp)VX - VX(VX + X_\perp) \\ &= X_\perp VX - VXX_\perp \\ &= [X_\perp, VX] \\ &= [X_\perp, V]X + V[X_\perp, X] \\ &= -X^2. \end{aligned} \quad (75)$$

Here we used the commutator property $[A, BC] = [A, B]C + B[A, C]$, which is trivial to verify by hand. There are several ways to go about this calculation; the intermediate steps here are but an example. □

The commutator of two second order operators is typically of third order. In this particular case it happens to be second order because XV and VX only differ by the first order operator X_\perp .

¹⁰

¹⁰See section A.14 for more exercises.

8.3 Integration on the sphere bundle

The sphere bundle is a product space, and we can naturally use the product measure Σ . Therefore the integral of $g \in C(S\bar{\Omega})$ is

$$\int_{S\Omega} g(x, v) d\Sigma(x, v) = \int_{\Omega} \int_{S^1} g(x, v) dS(v) dx. \quad (76)$$

Alternatively, the S^1 integral can be written as $\int_0^{2\pi} g(x, v_\theta) d\theta$.

The boundary $\partial\Omega$ is a closed smooth curve, and we have a natural measure on it. One way to describe it is to write the curve as $\alpha: [0, L] \rightarrow \mathbb{R}^2$ with arc length parametrization and then integrate on the interval $[0, L]$.

This gives rise to a measure $\tilde{\sigma}$ on $\partial(S\Omega)$, given by

$$\int_{\partial(S\Omega)} g d\tilde{\sigma} = \int_{S^1} \int_0^L g(\alpha(t), v) dt dS(v) \quad (77)$$

for any $g \in C(S\bar{\Omega})$.

It turns out that the measure $\sigma = |v \cdot \nu(x)| \tilde{\sigma}$ is more natural. It will appear in a change of variables formula for integration over the sphere bundle.

Proposition 8.2 (Santaló's formula). *Let $\Omega \subset \mathbb{R}^2$ be a convex, bounded, and smooth domain. For $g \in C(S\bar{\Omega})$ we have*

$$\int_{S\Omega} g d\Sigma = \int_{\partial_{in}(S\Omega)} \int_0^{\tau(x,v)} g(\varphi_t(x, v)) dt d\sigma(x, v). \quad (78)$$

Alternatively, the integral can be taken over the entire $\partial(S\Omega)$ since τ vanishes outside $\partial_{in}(S\Omega)$.

Proof. First, we change the order of integration in the integral over $S\Omega$:

$$\begin{aligned} \int_{S\Omega} g d\Sigma &= \int_{\Omega} \int_{S^1} g(x, v) dS(v) dx \\ &= \int_{S^1} \int_{\Omega} g(x, v) dx dS(v). \end{aligned} \quad (79)$$

Now fix any $v \in S^1$ and consider the inner integral

$$I(v) = \int_{\Omega} g(x, v) dx. \quad (80)$$

We extend g to $S\mathbb{R}^2$ by zero for convenience. We write the plane as an orthogonal direct sum $\mathbb{R}^2 = v\mathbb{R} \oplus v^\perp\mathbb{R}$. With this decomposition, we have

$$I(v) = \int_{\mathbb{R}} \int_{\mathbb{R}} g(sv^\perp + tv, v) dt ds. \quad (81)$$

The inner integral is an integral along the geodesic flow as desired. We will turn the outer integral into an integral over the boundary.

Let us denote

$$\partial_v \Omega = \{x \in \partial \Omega; v \cdot \nu(x) < 0\}. \quad (82)$$

We parametrize this part of the boundary by a (counterclockwise) unit speed curve $\beta: [0, L_v] \rightarrow \partial_v \Omega$. Observe that $\{(x, v); v \in S^1, x \in \partial_v \Omega\} = \partial_{\text{in}}(S\Omega)$.

Let us denote

$$a(v) = \min\{s \in \mathbb{R}; (sv^\perp + v\mathbb{R}) \cap \bar{\Omega} \neq \emptyset\} \quad (83)$$

and

$$b(v) = \max\{s \in \mathbb{R}; (sv^\perp + v\mathbb{R}) \cap \bar{\Omega} \neq \emptyset\}. \quad (84)$$

Now, there is a function $w: (a(v), b(v)) \rightarrow (0, L_v)$ so that $\beta(w(s)) - sv^\perp \in v\mathbb{R}$. In fact, w is a C^1 diffeomorphism with $w'(s) = -1/v \cdot \nu(\beta(w(s))) > 0$. The details are left as an exercise.

We change the variable of integration from s to $z = w(s)$ and obtain

$$\begin{aligned} I(v) &= \int_{\mathbb{R}} \int_{\mathbb{R}} g(sv^\perp + tv, v) dt ds \\ &= \int_{a(v)}^{b(v)} \int_{\mathbb{R}} g(sv^\perp + tv, v) dt ds \\ &= \int_0^{L_v} \int_{\mathbb{R}} g(w^{-1}(z)v^\perp + tv, v) dt (-v \cdot \nu(\beta(z))) dz \\ &= \int_0^{L_v} \int_0^{\tau(\beta(z), v)} g(\beta(z) + tv, v) dt |v \cdot \nu(\beta(z))| dz. \end{aligned} \quad (85)$$

Since β is a subcurve of α (restriction to a subinterval, possibly after rechoosing the initial and final point on α) and $\tau = 0$ on the part $\alpha \setminus \beta$, we get

$$I(v) = \int_0^L \int_0^{\tau(\alpha(z), v)} g(\alpha(z) + tv, v) dt |v \cdot \nu(\alpha(z))| dz. \quad (86)$$

Combining (79), (80), and (86), we find

$$\begin{aligned} \int_{S\Omega} g d\Sigma &= \int_{S^1} \int_0^L \int_0^{\tau(\alpha(z), v)} g(\alpha(z) + tv, v) dt |v \cdot \nu(\alpha(z))| dz dS(v) \\ &= \int_{\partial(S\Omega)} \left(\int_0^{\tau(x, v)} g(x + tv, v) dt \right) |v \cdot \nu(x)| d\tilde{\sigma}(x, v) \\ &= \int_{\partial(S\Omega)} \left(\int_0^{\tau(x, v)} g(\varphi_t(x, v)) dt \right) d\sigma(x, v) \end{aligned} \quad (87)$$

as claimed. \square

Exercise 82. Draw a picture or two and explain what the function w does. ○

Exercise 83. Explain why $w \in C^1$ and $w'(s) = -1/v \cdot \nu(\beta(w(s))) > 0$. ○

Exercise 84. Show that the measure of Ω is $\frac{1}{2\pi} \int_{\partial_{\text{in}}(S\Omega)} \tau(x, v) \, d\sigma(x, v)$. ○

Santaló's formula states that the integral over the sphere bundle can be calculated by calculating it one geodesic at a time, first integrating over the (lifted) geodesic and then integrating over the initial points and directions of these geodesics at $\partial_{\text{in}}(S\Omega)$. The space of all geodesics through Ω can be identified with $\partial_{\text{in}}(S\Omega)$. The measure spaces (Γ, μ) and $(\partial_{\text{in}}(S\Omega), \sigma)$ (with Borel σ -algebras) are two descriptions of the same thing.

The formula is a change of variables. It will help us find Green-type formulas for our three vector fields (see exercises below), and those will lead to integration by parts on the sphere bundle.

The formula also gives rise to some integral properties which will be convenient.

Exercise 85. Show that if $g \in C^1(S\bar{\Omega})$, then

$$\int_{S\Omega} Vg \, d\Sigma = 0. \quad (88)$$

Santaló is not needed. ○

Exercise 86. Show that if $g \in C^1(S\bar{\Omega})$, then

$$\int_{S\Omega} Xg \, d\Sigma = \int_{\partial_{\text{in}}(S\Omega)} (g(\varphi_{\tau(x,v)}(x, v)) - g(x, v)) \, d\sigma(x, v). \quad (89)$$

Santaló is useful. ○

Exercise 87. Show that if $g \in C^2(S\bar{\Omega})$ and $g|_{\partial(S\Omega)} = 0$, then

$$\int_{S\Omega} X_{\perp}g \, d\Sigma = 0. \quad (90)$$

Use the previous two exercises and the commutator formulas. ○

Exercise 88. Do you have any questions or comments regarding section 8? ○

9 X-ray tomography and the transport equation

9.1 Integration revisited

For $g, h \in L^2(S\Omega, \Sigma)$ we write

$$\langle g, h \rangle = \int_{S\Omega} gh \, d\Sigma \quad (91)$$

and $\|g\| = \sqrt{\langle g, g \rangle}$. Our functions in this section are real-valued so no conjugation is needed. The complex case is not harder, but the real approach is technically convenient.

Santaló's formula (proposition 8.2) makes it easy to find integration by parts formulas for our three vector fields.

Lemma 9.1. *For any $g, h \in C_c^\infty(S\Omega)$ we have*

$$\begin{aligned} \langle g, Xh \rangle &= -\langle Xg, h \rangle, \\ \langle g, Vh \rangle &= -\langle Vg, h \rangle, \quad \text{and} \\ \langle g, X_\perp h \rangle &= -\langle X_\perp g, h \rangle. \end{aligned} \quad (92)$$

Exercise 89. Prove the lemma using results from the previous section. The next exercise is also useful. \bigcirc

11

9.2 A second order PDE

Now we finally begin our analysis of the X-ray transform. We want to show that if $f \in C_c^\infty(\Omega)$ integrates to zero over all lines, then $f = 0$.

First, let us define the integral function u^f as above. Recall from exercise 77 that $VXu^f = 0$. In addition, u^f satisfies the boundary condition $u^f|_{\partial(SM)} = 0$. For outward pointing and tangential directions, this is because $\tau = 0$. For inward pointing directions, this is because $\mathcal{I}f = 0$.

Exercise 90. Explain how to identify the functions $\mathcal{I}f$ and $u^f|_{\partial_{\text{in}}(S\Omega)}$. \bigcirc

The function $u = u^f$ solves the boundary value problem

$$\begin{cases} VXu = 0 & \text{in } S\Omega \\ u = 0 & \text{on } \partial(S\Omega). \end{cases} \quad (93)$$

¹¹See section A.14 for more exercises.

Clearly $u = 0$ is a solution. If the solutions of this second order PDE are unique, then it follows that $u^f = 0$ and therefore $f = -Xu^f = 0$. This leads to injectivity of the X-ray transform.

The operator XV is not elliptic, hyperbolic, or parabolic. Therefore we do not have access to standard uniqueness theorems, and we have to show uniqueness by hand.

¹²

The order of the operators is crucial. The PDE $VXu = 0$ will turn out to have unique solutions, but $XVu = 0$ never does. If $u(x, v) = h(x)$ for $h \in C^\infty(\bar{\Omega})$ with $h|_{\partial\Omega} = 0$, then $Vu = 0$ and the boundary condition is also satisfied. In fact, functions of this form are precisely the solutions of $XVu = 0$ with zero boundary values, as follows from the next exercise.

Exercise 91. Suppose $g \in C^1(S\bar{\Omega})$. Show that if $g|_{\partial_{\text{in}}(S\Omega)} = 0$ and $Xg = 0$, then $g = 0$.

Also, explain why $g|_{\partial_{\text{in}}(S\Omega)} = 0$ implies $Vg|_{\partial_{\text{in}}(S\Omega)} = 0$. ○

¹³

9.3 Properties of the integral function

Before studying the boundary value problem (93) further, it is good to verify that u^f is sufficiently regular.

Lemma 9.2. *The function u^f defined above is in $C_c^\infty(S\Omega)$.*

Proof. Recall exercise 71. If the boundary $\partial\Omega$ is smooth, it follows with the same argument and the smooth version of the implicit function theorem that τ is smooth near $(x, v) \in SM$ when $\varphi_{\tau(x,v)}(x, v)$ points outward. The function u^f is defined by

$$u^f(x, v) = \int_0^{\tau(x,v)} f(x + tv) dt. \quad (94)$$

Since f and τ are smooth, so is u^f . (We omit some technical details here, but the statement is hopefully plausible to the reader. One can also extend f by zero to \mathbb{R}^2 to get rid of the τ .)

So far we have not used the fact that f is compactly supported nor tried to prove that so is u^f . Also, smoothness at tangential exits has not been established yet.

¹²See section A.14 for more exercises.

¹³See section A.14 for more exercises.

Assume f is supported in a compact set $K \subset \Omega$. By exercise 92 we may assume that K is convex. We will show that u^f is supported in $SK = K \times S^1$.

This will also prove smoothness near points $(x, v) \in S\Omega$ where $\varphi_{\tau(x,v)}(x, v)$ is tangential to $\partial\Omega$; see exercise 93. Therefore it only remains to prove the support condition for u^f .

Take any $x \in \bar{\Omega} \setminus K$ and $v \in S^1$. Let

$$\gamma(x, v) = \{x + tv; t \in [0, \tau(x, v)]\} \quad (95)$$

be the line from x to $\partial\Omega$ in the direction of v . If $\gamma(x, v) \cap K = \emptyset$, then $u^f(x, v) = 0$ since $u^f(x, v)$ is the integral of f over $\gamma(x, v)$.

Because K is convex and $x \notin K$, at most one of the line segments $\gamma(x, v)$ and $\gamma(x, -v)$ can meet K . Thus if $\gamma(x, v) \cap K \neq \emptyset$, then $\gamma(x, -v) \cap K = \emptyset$. By the argument given above, $u^f(x, -v) = 0$.

On the other hand, $u^f(x, v) + u^f(x, -v) = 0$ for all $(x, v) \in S\Omega$ since $\mathcal{I}f = 0$, so $u^f(x, -v) = 0$ implies $u^f(x, v) = 0$. We have thus shown that $u^f(x, v) = 0$ when $x \notin K$. \square

Exercise 92. Suppose $\Omega \subset \mathbb{R}^n$ is a convex open set and $K \subset \Omega$ compact. Show that the convex hull of K is compact and contained in Ω . \circ

Exercise 93. Suppose $\Omega \subset \mathbb{R}^n$ is a bounded and convex C^1 domain. Suppose $(x, v) \in S\bar{\Omega}$ is such that $\varphi_{\tau(x,v)}(x, v)$ is tangential to $\partial\Omega$. Show that $x \in \partial\Omega$ and $v \cdot \nu(x) = 0$. \circ

We have chosen to work with compactly supported smooth functions to avoid technical difficulties. The same method works for $f \in C^2(\bar{\Omega})$ as well, with no assumptions on boundary values. This would require more delicate analysis of boundary behaviour, since in general $u^f \notin C^2(S\bar{\Omega})$ even if $f \in C^2(\bar{\Omega})$. In fact, without assuming $\mathcal{I}f = 0$, one only has $u^f \in C^{1/2}(S\bar{\Omega})$; see exercise 72.

We will next prove an integral identity. The statement concerns second and first order derivatives, but the proof uses derivatives up to order four. In cases like this the theorem can be shown to hold in C^2 using the density of C^∞ in C^2 .

9.4 The Pestov identity

The key to proving uniqueness of (93) is an integral identity known as the Pestov identity. It was introduced by Mukhometov, and it could well be called the Mukhometov–Pestov identity.

Proposition 9.3 (Pestov identity). *If $u \in C_c^\infty(S\Omega)$, then*

$$\|VXu\|^2 = \|XVu\|^2 + \|Xu\|^2. \quad (96)$$

Proof. Using lemmas 9.1 and 8.1 we find

$$\begin{aligned}
 \|VXu\|^2 - \|XVu\|^2 &= \langle VXu, VXu \rangle - \langle XVu, XVu \rangle \\
 &= -\langle VVXu, Xu \rangle + \langle XXVu, Vu \rangle \\
 &= \langle XVVXu, u \rangle - \langle VXXVu, u \rangle \\
 &= \langle (XVVX - VXXV)u, u \rangle \\
 &= \langle [XV, VX]u, u \rangle \\
 &= \langle -X^2u, u \rangle \\
 &= \langle Xu, Xu \rangle
 \end{aligned} \tag{97}$$

as claimed. \square

As mentioned earlier, the assumption of compact support is not necessary. It is enough that $u|_{\partial(S\Omega)} = 0$, but the proof would be somewhat more technical. Smoothness is not necessary either, it is just convenient.

¹⁴

Bonus exercise 5. Stare at the Pestov identity, experience enlightenment, and explain what it means and why it should hold true. \bigcirc

The Pestov identity makes proving our injectivity result easy:

Theorem 9.4. *Let $\Omega \subset \mathbb{R}^2$ be a bounded, smooth, and strictly convex domain. If $f \in C_c^\infty(\Omega)$ satisfies $\mathcal{I}f = 0$, then $f = 0$.*

Exercise 94. Prove theorem 9.4 by applying proposition 9.3 to u^f . \bigcirc

This argument provides us with yet another uniqueness proof. However, it does not give an inversion formula for the X-ray transform. There are inversion formulas within this framework, but finding one requires considerably more work than proving uniqueness.

9.5 Remarks about manifolds

This method can also be used to prove injectivity results on many Riemannian manifolds with boundary. The sphere bundle and the derivatives on it are still well defined and useful. The Santaló formula still holds true. To be able to use the Pestov identity, the integral function u^f needs to be regular enough, and the right-hand side of the identity needs to be non-negative.

To obtain convenient regularity, the manifold is typically assumed to be compact and have strictly convex boundary. Strict convexity is defined in terms of the curvature of the boundary.

¹⁴Another exercise belongs here. See the appendix.

The Pestov identity on a two-dimensional manifold M with boundary reads

$$\|VXu\|^2 = \|XVu\|^2 + \|Xu\|^2 - \int_{SM} K |Vu|^2 d\Sigma, \quad (98)$$

where K is the Gaussian curvature of the surface. If $K \leq 0$, then the desired positivity result follows. Indeed, the X-ray transform is injective on non-positively curved surfaces with strictly convex boundary.

In higher dimensions things are somewhat more complicated. Vertical and horizontal derivatives are no longer given by vector fields, and gradient-like operators are needed instead. In addition, curvature can no longer be adequately described with a scalar function. However, there is a Pestov identity and it can be used to prove similar results. Positivity now depends on the sectional curvature.

These results can be generalized in various ways. Development and application of the relevant tools in differential geometry require a separate course.

Exercise 95. Do you have any questions or comments regarding section 9? ○

10 X-ray tomography of vector fields

10.1 Definition of the ray transform

So far we have only discussed X-ray tomography of scalar functions $f: \mathbb{R}^n \rightarrow \mathbb{R}$. We can ask a similar question for other kinds of functions as well, and we will only explore one generalization in this course: X-ray tomography of vector fields.

A vector field in the Euclidean space is a function $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$. The integral of f over a line $\gamma: \mathbb{R} \rightarrow \mathbb{R}^n$ is the X-ray transform

$$\mathcal{I}f(\gamma) = \int_{\mathbb{R}} f(\gamma(t)) \cdot \dot{\gamma}(t) dt \quad (99)$$

whenever this integral exists. (Some people call the transform something else in the case of vector fields. We do not.) We will continue to use unit speed parametrization, although in this particular case it does not make a difference. Those familiar with differential forms may identify a vector field with a one-form, and the integral of a k -form over an oriented k -dimensional manifold is parametrization invariant.

Exercise 96. Prove that formula (99) for the X-ray transform of a vector field is in fact invariant under any orientation-preserving reparametrization.

What happens if orientation is flipped? Does the integral of a scalar function change if you reparametrize it or change orientation? \bigcirc

We can now ask our main question in this new setting: Does $\mathcal{I}f$ determine f uniquely? In other words, if $\mathcal{I}f = \mathcal{I}g$, do we then necessarily have $f = g$?

10.2 An application

Generalizing mathematical questions is commonplace, and a mathematician may not need any further motivation for this variant of the problem. While mathematical interest might be sufficient reason for this detour, we will also present one physical application. The applications of ray tomography of scalar and vector fields are not limited to what is mentioned in this course, and some applications call for further generalizations.

Consider a stationary flow of a liquid, described by the flow field $u: \mathbb{R}^3 \rightarrow \mathbb{R}^3$. That is, at the point x the liquid flows with velocity $u(x)$. The speed of sound can be described by a scalar field $c: \mathbb{R}^3 \rightarrow (0, \infty)$, but we assume that it is constant.

If $|u(x)| \ll c$ for all x , then sound waves in the moving liquid travel at roughly straight lines, but their speed along those lines are changed by u . (We can consider the flow to be a small perturbation to the completely still reference situation. Travel time has first order dependence on u , but trajectories depend on u only in the second order. Therefore in the linearized problem geometry is unchanged but travel times change. We will not attempt to make this linearization procedure precise.)

Consider a straight line $\gamma: [0, L] \rightarrow \mathbb{R}^3$ parametrized by arc length. If $u = 0$, then the time to travel from $\gamma(0)$ to $\gamma(L)$ is

$$\int_0^L \frac{1}{c} ds. \quad (100)$$

The presence of u changes this to

$$\int_0^L \frac{1}{c + u(\gamma(s)) \cdot \dot{\gamma}(s)} ds \approx L - \int_0^L \frac{u(\gamma(s)) \cdot \dot{\gamma}(s)}{c^2} ds. \quad (101)$$

Therefore the (linearized) travel time measurement determines $\mathcal{I}u$.

Linearized travel time tomography often leads to X-ray tomography, but the unknown objects may or may not be scalar functions.

The physical problem is then whether such time-of-flight measurements determine the flow field u . Does it help if the fluid is incompressible ($\nabla \cdot u = 0$)?

10.3 Non-uniqueness and potentials

It turns out that the answer to our main question is no: a vector field f is not uniquely determined by $\mathcal{I}f$. There are vector fields f that are not identically zero but for which $\mathcal{I}f = 0$.

The next best thing to ask for is a characterization of the kernel of the X-ray transform. Can we characterize the set of those f for which $\mathcal{I}f = 0$?

There is a special class of vector fields we study first: gradient fields. If $h: \mathbb{R}^n \rightarrow \mathbb{R}$ is a smooth scalar function, then ∇h is a smooth vector field. Let us calculate the X-ray transform of such a vector field.

Exercise 97. Let $h: \mathbb{R}^n \rightarrow \mathbb{R}$ be a smooth scalar function and $\gamma: [0, L] \rightarrow \mathbb{R}^n$ a line. Show that

$$\int_0^L \nabla h(\gamma(t)) \cdot \dot{\gamma}(t) dt = h(\gamma(L)) - h(\gamma(0)). \quad (102)$$

Explain why, if $h \in C_c^\infty(\mathbb{R}^n)$, then $\mathcal{I}(\nabla h)$ is well defined and identically zero. \bigcirc

This means that there is a freedom to change a vector field f to $f + \nabla h$ without changing $\mathcal{I}f$ at all. This is called a gauge freedom.

We now ask a refined question: If a sufficiently nice vector field $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisfies $\mathcal{I}f = 0$, then is there a scalar function h so that $f = \nabla h$?

The answer to this refined question is indeed positive, and we will prove it in one special case. For simplicity, we will only prove the result in two dimensions. In exercise 10 we saw that for scalar functions the higher dimensional result follows from the two one in dimension two. The same argument works here, too:

Exercise 98. Suppose this is known: If a compactly supported smooth vector field f on \mathbb{R}^2 satisfies $\mathcal{I}f = 0$, then there is a smooth compactly supported scalar function h on the plane so that $f = \nabla h$.

Show this: A smooth compactly supported vector field f on \mathbb{R}^n , $n \geq 2$, satisfies $\mathcal{I}f = 0$ if and only if there is a smooth compactly supported scalar function h so that $f = \nabla h$. \bigcirc

In three dimensions one can write a vector field as a sum of a gradient field and a solenoidal (divergence-free) vector field in a unique way. This is known as the Helmholtz decomposition. There is an analogous decomposition in higher dimensions and also on manifolds, known as the Hodge decomposition.

The X-ray transform of the gradient component is always zero, but the rest is uniquely determined as we shall see. This kind of result is known as solenoidal injectivity. In particular, it follows that a solenoidal vector field is uniquely determined by its X-ray transform. Our physical example problem

is indeed uniquely solvable under the additional assumption that the flow is incompressible (solenoidal).

10.4 Solenoidal injectivity

We will now prove solenoidal injectivity in two dimensions by making use of the Pestov identity.

Let $\Omega \subset \mathbb{R}^2$ be a bounded, smooth, and strictly convex domain. A vector field $f: \Omega \rightarrow \mathbb{R}^2$ can be regarded as a function \tilde{f} on $S\mathbb{R}^2$ as $\tilde{f}(x, v) = f(x) \cdot v$. We can define the integral function u^f in two ways, by considering f as a function on $S\Omega$ or by using an integral formula like (99). These two approaches lead to exactly the same function: $u^f = u^{\tilde{f}}$. (These notes attempt to distinguish the vector field f on Ω and the function \tilde{f} on $S\Omega$ consistently, but be prepared for failures.)

We assume that $\mathcal{I}f = 0$. An inspection of the proof of lemma 9.2 shows that $u^f \in C_c^\infty(S\Omega)$ also in the case of vector fields. The fundamental theorem of calculus of exercise 73 is still valid when f is seen as a function on $S\Omega$. The same proof gives that $Xu^f(x, v) = -\tilde{f}(x, v) = -f(x) \cdot v$. However, now f does depend on direction, and so typically $V\tilde{f} \neq 0$. This causes a major change in our proof and result.

Exercise 99. Let f be a smooth vector field on \mathbb{R}^2 , and define a function $\tilde{f}: S\mathbb{R}^2 \rightarrow \mathbb{R}$ by $\tilde{f}(x, v) = f(x) \cdot v$. Calculate $V\tilde{f}(x, v)$ and interpret the result geometrically. \circ

Exercise 100. Consider the function u^f on $S\Omega$ defined above for a vector field f with $\mathcal{I}f = 0$. Show that $\|VXu^f\| = \|Xu^f\|$. \circ

Theorem 10.1. *Let $\Omega \subset \mathbb{R}^2$ be a bounded, smooth and convex domain. If a compactly supported smooth vector field $f: \Omega \rightarrow \mathbb{R}^2$ integrates to zero over all lines, then there is $h \in C_c^\infty(\Omega)$ so that $f = \nabla h$.*

Proof. As discussed above, the integral function u^f is compactly supported and smooth, so we may apply the Pestov identity:

$$\|VXu^f\|^2 = \|XVu^f\|^2 + \|Xu^f\|^2. \quad (103)$$

By exercise 100, this leads to

$$0 = \|XVu^f\|^2. \quad (104)$$

This implies that the function $XVu^f \in C_c^\infty(S\Omega)$ is identically zero.

Since u^f vanishes in a neighborhood of $\partial(S\Omega)$, so does Vu^f . The function Vu^f has zero boundary values and is annihilated by the geodesic vector

field, so it has to vanish identically. Both of these conclusions follow from exercise 91.

Because $Vu^f = 0$, the function $u^f(x, v)$ is in fact independent of v . Recall that V is the derivative with respect to $v \in S^1$ and S^1 is connected. Therefore there is a scalar function $h \in C_c^\infty(\Omega)$ so that $u^f(x, v) = -h(x)$.

Using $Xu^f = -\tilde{f}$, it follows that $f = \nabla h$. The details are left as an exercise. \square

Exercise 101. Complete the proof above by showing that $f = \nabla h$. \circ

In the proof above we needed to produce a potential h for the vector field f . The potential turned out to be essentially the integral function u^f . The hardest part was showing that u^f can be considered as a scalar function on Ω .

Exercise 102. Let f_0 be a scalar function and f_1 a vector field. Their sum is not a very reasonable object at first, and it can be considered just as a formal sum. How can you consider $f_0 + f_1$ as a function on $S\mathbb{R}^n$? How should we define $\mathcal{I}(f_0 + f_1)$?

Assume now that f_0 and f_1 are smooth and compactly supported. Using previously obtained results, argue why $\mathcal{I}(f_0 + f_1) = 0$ implies that $f_0 = 0$ and $f_1 = \nabla h$ for some $h \in C_c^\infty(\mathbb{R}^n)$. \circ

Let us then see what this solenoidal injectivity result means for reconstructing a vector field from data. If f and g are vector fields (or sums of scalars and vector fields) and $\mathcal{I}f = \mathcal{I}g$, then there is a scalar potential h vanishing at the boundary so that $f = g + \nabla h$ (and the scalar parts of f and g coincide).

We chose to use the Pestov identity, but it is not the only way to prove this statement in a Euclidean space. We remark that the same proof works for non-positively curved Riemannian manifolds of dimension two with strictly convex boundary.

Let us see what happens in one dimension. As we have seen before, the X-ray transform is not injective in one dimension scalar functions. There are non-trivial continuous functions $f: [0, 1] \rightarrow \mathbb{R}$ so that the X-ray transform vanishes.

In one dimension scalar functions and vector fields are essentially the same object, but the solenoidal injectivity requires less than full injectivity. The conclusion might be surprising: we have no injectivity for scalars, but we do have solenoidal injectivity for vector fields.

¹⁵Another exercise belongs here. See the appendix.

Exercise 103. Do you have any questions or comments regarding section 10?



11 The Fourier transform

11.1 A general view to Fourier transforms

Previously we studied the Fourier transform on a torus \mathbb{T}^n . The Fourier transform took a function on the torus \mathbb{T}^n to a function on the lattice \mathbb{Z}^n , and the inverse Fourier transform did the opposite. One could in fact define a the Fourier transform on the lattice, and that would turn out to be essentially the same as the inverse Fourier transform for the torus.

In this section we will study the Fourier transform on \mathbb{R}^n . It will take a function $\mathbb{R}^n \rightarrow \mathbb{C}$ to another function $\mathbb{R}^n \rightarrow \mathbb{C}$, and the inverse transform is very similar to the transform itself. Before going any deeper into this, we will look at the Fourier transforms in far greater generality to see that the two Fourier transforms in this course are merely two special cases of a far more general structure.

Let G be a topological group. It means that it is a topological space and a group so that the group operations $G \rightarrow G, x \mapsto x^{-1}$ and $G \times G \rightarrow G, (x, y) \mapsto xy$ are continuous. We assume that G is abelian and locally compact.

Let \hat{G} denote the set of all continuous homomorphisms $G \rightarrow S^1$. Here $S^1 \subset \mathbb{C}$ is considered as the multiplicative group of unit complex numbers. Elements of \hat{G} are called characters of G .

The set \hat{G} can be endowed with a group structure by pointwise multiplication of characters: $(\alpha\beta)(x) = \alpha(x)\beta(x)$. The set \hat{G} is a set of functions, and it can be equipped with the topology of locally uniform convergence. These structures make \hat{G} into a topological group.

What is important is that $\hat{\hat{G}}$ is also a locally compact abelian group and that $\hat{\hat{G}}$ is naturally isomorphic to G . This result is known as the Pontryagin duality theorem and \hat{G} is called the dual group of G .

The Fourier transform takes a function on G into a function on \hat{G} , and the inverse Fourier transform reverses this. In light of the duality theorem, it is not surprising that the inverse Fourier transform for G resembles the Fourier transform for \hat{G} . The Fourier transform requires a measure on the underlying space, and that is the Haar measure. When G and \hat{G} are equipped with compatible Haar measures, the L^2 theory (and much more) of Fourier transforms can be extended to any locally compact abelian groups.

Let us make the dual groups a little more concrete with examples. The

dual group of \mathbb{T}^n is \mathbb{Z}^n and vice versa, and this we encountered earlier with Fourier series. The dual group of \mathbb{R}^n is \mathbb{R}^n itself, and this we will study now. The measures on \mathbb{T}^n and \mathbb{R}^n are the Lebesgue measures, and the one on the lattice is the counting measure.

Exercise 104. An element $k \in \mathbb{Z}^n$ can be identified with a character $\chi_k: \mathbb{T}^n \rightarrow S^1$ by $\chi_k(x) = e^{ik \cdot x}$.

Show that for any $k \in \mathbb{Z}^n$ the corresponding character χ_k is indeed a well defined and continuous homomorphism. Recall that $\mathbb{T}^n = \mathbb{R}^n / 2\pi\mathbb{Z}^n$.

How can you identify a point x on the torus \mathbb{T}^n with a character $\psi_x \in \widehat{\mathbb{Z}^n}$? No need to prove anything; just give the formula. \bigcirc

Note that these characters were used in the formulas for the Fourier transform and its inverse on the torus. This is how Fourier transforms work in general, by integrating a function against a character.

If G is not abelian, then \hat{G} should be replaced with (equivalence classes of) irreducible representations of G . This coincides with the dual group in the abelian case since irreducible complex representations of abelian groups are one-dimensional. Moreover, one-dimensional representations coincide with their characters, so the characters introduced here are the same as the representation theoretic characters. Fourier analysis on non-abelian groups is possible via representation theory.

Finally, we remark that there are several different conventions for the Fourier transform on a torus or a Euclidean space. The differences concern the placement of factors of 2π . It is impossible to get completely rid of the factors.

11.2 The Fourier transform on a Euclidean space

It is typical to call the Fourier transform on a torus the Fourier series and the one on a Euclidean space the Fourier transform. Fourier analysis on other groups is much rarer.

The Fourier transform of a function $f: \mathbb{R}^n \rightarrow \mathbb{C}$ is $\mathcal{F}f: \mathbb{R}^n \rightarrow \mathbb{C}$ defined by

$$\mathcal{F}f(\xi) = \int_{\mathbb{R}^n} e^{-i\xi \cdot x} f(x) \, dx \quad (105)$$

whenever this integral makes sense. Again, we are purposely vague since the definition can be extended to various classes of functions or distributions.

Theorem 11.1. *The Fourier transform is a bijection $\mathcal{F}: L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)$, given by*

$$\mathcal{F}f(\xi) = \int_{\mathbb{R}^n} e^{-i\xi \cdot x} f(x) \, dx \quad (106)$$

for $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and extended by continuity to the rest of $L^2(\mathbb{R}^n)$.
The inverse Fourier transform $\mathcal{F}^{-1}: L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)$ is given by

$$(\mathcal{F}^{-1}f)(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{i\xi \cdot x} f(\xi) \, d\xi. \quad (107)$$

The Fourier transform is unitary in the sense that

$$\int_{\mathbb{R}^n} \overline{g(x)} f(x) \, dx = (2\pi)^{-n} \int_{\mathbb{R}^n} \overline{\mathcal{F}g(\xi)} \mathcal{F}f(\xi) \, d\xi. \quad (108)$$

Again, the proof will be omitted.

Exercise 105. What is the relation between $\|f\|_{L^2}$ and $\|\mathcal{F}f\|_{L^2}$? ○

To simplify matters, we will apply the Fourier transform to compactly supported continuous functions. What we need to know is that $\mathcal{F}f = 0$ implies $f = 0$. Under additional assumptions very little information on $\mathcal{F}f$ is needed to conclude that $f = 0$, and we will study this next.

11.3 A Paley–Wiener theorem

A general and important phenomenon in Fourier analysis is the correspondence between decay and regularity. Fast decay of $f(x)$ as $|x| \rightarrow \infty$ corresponds to high regularity of $\mathcal{F}f$ and vice versa. For a famous example the Schwartz space contains by definition functions which have high regularity (infinitely differentiable) and fast decay (all derivatives vanish faster than $|x|^{-N}$ for any $N \in \mathbb{N}$), and the Fourier transform of the Schwartz function space is precisely the space itself.

We will study the ultimate form of decay at infinity: compact support. This should lead to very high regularity, and that turns out to be the case. Our theorem in this subsection is a version of the Paley–Wiener theorem.

Definition 11.2. A function $f: \mathbb{R}^n \rightarrow \mathbb{C}$ is called real analytic if it is smooth and for every point $x \in \mathbb{R}^n$ there is $r > 0$ so that the Taylor series of f around x converges to f in $B(x, r)$.

In complex analysis one can define analyticity in a similar fashion by demanding that a complex Taylor series converges to the function in a small neighborhood of any point. This turns out to be equivalent with complex differentiability (the existence of the derivative as a limit of a difference quotient). When working over the reals this is no longer the case; real analyticity is far stronger than differentiability.

Exercise 106. Show that if a real analytic function $f: \mathbb{R}^n \rightarrow \mathbb{C}$ vanishes in a non-empty open set $U \subset \mathbb{R}^n$, then f is identically zero. ○

The main result of this section is this:

Theorem 11.3. *The Fourier transform of a compactly supported function $f \in L^1(\mathbb{R}^n)$ is real analytic.*

Let us collect some tools before the proof. First, recall a lemma from measure and integration theory:

Lemma 11.4. *Fix integers $1 \leq j \leq n$. Consider a function $g: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{C}$. Suppose that for every $y \in \mathbb{R}^n$ we have $g(\cdot, y) \in L^1(\mathbb{R}^n)$, that for every $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$ the partial derivative $\partial_{y_j} g(x, y)$ exists, and that there is a function $h \in L^1(\mathbb{R}^n)$ so that $|\partial_{y_j} g(x, y)| \leq h(x)$ for all $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$. Then the function*

$$G(y) = \int_{\mathbb{R}^n} g(x, y) \, dx \quad (109)$$

has the partial derivative $\partial_{y_i} G(y)$ everywhere and

$$\partial_{y_i} G(y) = \int_{\mathbb{R}^n} \partial_{y_i} g(x, y) \, dx, \quad (110)$$

where the last integral is a well-defined Lebesgue integral.

Exercise 107. Suppose $f \in L^1(\mathbb{R}^n)$ vanishes outside a compact set K . Denote by f_j the function $f_j(x) = x_j f(x)$. Show that $\partial_{\xi_j} \mathcal{F}f(\xi) = -i\mathcal{F}f_j(\xi)$ and the partial derivative exists. \circ

Similarly, one can find that for any vector $v \in \mathbb{C}^n$ one has

$$v \cdot \nabla \mathcal{F}f(\xi) = -i\mathcal{F}f_v(\xi), \quad (111)$$

where $f_v(x) = (v \cdot x)f(x)$.

Exercise 108. Suppose $f \in L^1(\mathbb{R}^n)$ vanishes outside a compact set K . Argue that $\mathcal{F}f \in C^\infty$. \circ

Exercise 109. Suppose $f \in L^1(\mathbb{R}^n)$ vanishes outside a compact set K . Show that $\mathcal{F}f(\xi) = \mathcal{F}g(\zeta)$, where $g(x) = e^{i(\zeta - \xi) \cdot x} f(x)$. \circ

Proof of theorem 11.3. Repeated application of equation (111) gives

$$(v \cdot \nabla)^m \mathcal{F}f(\xi) = \mathcal{F}(\mu_v^m f)(\xi), \quad (112)$$

where μ_v^m is the multiplication operator defined by $\mu_v^m f(x) = (v \cdot x)^m f(x)$. Notice that $v \cdot \nabla$ is a derivative in the direction v , and it makes sense to take the m th iterated derivative.

¹⁶Another exercise belongs here. See the appendix.

Take any $\rho \in \mathbb{C}^n$. We have $e^{\rho \cdot x} = \sum_{k \in \mathbb{N}} \frac{1}{k!} (\rho \cdot x)^k$. It is clear that each partial sum is dominated by $\sum_{k \in \mathbb{N}} \frac{1}{k!} |\rho \cdot x|^k = e^{|\rho \cdot x|}$. This majorant is uniformly bounded when $x \in K$ and $|\rho|$ is uniformly bounded.

Therefore, by the dominated convergence theorem and exercise 109,

$$\mathcal{F}f(\xi) = \sum_{k=0}^{\infty} \frac{1}{k!} \mathcal{F}(\mu_{i(\xi_0-\xi)}^k f)(\xi_0). \quad (113)$$

Applying (112) to each term gives

$$\mathcal{F}f(\xi) = \sum_{k=0}^{\infty} \frac{1}{k!} ((\xi - \xi_0) \cdot \nabla)^k \mathcal{F}f(\xi_0). \quad (114)$$

This is precisely the Fourier series of $\mathcal{F}f$ about the point ξ_0 evaluated at ξ . We have shown that this series converges to $\mathcal{F}f(\xi)$ as desired. \square

Exercise 110. We proved above that the Taylor series of $\mathcal{F}f$ at ξ_0 converges. What can you deduce about the radius of convergence? \circ

For the fun of it, let us see a couple of examples of real analytic functions.

Exercise 111. Calculate the Fourier transform of the characteristic function of the cube $[-1, 1]^3 \subset \mathbb{R}^3$. \circ

Exercise 112. Calculate the Fourier transform of the characteristic function of the unit ball $B \subset \mathbb{R}^3$. \circ

Exercise 113. Do you have any questions or comments regarding section 11? \circ

12 Partial data

In this section we will give our fifth injectivity proof based on Fourier analysis on \mathbb{R}^n . This method will also give a partial data result.

12.1 Various kinds of limitations

In real life measurement situations there are various kinds of limitations to the measurements. Sometimes one can fire X-rays through an object in any position and direction, but not always. One might need to avoid hitting something sensitive, or the geometry of the measurement situation restricts the available directions. In general, one might ask how large a set of lines is needed so that a function is uniquely determined by its integrals over them.

We have mostly studied X-ray tomography with full data so far. We had one result with partial data so far, namely Helgason's support theorem (theorem 5.4), which concerns tomography around a convex obstacle. In this section we will study a particular partial data scenario, where the set of directions of X-rays is not the whole sphere. This is called limited angle tomography.

12.2 Full data with the Fourier transform

Before embarking on the study of partial data, let us first solve the simpler full data problem with these tools. Consider a function $f \in C_c(\mathbb{R}^n)$ and define the X-ray transform as

$$\mathcal{I}f(x, v) = \int_{\mathbb{R}} f(x + tv) dt. \quad (115)$$

We will restrict the parameters $(x, v) \in \mathbb{R}^n \times S^{n-1}$ so that $x \cdot v = 0$. In other words, $x \in v^\perp$, where v^\perp denotes the subspace orthogonal to v . Since $\mathcal{I}f(x + sv, v) = \mathcal{I}f(x, v)$ for any $s \in \mathbb{R}$, this restriction does not reduce our data. In other words, $\mathcal{I}f(x, v)$ for all parameters $(x, v) \in \mathbb{R}^n \times S^{n-1}$ is uniquely determined by the restriction to $v \in S^{n-1}$ and $x \in v^\perp$.

Fix any $v \in S^{n-1}$ and consider the function $\mathcal{I}f(\cdot, v)$ on the $(n-1)$ -dimensional space $v^\perp \subset \mathbb{R}^n$. We can calculate the Fourier transform of $\mathcal{I}f(\cdot, v)$ on this space v^\perp . This function is continuous and compactly supported. For $\xi \in v^\perp$, we denote

$$(\mathcal{F}_{v^\perp} \mathcal{I}f(\cdot, v))(\xi) = \int_{v^\perp} e^{-i\xi \cdot x} \mathcal{I}f(x, v) d\mathcal{H}^{n-1}(x). \quad (116)$$

Exercise 114. Fix any $v \in S^{n-1}$. Suppose $\xi \in \mathbb{R}^n$ is orthogonal to v . Show that $(\mathcal{F}_{v^\perp} \mathcal{I}f(\cdot, v))(\xi) = \mathcal{F}f(\xi)$. ○

Theorem 12.1. *If $f \in C_c(\mathbb{R}^n)$ integrates to zero over all lines, then $f = 0$.*

Proof. Take any $\xi \in \mathbb{R}^n$ and choose $v \in S^{n-1}$ so that $v \cdot \xi = 0$. Since $\mathcal{I}f(\cdot, v) = 0$, exercise 114 gives $\mathcal{F}f(\xi) = 0$. Therefore $\mathcal{F}f = 0$, and by injectivity of the Fourier transform also $f = 0$. □

12.3 Limited angle tomography

Now we turn to our partial data problem. Let $D \subset S^{n-1}$ be the set of allowed directions. The question is whether $f \in C_c(\mathbb{R}^n)$ is uniquely determined by $\mathcal{I}f(x, v)$ for all $x \in \mathbb{R}^n$ and $v \in D$.

If $D = S^{n-1}$, then the result is stated in theorem 12.1 — and our other injectivity theorems. However, if D is finite, it turns out that there is an infinite dimensional subspace of functions $f \in C_c(\mathbb{R}^n)$ for which this data $\mathcal{I}f|_{\mathbb{R}^n \times D} = 0$. The simplest case is not hard to see.

Exercise 115. Suppose $D = \{v\}$ is a singleton. Show that there is a function $f \in C_c(\mathbb{R}^n) \setminus 0$ for which $\mathcal{I}f(\cdot, v) = 0$. \bigcirc

The general case for any finite set follows from a convolution argument:
¹⁷

Let us denote

$$D^\perp = \{\xi \in \mathbb{R}^n; \xi \cdot v = 0 \text{ for some } v \in D\}. \quad (117)$$

Notice that D^\perp is not the orthogonal complement of the linear space spanned by D . Instead, it is the union of the orthogonal complements of the elements in D .

Now we are ready to prove our theorem.

Theorem 12.2. *Let $f \in C_c(\mathbb{R}^n)$. Suppose $D \subset S^{n-1}$ is such that $D^\perp \subset \mathbb{R}^n$ contains an interior point. If $\mathcal{I}f(x, v) = 0$ for all $x \in \mathbb{R}^n$ and $v \in D$, then $f = 0$.*

Proof. Take any $\xi \in D^\perp$. Then there is $v \in D$ so that $v \cdot \xi = 0$. Since $\mathcal{I}f(\cdot, v) = 0$, exercise 114 gives $\mathcal{F}f(\xi) = 0$. Therefore the Fourier transform $\mathcal{F}f$ vanishes in D^\perp .

The continuous function f is compactly supported, so by theorem 11.3 $\mathcal{F}f$ is analytic. By assumption D^\perp contains a non-empty open set, and $\mathcal{F}f$ vanishes in it. Now exercise 106 implies that $\mathcal{F}f$ has to vanish identically due to analyticity. Since the Fourier transform is injective as mentioned in theorem 11.1, we conclude that the function f vanishes identically. \square

A new question arises: How much is needed about the set D of admissible directions to ensure that D^\perp contains an interior point? We will look at a couple of examples.

First, it is clear that D needs to be uncountable. If D is countable, then D^\perp is a union of countably many hyperplanes, and such a union cannot have interior points. We a simple approximation argument one can replace D with \bar{D} , so we may in fact assume that D is closed if we want to.

If D contains an interior point, so does D^\perp (exercise). In fact, much less is needed. For example, if D contains all the unit vectors $(v_1, v') \in \mathbb{R} \times \mathbb{R}^{n-1}$ for which $|v_1| = |v'|$, then D^\perp contains interior points if $n \geq 2$. This set D is a light cone, and this is the set needed for X-ray tomography in Minkowski spaces.

¹⁷Another exercise belongs here. See the appendix.

Exercise 116. Show that if $v \in S^{n-1}$ is an interior point of D , then any non-zero $\xi \in \mathbb{R}^n$ orthogonal to v is an interior point of D^\perp . \bigcirc

12.4 On stability and singularities

We found out above that D^\perp having an interior point is sufficient for injectivity. However, it is insufficient for stability. No matter which Sobolev norms one chooses for functions on \mathbb{R}^n and the relevant subset of Γ , there is no continuous left inverse for the partial data X-ray transform. Stable inversion is possible if $D = S^{n-1}$, and even if $D^\perp = \mathbb{R}^n$.

The reason for instability is that some kinds of singularities are undetected. Using microlocal analysis one can have very fine control over singularities, and it is possible to ask whether a distribution is smooth or singular at a given point in a given direction. To be able to detect a singularity at a point $x \in \mathbb{R}^n$ in direction $v \in S^{n-1}$, the data must contain a line through x in a direction orthogonal to v . A precise formulation of results of this kind of result is way beyond our reach here.

In our limited angle tomography situation the condition mentioned above amounts to $D^\perp = \mathbb{R}^n$. In this case one can reconstruct the Fourier transform everywhere directly, and invert the Fourier transform to obtain the original function. This is stable. On the other hand, if D^\perp contains interior points but is not the whole space, then the data needs to be analytically continued (by virtue of exercise 106), and analytic continuation is unstable without strong a priori estimates.

Exercise 117. Let $\Gamma' \subset \Gamma$ be a set of unoriented lines in \mathbb{R}^n . Suppose Γ' satisfies the stability condition mentioned above: For every $x \in \mathbb{R}^n$ and $v \in S^{n-1}$ there is a line $\gamma \in \Gamma'$ going through x in a direction orthogonal to v . Show that if $n = 2$, then $\Gamma' = \Gamma$, but if $n \geq 3$, this is not necessarily the case. \bigcirc

12.5 Local reconstruction

Another interesting question is whether a function can be reconstructed at a point from integrals of lines near that point. More precisely, let $x \in \mathbb{R}^n$ and let $U \ni x$ be a neighborhood (a region of interest). Does the knowledge of $\mathcal{I}f(\gamma)$ for all γ that meet U determine $f|_U$ for some class of functions f ?

It turns out that this is not possible. However, this data is enough to detect the singularities of $f|_U$. That is, one can locally reconstruct jumps and other singularities accurately, but not a smooth function. Microlocal reconstruction is possible, local is not. In many practical applications it is

indeed important to find the singularities of the unknown to identify sharp features, and it is not a big issue if the smooth part remains beyond reach.

Exercise 118. Let $f \in C_c(\mathbb{R}^n)$ and let $U \subset \mathbb{R}^n$ be an open set. Explain why the integrals of f over all lines that meet U determine $\mathcal{I}^*\mathcal{I}f|_U$. \bigcirc

However, local reconstruction is possible for the Radon transform in three dimensions. Consider a point $x \in \mathbb{R}^3$ and a neighborhood $U \ni x$. Then the integrals of $f \in C_c(\mathbb{R}^3)$ over all hyperplanes that meet U determine $f|_U$. As in exercise 118, this data determines $R^*Rf|_U$, where R stands for the Radon transform. It turns out that for some constant $c > 0$ we have $-c\Delta R^*Rf = f$, where Δ is the Laplace operator, so that f can be recovered from R^*Rf by differentiation.

In \mathbb{R}^n , the normal operator of the X-ray transform can be inverted by $(-\Delta)^{1/2}$ and that of the Radon transform by $(-\Delta)^{(n-1)/2}$. Local reconstruction is possible for the Radon transform in odd dimensions starting at three. Non-integer powers of the Laplace operator can be defined via Fourier transform.

Exercise 119. We saw above that local reconstruction for the Radon transform is possible in \mathbb{R}^3 . On the other hand, we saw in exercise 7 that injectivity of the Radon transform implies injectivity for the X-ray transform. Why does this not lead to local reconstruction for the X-ray transform in \mathbb{R}^3 ? \bigcirc

Exercise 120. Do you have any questions or comments regarding section 12? \bigcirc

13 Outlook

To conclude the course we review some directions of further study on the subject. The written descriptions are brief; they will be elaborated on and discussed in the lecture. The statements will not be made fully precise; the purpose is to give a flavor of what is known, not all details. Some results have been weakened for technical convenience.

13.1 Geodesic X-ray tomography

So far we have studied functions in Euclidean domains and integrated them over straight lines. But what if the domain is replaced by a Riemannian manifold with boundary and lines by geodesics? Most of our methods will be inapplicable on manifolds, but not all.

Most of our tools (Fourier series, Fourier transform, convolutions, polar coordinates) fail on a general manifold. If the manifold happens to be spherically symmetric, then our radial Fourier series approach works:

Theorem 13.1. *Equip the closed unit ball $\bar{B} \subset \mathbb{R}^n$ with a rotation symmetric Riemannian metric so that every maximal geodesic meets the boundary and therefore has finite length. Such manifolds are called non-trapping. On a non-trapping spherically symmetric Riemannian manifold a function is uniquely determined by its integrals over all geodesics.*

Spherical symmetry is a strong requirement. Our proof idea with the sphere bundle and the Pestov identity works in more generality:

Theorem 13.2. *A compact manifold with boundary is called simple if any two points can be joined with a unique geodesic and the geodesic depends smoothly on its endpoints. On a simple Riemannian manifold a function is uniquely determined by its integrals over all geodesics.*

The second theorem does not contain the first one; there are non-simple but non-trapping rotation symmetric manifolds.

13.2 Tensor tomography

A scalar function can be replaced with a tensor field of any rank. So far we have studied only rank zero (scalar fields) and rank one (vector fields). To go further, one must first understand what a tensor field is in the first place, and then figure out how to integrate them along lines (or geodesics). As in the case of vector fields, there is non-uniqueness for any non-zero rank. The goal is then to characterize this non-uniqueness.

A symmetric rank m tensor field on \mathbb{R}^n is a function $f: \mathbb{R}^n \times \mathbb{R}^{nm} \rightarrow \mathbb{R}$ so that $f(x; v_1, \dots, v_m)$ is smooth in x , linear in each v_i and invariant under changes of any two v_i and v_j . The integral of such a tensor field over a line $\gamma: \mathbb{R} \rightarrow \mathbb{R}^n$ is

$$\int_{\gamma} f = \int_{\mathbb{R}} f(\gamma(t); \dot{\gamma}(t), \dots, \dot{\gamma}(t)) dt. \quad (118)$$

That is, the velocity $\dot{\gamma}$ is plugged into each of the m slots.

A scalar function is a tensor field of rank zero, and a vector field of rank one.

If $m = 1$, there is only one slot for v , and this is the integral of a vector field along a line as defined earlier. If $m = 0$, there are no slots at all and the function only depends on x . The resulting integral is the usual integral of a scalar function we have studied in this course.

Theorem 13.3. *On a two-dimensional simple manifold a tensor field f of order m integrates to zero over all geodesics if and only if there is a tensor field h of order $m - 1$ which vanishes at the boundary and satisfies $f = d^s h$, where d^s is a symmetrized covariant derivative.*

13.3 Other classes of curves

So far we have only integrated functions (and vector fields) over straight lines. We also mentioned Riemannian geodesics, but there several other options as well. Of course, this can be seen as the mathematical art of (over)generalization, but a great number of different geometrical situations turn out to be physically relevant.

As a broad term, inverse problems in integral geometry ask to recover an object from its integrals over a collection of subsets of the space.

Some classes of curves to consider:

- Lines in \mathbb{R}^n .
- Circles in \mathbb{R}^2 .
- Geodesics on a Riemannian manifold.
- Magnetic geodesics on a manifold or \mathbb{R}^n .
- Geodesics on a Finsler manifold.
- Integral curves of a dynamical system.
- Curves which reflect or split in some way.

13.4 Linearization of travel time

For one specific example of applications, we can consider travel time tomography in seismology. The problem can be recast as a geometrical one, once the Earth is treated as a geometrical object. The linearized problem has non-uniqueness, but it corresponds to the non-uniqueness inherent to the geometrical problem. This has physical repercussions.

Theorem 13.4. *Let M be a manifold and g_s a family of Riemannian metrics on it, depending on a parameter $s \in \mathbb{R}$. Consider two points $x, y \in M$. Let γ_s be a geodesic with respect to the metric g_s joining these two points. Suppose γ_s depends smoothly on s and denote the length of γ_s by $\ell_s(\gamma_s)$. Denote by $f_s = \partial_s g_s$ the second order tensor field obtained by differentiating the metric with respect to the parameter. Then*

$$\frac{d}{ds} \ell_s(\gamma_s) = \frac{1}{2} \int_{\gamma_s} f_s. \quad (119)$$

Linearized travel time tomography is tensor tomography of rank two.

If one studies conformal variations of a metric, then each g_s is a conformal multiple of g_0 . In that case f_s is also a conformal multiple of g_s , and the tensor tomography problem reduces to a scalar tomography problem for the conformal factor.

13.5 Overview

It is now time to look back and see what, if anything, we have accomplished during the course. The course started with the physical problem of X-ray tomography and its mathematical formulation. We then proved a uniqueness result in five different ways. These results are collected in theorems 3.4, 5.2, 6.2, 9.4, and 12.1. However, these do not exhaust all known inversion methods. For example, we excluded Radon's original inversion method from 1917.

Bonus exercise 6. Find Radon's inversion formula and explain how and why it works. \bigcirc

In addition, we proved a number of related results. We proved an injectivity result for the X-ray transform on tori (theorem 3.3) and solenoidal injectivity for vector field tomography (theorem 10.1). We also gave a proof of two partial data results, namely Helgason's support theorem (theorem 5.4) and a result in limited angle tomography (theorem 12.2).

Exercise 121. We gave five uniqueness proofs for the X-ray transform. Give a quick overview of the mathematical tools needed for each of the five proofs. Give five lists, one for each proof. Which proofs did you find most accessible and simple, and which ones hardest to follow? \bigcirc

13.6 Further reading

The following selection of books, lecture notes and review articles is by no means complete. The sources listed here are available online.

- Gunther Uhlmann, "Inverse problems: seeing the unseen", Bulletin of Mathematical Sciences, volume 4, issue 2, pp. 209–279, 2014: An overview of two important inverse problems, namely travel time tomography and Calderón's problem. Both are related to X-ray transforms of some kind.
<https://link.springer.com/article/10.1007/s13373-014-0051-9>
- Sigurður (Sigurdur) Helgason, "The Radon Transform", second edition, 1999 and "Integral Geometry and Radon Transforms", first edition,

2011: The books are freely available on the author's homepage. They give a very thorough treatment of the theory of X-ray transforms and particularly Radon transforms.

<http://www-math.mit.edu/~helgason/publications.html>


- Vladimir Sharafutdinov, “Ray Transform on Riemannian Manifolds. Eight Lectures on Integral Geometry”, 1999: Lecture notes on X-ray tomography on Riemannian manifolds.
<http://www.math.nsc.ru/~sharafutdinov/publ.html>
- Gabriel Paternain, Mikko Salo, Gunther Uhlmann, “Tensor tomography: progress and challenges”, Chinese Ann. Math. Ser. B 35, no. 3, 399–428, 2014: A review article on X-ray tomography on manifolds, with focus on tensor fields.
<https://arxiv.org/abs/1303.6114>
- Will Merry, Gabriel Paternain, “Inverse Problems in Geometry and Dynamics”, 2011: Lecture notes on geodesic X-ray tomography with a dynamical focus.
[https://www.dpmms.cam.ac.uk/~gpp24/ipgd\(3\).pdf](https://www.dpmms.cam.ac.uk/~gpp24/ipgd(3).pdf)
- Gunther Uhlmann, Hanming Zhou, “Journey to the Center of the Earth”, 2016: A review article of travel time tomography, which is closely related to tensor tomography.
<https://arxiv.org/abs/1604.00630>

Exercise 122. Do you have any questions or comments regarding section 13?



13.7 Feedback

This course is somewhat unusual, compared to most courses in mathematics. To make the course more suitable for students in the future, the last exercises concern the course itself.

Exercise 123. This course introduced a physical problem and introduced a number of mathematical tools related to that application. Should there have been more focus on physics — more applications, more details, deeper explanations, or something else? Was the balance between physics and mathematics good for your interests? (Also, what is your major?) 

Exercise 124. The course was designed to be broad but shallow. We discussed a number of different mathematical tools and ideas related to X-ray tomography, but we did not go very deep into any of them. This was done so as to

give you a broad overview of the topic and an example of how various different tools in analysis can be used to tackle the same applied problem. On the other hand, we could have developed a theory with optimal regularity (not restricting to continuous functions all the time), stability estimates, range characterizations, various function spaces and the like. Should the course have been deeper as opposed to broad? Would you have preferred to see one theory developed in detail rather than several independent ideas? It would also be possible to give a broad introductory course like this one and then a more in-depth follow-up course. \bigcirc

Exercise 125. Answer the following questions:

- (a) How much time did you spend on the course?
- (b) What were the worst things about the course?
- (c) Would you be interested in a course in the geometry of geodesics and geodesic X-ray tomography on Riemannian manifolds?
- (d) Do you have any feedback in mind that was not covered by other questions?

Thank you for your feedback! \bigcirc

A More exercises

These exercises give points normally. However, the grade limits do not change because of these exercises. These exercises belong in the indicated sections, but are listed here for the sake of consistent numbering throughout different versions. There have been slight changes to existing exercises, but any version is valid for the purpose of completing the course.

A.1 Section 1

Exercise 126. Show that the X-ray transform $\mathcal{I}: C_B \rightarrow C_b$ as defined above is not injective if $n = 1$. \bigcirc

A.2 Section 2

Exercise 127. When defining the X-ray transform on \mathbb{T}^1 and \mathbb{T}^n , we made use of the exponential functions $e^{ik \cdot x}$. Show that if $k \in \mathbb{Z}^n$ and $x \in \mathbb{T}^n = \mathbb{R}^n / 2\pi\mathbb{Z}^n$, the value of $e^{ik \cdot x}$ does not depend on the representative of x in \mathbb{R}^n . This means that the exponential function $e^{ik \cdot x}$ is indeed well defined. Is the exponent $k \cdot x$ well defined, too? \bigcirc

A.3 Section 3

Exercise 128. Show that if $v \in 2\pi\mathbb{Z}^n \setminus 0$, then \mathcal{I}_v is not injective. Use lemma 3.2 or take a function $f \in C^\infty(\mathbb{T}^n)$ and consider the function $v \cdot \nabla f(x)$. ○

Exercise 129. Let $v \in 2\pi\mathbb{Z}^n \setminus 0$ and $m \in \mathbb{Z} \setminus 0$. Show that $\mathcal{I}_{mv} = \mathcal{I}_v$. ○

Exercise 130. All of the results in this section are valid for $n = 1$ apart from exercise 27. When $n = 1$, one can only find an orthogonal $w \in \mathbb{Z}$ for $k = 0$. By exercise 129 all one can measure about $f \in C(\mathbb{T}^1)$ is $\mathcal{I}_1 f$. What does this mean for recovering the Fourier coefficients $\mathcal{F}f(k)$? ○

A.4 Section 4

Exercise 131. Show that if $f \in C^1(\bar{D}^*)$, then

$$\partial_\theta f(r, \theta) = \left. \frac{d}{d\varphi} (R_\varphi f)(r, \theta) \right|_{\varphi=0}. \quad (120)$$

In other words, prove equation (36). (Notice that while R_φ maps $C^1(\bar{D}^*) \rightarrow C^1(\bar{D}^*)$, the derivative ∂_θ only maps $C^1(\bar{D}^*) \rightarrow C^0(\bar{D}^*)$.) ○

A.5 Section 5

Exercise 132. Justify the formulas for T_0 and T_1 and recursion relation for T_k using the property that $\cos(kx) = T_k(\cos(x))$ for all $k \in \mathbb{N}$. ○

A.6 Section 6

Exercise 133. Using the definition, show carefully that $(A^*)^* = A$. ○

Exercise 134. Show that $B = P^{-1}A^TQ$. Therefore the normal operator is $BA = P^{-1}A^TQA$. ○

A.7 Section 7

A.8 Section 8

A.9 Section 9

Exercise 135. Use the Santaló formula to rewrite $\|Xw\|_{L^2(S\Omega)}^2$ when $w \in C_c^\infty(S\Omega)$. What is the integral you end up calculating over each geodesic? Two terms of this kind appear on the right-hand side of the Pestov identity.

○

A.10 Section 10

Exercise 136. Consider the one-dimensional set $(0, 1)$ or its closure. What does it mean if the X-ray transform is solenoidally injective on this space? Prove this solenoidal injectivity.

○

A.11 Section 11

Exercise 137. Define the function $f: \mathbb{R} \rightarrow \mathbb{R}$ by

$$f(x) = \begin{cases} 0, & x \leq 0 \\ \exp(-1/x), & x > 0. \end{cases} \quad (121)$$

Consider it known that $f \in C^\infty(\mathbb{R})$.

Explain and justify (or prove): For any $x \in \mathbb{R}$ the Taylor series of f at x converges in some open neighborhood of x . However, f is not real-analytic.

○

A.12 Section 12

Exercise 138. Fix any $v \in S^{n-1}$. Let $f, g \in C_c(\mathbb{R}^n)$. Show that $\mathcal{I}(f * g)(\cdot, v) = (\mathcal{I}f(\cdot, v)) * g$.

(It then follows that if $\mathcal{I}f(\cdot, v_1) = 0$ and $\mathcal{I}g(\cdot, v_2) = 0$, then $\mathcal{I}(f * g)(\cdot, v) = 0$ for both $v \in \{v_1, v_2\}$. The only thing left to worry about is that the convolution of two non-trivial compactly supported functions cannot vanish identically.)

○

A.13 Section 13

—

A.14 More exercises for version 3

These are the exercises written after version 2.

Section 2:

Exercise 139. One aspect of the basis functions used in the Fourier series is that they are eigenfunctions of the Laplace operator. What is the eigenvalue of the function $x \mapsto e^{ik \cdot x}$? ○

Bonus exercise 7. Show that the number 7 cannot be written as the sum of three integer squares. Recall the statement of Lagrange's four-square theorem.

Show that the set of eigenvalues of the Laplace operator Δ on the torus \mathbb{T}^n is \mathbb{N} if and only if $n \geq 4$. ○

Section 5:

Bonus exercise 8. One can also go the other way around. Assume that theorem 5.2 is true (this has been proved with other methods in the previous section). Use the tools developed in this and the previous section to prove that the Abel transforms \mathcal{A}_k are injective. ○

Section 8:

Exercise 140. Consider linear operators ($n \times n$ matrices, for example) A , B , and C . Show that $[A, BC] = [A, B]C + B[A, C]$ and $[AB, C] = A[B, C] + [A, C]B$. ○

Exercise 141. Compute $[XVV, VX_\perp]$ and $[V^2, X_\perp^2]$. ○

Section 9:

Exercise 142. An operator $A: C^\infty(S\Omega) \rightarrow C^\infty(S\Omega)$ is called a derivation if it is linear and satisfies $A(gh) = gAh + hAg$ and $A1 = 0$, where 1 stands for the constant function.

Show that the commutator of two derivations is a derivation. Explain why X , V , and X_\perp are derivations. ○

Exercise 143. Show that $VX = \frac{1}{4}(X+V)^2 - \frac{1}{4}(X-V)^2 + \frac{1}{2}X_\perp$ and $XV = \frac{1}{4}(X+V)^2 - \frac{1}{4}(X-V)^2 - \frac{1}{2}X_\perp$.

Since the three vector fields X , V , and X_\perp are orthonormal (with respect to the Sasaki metric), so are $\frac{1}{\sqrt{2}}(X+V)$, $\frac{1}{\sqrt{2}}(X-V)$, and X_\perp . Therefore our two operators look locally like the operators $\frac{1}{2}(\partial_x^2 - \partial_y^2 \pm \partial_z)$ in \mathbb{R}^3 . ○

(Replacement of old paragraph:) One might expect that the first order term is not that important, but it turns out to be very important. In other words, the order of the operators X and V is crucial. We will show that assuming zero boundary values the PDE $VXu = 0$ has unique solutions, but $XVu = 0$ never does. The uniqueness result will be proven in the next section. The non-uniqueness result is easier and is left as exercise 144.

(Exercise 91 here.)

Exercise 144. Suppose $u \in C^\infty(S\bar{\Omega})$ with $u|_{\partial(S\Omega)} = 0$. Show that $XVu = 0$ in $S\Omega$ if and only if there is $f \in C^\infty(\bar{\Omega})$ with zero boundary values so that $u = \pi^* f$. \circ

A.15 Final remarks

This is the last version of the notes to be used for this course. There are a total of 143 exercises, so getting the required 253 points for the grade 5/5 (or 125 points for 1/5) is easier than it first appeared — especially with the presentation or bonus exercises.

Some time after the course I will write a final version of these notes for future use. It will include bigger changes: new sections, reordering, newer exercises in their correct places. I will not be able to finish the final version before the course is over, I have seen your solutions to the exercises and read all your feedback, including questions comments section by section and overall feedback. Feedback has been very useful so far. Thank you all!