ON THE BROKEN RAY TRANSFORM

JOONAS ILMAVIRTA

To be presented, with the permission of the Faculty of Mathematics and Science of the University of Jyväskylä, for public criticism in Auditorium Agora Alfa on August 29th, 2014, at 12 o’clock noon.
Abstract

The classical problem of X-ray tomography asks whether one can reconstruct a function from its integrals over all lines. This problem has a wide range of applications and has been studied intensively. As an example of the applications we mention computerized tomography, which is an important tool in medical imaging. When we replace the lines with broken rays which reflect on some subset of the boundary of the domain of interest, the problem becomes much more difficult. The reflecting part of the boundary is inaccessible to measurements, so the problem at hand resembles X-ray tomography with partial data. This type of tomography is called broken ray tomography in this thesis.

The integral transform corresponding to broken ray tomography is called the broken ray transform, which can be thought of as a generalization of the X-ray transform. The fundamental question is whether this transform is injective. We employ four different methods to approach this question, and each of them gives interesting results.

Direct calculation can be used in a ball, where the geometry is particularly simple. If the reflecting part of the boundary is (piecewise) flat, a reflection argument can be used to reduce the problem to the usual X-ray transform. In some geometries one can use broken rays near the boundary to determine the values of the unknown function at the reflector, and even construct its Taylor series. One can also use energy estimates – which in this context are known as Pestov identities – to show injectivity in the presence of one convex reflecting obstacle. Many of these methods work also on Riemannian manifolds.

We also discuss the periodic broken ray transform, where the integrals are taken over periodic broken rays. The broken ray transform and its periodic version have applications in other inverse problems, including Calderón’s problem and problems related to spectral geometry.

A simplified introduction to inverse problems to the nonmathematical reader is given in appendices A (English), B (Finnish), and C (Latin).
Tiivistelmä


Murtosasadetomografiasta vastaava integraalimuunnos on murtosädemuunnos, jota voi ajatella Röntgen-muunnoksen yleistyksenä. Peruskysymys on, onko tämä muunnos injektiivinen. Käytämme neljää eri tapaa lähestyä tätä ongelmaa, ja niistä jokainen antaa mielenkiintoisia tuloksia.


Tarkastelemme myös jaksollista murtosasademuunnosta, jossa integrointi tehdään kaikkien jaksollisten murtosäteiden yli. Murtosasademuunoksella ja sen jaksollisella versiolla on sovelluksia muissa inversio-ongelmissa kuten Calderónin ongelmassa ja spektraaligeometriasta liittyvissä kysymyksissä.

Yksinkertaistettu johdanto inversio-ongelmii matematiikkaa tuntemattomalle lukijalle on liitteissä A (englanniksi), B (suomeksi) ja C (latinaksi).
Acknowledgements

I wish to express my greatest gratitude to my advisor, Professor Mikko Salo, for leading me to an interesting path and for giving support and advice on my journey along it. I have been able to complete my studies in a friendly and supportive atmosphere, for which I thank all of the staff and students at the Department of Mathematics and Statistics. Numerous colleagues have helped me with the problems I encountered, and I hope to have been able to offer them something in return.

I am most thankful to Professor Vladimir Sharafutdinov for making invaluable remarks on an early version of this thesis and for recommending several relevant articles that I was previously unaware of.

I would never have finished this work without life outside academia. Support from family and friends has made all of this possible. The people who have been there for me know who they are; I refrain from listing names as it would only lead to me having to decide whom not to include in the list.

My research was financially supported by the Academy of Finland, especially the Centre of Excellence in Inverse Problems Research, for which I am most grateful.

Jyväskylä, March 2014
Joonas Ilmavirta
List of included articles

This thesis is based on the work contained within the following publications:

[I] Broken ray tomography in the disc
    J. Ilmavirta,

[II] A reflection approach to the broken ray transform
    J. Ilmavirta,
    to appear in Mathematica Scandinavica.

[III] Boundary reconstruction for the broken ray transform
    J. Ilmavirta,

[IV] On Radon transforms on tori
    J. Ilmavirta,
    preprint.

[V] Broken ray transform on a Riemann surface with a convex obstacle
    J. Ilmavirta, M. Salo,
    preprint.

The author has participated actively in development of the joint paper [V]. In particular, the author is responsible for the regularity lemma and calculating the linearization.

The Latin introduction (appendix C) has been published in Melissa (number 180, June 2014), a magazine published in Latin.
Appendices A, B and C contain an introduction to inverse problems for the non-mathematical audience. The same content is presented in three languages.
1. Introduction

1.1. Inverse problems. To understand an inverse problem, we first need to understand direct problems. Consider, for example, the following direct problems:

- Given a $C_0$ function (continuous with compact support) $f : \mathbb{R}^n \to \mathbb{R}$, find its integral over every line.
- Given a closed Riemannian manifold, find the spectrum of its Laplace-Beltrami operator.
- Given a compact Riemannian manifold with boundary, find the distance between any two boundary points.
- Given a $C^2$ function $\gamma : \mathbb{R}^n \supset \bar{\Omega} \to (0, \infty)$, find the boundary values $u|_{\partial \Omega}$ and $\gamma \partial_\nu u|_{\partial \Omega}$ for all solutions $u \in W^{1,2} (\Omega)$ of $\text{div} (\gamma \nabla u) = 0$.

It is obvious that the solution to these problems exists uniquely, and in some cases it is relatively easy to find an explicit solution.

When turned around, these become inverse problems:

- Given the integral of a $C_0$ function over every line, find the function. (The X-ray tomography problem.)
- Given the spectrum of the Laplace-Beltrami operator on a closed manifold, find the manifold. (Related to the spectral rigidity problem.)
- Given the distance between any two boundary points on a manifold with boundary, find the manifold. (The boundary rigidity problem.)
- Given the boundary values $u|_{\partial \Omega}$ and $\gamma \partial_\nu u|_{\partial \Omega}$ for all solutions $u \in W^{1,2} (\Omega)$ of $\text{div} (\gamma \nabla u) = 0$, find the function $\gamma \in C^2 (\bar{\Omega})$. (Calderón’s problem.)

These famous inverse problems have been studied intensively. Solving these problems is far more difficult than solving the corresponding direct problems. In particular, it is not clear that a solution should exist, let alone be unique. If the solution exists uniquely, one would like to have an explicit reconstruction method and know that it is stable.

1.2. Ray transforms. We are interested in inverse problems similar to the first one in the list above. The question whether one can recover a function in a bounded Euclidean domain from its integral over all lines has applications in medical imaging. This problem was first solved by Radon [28] and Cormack [4]. For applications and further developments, see for example the book by Natterer [26].

This question can be generalized in several directions. First, lines can be replaced with another family of curves. This includes geodesics on a Riemannian manifold. Second, the integrals can be taken over hyperplanes instead of lines (the Radon transform) or more generally over all $d$-planes for any integer $d$ strictly between one and the dimension (the $d$-plane Radon transform). This second generalization does not make sense on all manifolds, but it can be defined in the Euclidean space and flat tori; the $d$-plane Radon transform on tori is discussed in [IV]. Third, the scalar function may be replaced with a tensor field (see [29]).
The first generalization is the one most important to us. Namely, we may replace the lines (or geodesics) with broken rays which are piecewise linear (or geodesic) curves. There are many different types of breaking points to consider. If the breaking points are in the interior of the domain, one can consider geodesics that break in all possible directions [16] or only those that break in a fixed angle [17, 32, 2, 3]. Both of these cases have applications in imaging. Recently Zhao, Schotland and Markel [34] (see also [13]) have also considered the case where a line breaks into several lines, forming a star rather than a broken line. One may also let the broken rays reflect from the boundary of the domain according to the usual law of geometrical optics. Versions of this problem have been considered by Mukhometov [20, 21, 22, 23, 24, 25] and Eskin [8], and this case is the main topic of this thesis.

1.3. The broken ray transform. Before going further, we need to introduce some notation. Let $M$ be a compact Riemannian manifold with boundary, and let its boundary be divided into disjoint sets $E$ and $R$. A broken ray is a curve on $M$ which is geodesic in the interior of $M$, has both endpoints on $E$ and reflects on other boundary points according to the usual reflection law. That is, the incoming direction $v_i$ and the outgoing direction $v_o$ are related via $v_o = v_i - 2 \langle \nu, v_i \rangle \nu$, where $\nu$ is the outer unit normal at the reflection point. In dimension two this amounts to demanding that the angle of incidence equals the angle of reflection. Note that a broken ray must reflect when it hits $R$, but on $E$ it can either reflect or stop. Allowing reflections on $E$ will be convenient.

We refer to the set $E$ as the set of tomography, since measurements can only be done on it. The set $R$ is a reflecting part of the boundary which is inaccessible to measurements.

There are two main types of geometrical situations that we are interested in. The boundary $\partial M$ can be connected, when $E$ and $R$ necessarily meet at some boundary point. If the boundary is not connected, $E$ and $R$ may consist of connected components of the boundary, making them well separated.

For the sake of simplicity, we assume all broken rays to have finite length. We allow reflections to be tangential (in which case the direction is not changed at all), although it will pose a problem that the reflection fails to be smooth at tangential directions.

**Definition 1.** Let $M$ be a manifold as described above and let $\Gamma$ be the set of all broken rays on it. Denote by $\mathbb{R}^\Gamma$ the set of all mappings $\Gamma \to \mathbb{R}$. We define the broken ray transform as a mapping $\mathcal{G} : C(M) \to \mathbb{R}^\Gamma$ by letting

$$\mathcal{G}f(\gamma) = \int_0^L f(\gamma(t))dt =: \int_\gamma f ds$$

for a broken ray $\gamma : [0, L] \to M$. 
The transform can be also be defined for many noncontinuous functions in the same way. One can also introduce a weight or attenuation to the broken ray transform (see [III, Section 2.1]), but for the sake of simplicity we omit them here.

**Question 2.** Regarding the broken ray transform defined in definition 7 above, we ask the following questions:

- **Uniqueness:** Is a function uniquely determined by its broken ray transform? In other words, is the broken ray transform injective?
- **Reconstruction:** Is there an explicit numerical algorithm or analytic formula to recover a function from its broken ray transform?
- **Stability:** If two functions have almost the same broken ray transform, are the functions almost the same?
- **Regularity:** How do the answers to these questions depend on the regularity of the unknown function and the domain?

The goal of this thesis is to answer these questions. A complete answer is by far beyond reach, but partial answers have been obtained. We will give several examples of domains where the broken ray transform is injective. We use minimal regularity assumptions for the methods of proof used, but there is no indication that the obtained regularity is optimal. A partial stability result is shown in one case (see section 2.4), but otherwise stability and reconstruction are only inherited by reducing the broken ray transform to the usual X-ray transform.

We also give counterexamples to uniqueness. But since none of the counterexamples is particularly close to the positive results, we cannot claim that any of the theorems would be sharp.

### 1.4. The periodic broken ray transform

As the name suggests, a periodic broken ray is a periodic curve on a manifold \( M \) with boundary which reflects at the boundary according to the aforementioned reflection law. Note that in this case all of the boundary \( \partial M \) is reflective, since periodic broken rays have no endpoints.

The periodic broken ray transform is then defined analogously with the broken ray transform. We pose question 2 also for the periodic broken ray transform. The answers given here are limited to some examples and counterexamples.

### 2. How to approach the broken ray transform

We present four methods to approach the broken ray transform described in section 1.3:

1. Compute the broken ray transform explicitly.
2. Reduce to X-ray transform by reflection.
3. Reduce boundary determination for the broken ray transform to injectivity of the X-ray transform on boundary via broken rays staying near the boundary.
(4) Reduce injectivity of the broken ray transform to the unique solvability of a PDE.

Details of these methods are given in the correspondingly numbered subsections 2.1–2.4, but we give a short overview here. The proofs given in the subsections are not complete, but merely attempt to convey the main ideas. For detailed proofs, see the corresponding articles included in this thesis.

The first approach requires the domain to have very explicit geometry, at least for the reflecting part $R$ of the boundary. Such approach can be taken in a Euclidean disc, where a broken ray consists of rotated copies of a chord. This approach was taken in [1] and will be outlined in section 2.1.

The second approach is best illuminated with the example of a half plane: If $f : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ is a compactly supported continuous function, it can be reflected to a function $\tilde{f} : \mathbb{R}^2 \to \mathbb{R}$ by setting $\tilde{f}(x, y) = f(x, |y|)$. If we know the integral of $f$ over all broken rays that reflect on $\mathbb{R} \times \{0\}$, then we know the integral of $\tilde{f}$ over all lines in $\mathbb{R}^2$. Thus the broken ray transform is reduced to the X-ray transform, for which injectivity is known.

This approach has been taken in [I, 12, 11]. The reflection method works best when the reflecting part $R$ of $\partial \Omega$ is flat. More details of this method (including the role of flatness) are given in section 2.2.

The third approach is based on an intuitive geometrical observation: If $\partial \Omega$ is strictly convex, then a broken ray starting almost tangent to $\partial \Omega$ remains almost tangent to $\partial \Omega$, and if a sequence of broken rays becomes closer and closer to being tangent to the boundary, the limit curve is a geodesic on $\partial \Omega$. In fact, any boundary geodesic along which $\partial \Omega$ is strictly convex can be approached uniformly with broken rays.

Thus, if the unknown function is continuous, we may recover its integral over all boundary geodesics from the broken ray data. If we can invert the X-ray transform on the manifold $\partial \Omega \setminus E$, we can reconstruct the unknown function at $\partial \Omega$. A similar argument can be used to recover the normal derivatives at the boundary as well, but the corresponding X-ray transform on the boundary manifold picks a weight depending on curvature. A more detailed account of this method is given in section 2.3 which is based on [III].

The last approach is not as obvious as the three other ones. For an unknown function $f : \Omega \to \mathbb{R}$ we define another function $u^f$ on the sphere bundle $S\Omega$. Assuming that $f$ has a vanishing broken ray transform, this function $u^f$ is shown to satisfy a PDE with some boundary conditions on $E$ and $R$. An energy estimate (called Pestov identity in this context) can be used to show that this PDE has a unique solution and thus $u^f$ vanishes. The function $u^f$ is defined so that this implies that $f$ vanishes.

This kind of a PDE approach has been previously used for the X-ray transform (see for example [18, 27], the first of which is translated from [19]) and also for
the broken ray transform under very special geometrical assumptions by Eskin [8]. Mukhometov [20, 21, 22, 23, 24, 25] has also used this approach to study the broken ray transform. This approach was used in [V] to show injectivity of the broken ray transform on some Riemannian surfaces with one convex obstacle. This approach can also be used to obtain a stability estimate: if the broken ray transform of \( f \) is small, then \( f \) itself is small. All the necessary definitions and basic properties related to this method are given in section 2.4.

The main result in [V] is very close to that of [23]. A major motivation for [V] is providing a basis for analysis of the broken ray transform in the presence of several convex obstacles. Therefore regularity of the function \( u^f \) is studied in detail, to which end we also introduce Jacobi fields along broken rays.

It is important to note that the different approaches work in different geometrical settings. The reflection approach works best when the reflecting part \( R \) of \( \partial \Omega \) is flat; the Riemannian metric of the constructed auxiliary manifold will fail to be \( C^1 \) precisely when the second fundamental form does not vanish identically on \( R \). The boundary determination result presented in section 2.3 requires that \( R \) is strictly convex along sufficiently many boundary geodesics – preferably all of them. The PDE approach requires that \( R \) is strictly concave, and the reason is twofold: the billiard map is more regular in this case and an important term in the Pestov identity [3] has wrong sign otherwise.

We have thus ways to approach the broken ray transform if \( R \) is strictly convex, flat, or strictly concave. When different parts of \( R \) have types, the methods may be combined to some extent: the reflection argument can easily be combined with any other one. These reflector geometries do not, however, exhaust all cases. Saddle-like \( R \), for example, does not yield to any of these approaches.

2.1. Explicit calculation. The case when \( \Omega \) is a Euclidean disc is geometrically very convenient. We may take \( \Omega \) to be the unit disc centered at the origin. A broken ray in the disc consists of rotated copies of a chord. Also, we may use polar coordinates \((r, \vartheta)\) and write functions as Fourier series in the angular variable \( \vartheta \).

This section follows the article [I]. The important property of the disc used here is rotational symmetry in polar coordinates.

We begin by observing that the broken ray transform in an \( n \)-dimensional Euclidean ball \( B^n \) for \( n \geq 2 \) can be reduced to the case \( n = 2 \). If \( F \) is a two-dimensional subspace of \( \mathbb{R}^n \), \( n \geq 3 \), then any broken ray with one segment in \( F \) is completely contained in \( F \). If we know injectivity for \( n = 2 \), we know that the broken ray transform of \( f : B^n \to \mathbb{R} \) determines \( f \) in \( F \cap B^n \). And since this holds for any \( F \), we have injectivity for every \( n \geq 2 \). We will therefore only prove the following theorems in the case \( n = 2 \).

The main results obtained in [I] are the following:

**Theorem 3.** If \( f : B^n \to \mathbb{C} \) is continuous and \( E \) is a singleton, the broken ray transform of \( f \) uniquely determines the integral of \( f \) over any circle centered at the
origin with the singleton in the circle’s plane. In particular, if \( n = 2 \), the integral of \( f \) is determined over any circle centered at the origin.

**Theorem 4.** Suppose \( f : \mathbb{B}^n \to \mathbb{C} \) is uniformly quasianalytic in the angular variable in the sense of \([1, \text{Definition 14}]\) and its angular derivatives of all orders satisfy the Dini-Lipschitz condition. If \( E \) is open, \( f \) is uniquely determined by its broken ray transform.

A simple example of a function in two dimensions satisfying the regularity assumption of theorem 4 is

\[
 f(r, \vartheta) = \sum_{|k| \leq K} a_k(r)e^{ik\vartheta},
\]

where \( K \in \mathbb{N} \) and each \( a_k \) is Hölder continuous.

Both of these results are based on passing to the limit of infinitely many reflections. To explain this in more detail, we must fix some notation. We identify angles with points on the boundary. A broken ray \( \gamma \)

- contains \( n_\gamma \) line segments of length \( d_\gamma \) and distance \( z_\gamma \) from the origin,
- winds around the origin \( m_\gamma \) times,
- has initial and final points \( \iota_\gamma \) and \( \kappa_\gamma \), and
- each line segment opens at an angle \( \alpha_\gamma \) as viewed from the origin.

These seven parameters define \( \gamma \) uniquely, and although this parametrization is redundant, it is intuitive and useful. The winding number \( m_\gamma \) is defined by the relation

\[
 n_\gamma \alpha_\gamma = 2\pi m_\gamma + \kappa_\gamma - \iota_\gamma.
\]

The lengths can be expressed in terms of the other parameters: \( z_\gamma = \cos(\alpha_\gamma/2) \) and \( d_\gamma = 2|\sin(\alpha_\gamma/2)| \). The parameters are illustrated in figure \([1]\).

Suppose we have a sequence \( (\gamma_i) \) of broken rays such that \( z_{\gamma_i} \to z \in (0,1) \) and \( n_{\gamma_i} \to \infty \) as \( i \to \infty \). Broken rays with many reflections are almost spherically symmetric: the integral of \( f \) over \( \gamma_i \) tends to the integral of \( f \) against a spherically symmetric function. For theorem \([3]\) this observation is enough, since it turns the broken ray transform \( f : B^2 \to \mathbb{C} \) essentially to the Abel transform of the spherical mean function \( a_0(r) = \int_{|x|=r} f(x) \). Injectivity of the Abel transform finishes the proof.

Theorem \([4]\) requires more. First we write \( f \) as a Fourier series:

\[
 f(r, \vartheta) = \sum_{k \in \mathbb{Z}} e^{ik\vartheta}a_k(r).
\]

Under the assumption of Dini-Lipschitz-continuity this series converges uniformly and therefore the broken ray transform may be computed term by term.
Figure 1. An example of a broken ray $\gamma$ in the disc. Here $m_\gamma = 1$ and $n_\gamma = 3$. The parameters $\alpha_\gamma$, $z_\gamma$, $d_\gamma$, $\iota_\gamma$, and $\kappa_\gamma$ as well as the set of tomography $E$ are illustrated in the figure. The angles $\iota_\gamma$ and $\kappa_\gamma$ are identified with the corresponding points on the boundary.

Another important observation is based on rotational symmetry: If $E$ is open, a slightly rotated broken ray is still a broken ray with endpoints in the same set $E$. But since rotating a broken ray clockwise has precisely the same effect on the broken ray transform than rotating the function counterclockwise, small rotations retain vanishing broken ray transform for a slightly smaller set of tomography. Passing to the limit of zero rotation we find that the angular derivative of a function with vanishing broken ray transform has vanishing broken ray transform. The argument can be iterated to show that angular derivatives of all orders have vanishing broken ray transform.

In theorem 3 we only recovered the Fourier component $a_0$ by using the Abel transform (denoted here by $A_0$). In attempt to recover all components and thus the whole function, we end up with generalized Abel transforms $A_k$. These transforms are also injective on continuous functions and arise naturally in the Radon transform: if $f(r, \vartheta) = e^{ik\vartheta}a(r)$, then the integral of $f$ over the line $\{x; x_1 \cos \varphi + x_2 \sin \varphi = \rho\}$ is $e^{ik\varphi}A_k|_{k\|a(\rho)}$ if $\rho > 0$ and $a$ is continuous.

2.2. Reflection. The only important difference between the X-ray transform and the broken ray transform is the existence of reflections. As reflections can be difficult to handle in generic geometry, it is lucrative to look for methods that get rid of them altogether.

We reflect the domain $\Omega$ at $R$ to construct a new domain $\tilde{\Omega}$ so that the natural projection map $\tilde{\Omega} \to \Omega$ takes lines into broken rays. This construction is most clear
Figure 2. The half disc is first embedded in a stadium-shaped domain (solid line) so that the reflecting part (thick line) lies on the boundary. It is technically convenient to embed a manifold with corners to a manifold with smooth boundary so that the reflecting part still lays on the boundary. Then one can glue another copy of the larger manifold (dashed line) along a part of the boundary. The original boundary with corners is thus reflected to a manifold with smooth boundary – the whole disc in this case. [II]

in the half plane example given in section 2 where $\Omega = \mathbb{R} \times [0, \infty)$ and $\tilde{\Omega} = \mathbb{R}^2$. This correspondence between broken rays in $\Omega$ and lines in $\tilde{\Omega}$ allows one to reduce injectivity (and reconstruction and stability) of the broken ray transform on $\Omega$ to that of the X-ray transform on $\tilde{\Omega}$. If $R$ is something else than part of a hyperplane in a Euclidean space, the reflected domain $\tilde{\Omega}$ is not Euclidean. We recognize two types of “non-hyperplane geometry” for the reflecting part: $R$ may be curved or may contain corners.

The reflection construction may also be carried out when $\Omega$ is a Riemannian manifold – this is done in the article [II]. We wish the new manifold $\tilde{\Omega}$ to be such a manifold that the X-ray transform on it is injective. Known injectivity results for Riemannian manifolds assume that the metric and the boundary are smooth.

If $\partial \Omega$ is to be smooth, we need to have a corner on $\partial \Omega$ where $E$ and $R$ meet. The need of corners is easily seen when one reflects a half disc (which has corners) into a whole disc (which has smooth boundary). This example is depicted in figure 2.

Smoothness of the metric on the reflected manifold $\tilde{\Omega}$ requires that $R$ is very flat. If the metric $g$ on $\Omega$ is $C^\infty$, then the reflected metric $\tilde{g}$ is $C^1$ if and only if the second fundamental form vanishes identically on $R$. By symmetry $\tilde{g} \in C^1$ if and only if $\tilde{g} \in C^2$. Higher order flatness is required on $R$ if higher order differentiability of $\tilde{g}$ is required. If $R$ is strictly concave or convex, solutions to the geodesic equation on $\tilde{\Omega}$ branch or do not exist. For more details on the effect of regularity of $R$, see [II, Lemma 13].

This reflection method is well known in the study of billiards [31], but has only recently been used for the broken ray transform. Hubenthal [12] used reflections to...
study the broken ray transform in the square using microlocal techniques. He has recently [III] used a similar methods for general planar domains where $R$ is flat but may contain many components. In this setting all broken rays in some neighborhood of a fixed one have the same reflection pattern (they hit the same components of $E$ in the same order) and the domain can be unfolded so that each of these broken rays becomes a line – note that for different patterns we obtain a different reflected domain $\tilde{\Omega}$.

The main theorem [III, Theorem 16] states that if the X-ray transform is injective on $\tilde{\Omega}$, then the broken ray transform is injective on $\Omega$. However, for the X-ray transform to make sense and to be injective on $\tilde{\Omega}$, it is useful to have $R$ flat.

Curved $R$ is at present untractable, but corners are not. Treating corners on Riemannian manifolds is technically difficult, but we can give Euclidean examples. Proposition [III, Proposition 6] states that the broken ray transform is injective in a domain $\Omega$ contained in a planar cone with any opening angle where $R$ is contained in the boundary of the cone. One example of such a domain is a square where $E$ consists of two adjacent sides.

If the opening angle is $\pi/m$ for some $m \in \mathbb{N}$, then $2m$ copies of $\Omega$ can be neatly glued together (along matching sides) to form the domain $\tilde{\Omega}$. The X-ray transform is injective in planar domains, and injectivity of the broken ray transform in $\Omega$ follows. For general angles the result reduces to Helgason’s support theorem [10] instead of injectivity of the X-ray transform. We glue together so many copies of $\Omega$ that $\tilde{\Omega}$ has opening angle at least $\pi$. Then the integral of the reflected function $\tilde{f}$ is known over all lines in the plane that evade a convex cone, and by the support theorem $\tilde{f}$ is determined outside the cone. These two constructions are illustrated in [III, Figures 2–5]. The simpler case is illustrated here in figures 3 and 4.

The reflection argument can also be used for the periodic broken ray transform. Examples of this are given in [III, Propositions 30 and 31], and here we mention the example of the square. Just like in the above construction for cones, we take four copies of the square and glue them together. Then we identify the opposite sides

**Figure 3.** A cone $\Omega$ with opening angle $\pi/3$ and a broken ray. [III]
Figure 4. A domain $\tilde{\Omega}$ obtained from the domain $\Omega$ in figure 3 by gluing six copies together. The broken ray unfolds into a straight line. [II]

of the square – we have thus constructed a flat torus. Now any periodic geodesic on the torus corresponds to a periodic broken ray in the square and we arrive at a new question: do the integrals of a function on the torus over all periodic geodesics determine the function? The answer to this question is affirmative [I, IV], whence the periodic broken ray transform on the square is injective.

2.3. Boundary determination. Let $\sigma : [0, L] \to \partial \Omega$ be a geodesic on $\partial \Omega$ with endpoints in $\interior E$. We want to approximate $\sigma$ with a sequence $(\gamma_n)$ of broken rays in $\Omega$. If $(\gamma_n(0), \dot{\gamma}_n(0)) \to (\sigma(0), \dot{\sigma}(0))$, the broken rays are closer and closer to being tangent to $\sigma$ at the starting point. To have the broken rays remain almost tangent to $\sigma$ at later times, $\partial \Omega$ has to be strictly convex along $\sigma$ – this means that $\mathbb{II}(\dot{\sigma}, \dot{\sigma}) > 0$, where $\mathbb{II}(\cdot, \cdot)$ is the second fundamental form. We call such boundary geodesics $\sigma$ admissible. Under these conditions $(\gamma_n, \dot{\gamma}_n) \to (\sigma, \dot{\sigma})$ uniformly on $[0, L]$ (see [III, Lemma 7]). If $\partial \Omega$ is strictly concave along $\sigma$, the broken rays $\gamma_n$ escape to the interior of $\Omega$.

Assuming that $f : \bar{\Omega} \to \mathbb{C}$ is continuous, this convergence implies that we may reconstruct the integral of $f$ over admissible boundary geodesics from its broken ray transform. If there is an admissible boundary geodesic through each point on $E$,
it is easy to reconstruct \( f \) on \( E \). If the X-ray transform (restricted to admissible geodesics) is injective on \( \bar{R} = \partial \Omega \setminus \text{int} \ E \), we may thus reconstruct \( f|_{\bar{R}} \) from the broken ray transform of \( f \). Thus under favourable circumstances we may reconstruct \( f|_{\partial \Omega} \) from \( \mathcal{G}f \).

This boundary determination method becomes more likely to work when the number of admissible boundary geodesics increases. If we assume \( \partial \Omega \) to be strictly convex, every boundary geodesic is admissible. There are some results for injectivity of the X-ray transform on Riemannian manifolds (see [III, p. 7]), but current results are insufficient in very general situations.

Somewhat surprisingly, this method allows to construct not only the values of the unknown function at the boundary, but also its normal derivatives of any order. For \( k \geq 1 \), knowledge of lower order derivatives and the broken ray transform allows one to reconstruct the integral of \( \mathcal{S}\partial \Omega \ni (x, v) \mapsto \Pi(v, v)^{-k/3} \partial^k \nu f(x) \) over all admissible geodesics. Thus, if the X-ray transform on \( \partial \Omega \) is injective when weighted with powers of the second fundamental form, one may recover the Taylor polynomial of a function at the boundary from its broken ray transform. This boundary determination result is given in [III Corollary 2].

Like the approach presented in section 2.1, this one is based on the limit of infinitely many reflections. Also, similarly to the approach in section 2.2, this one reduces the problem to the X-ray transform – this time with weight.

### 2.4. PDE approach.

This approach to the broken ray transform was introduced by Mukhometov [20, 21, 22, 23, 24, 25], and we describe the main idea. Let \( \Omega \subset \mathbb{R}^2 \) be an annular domain, that is, a strictly convex bounded domain from which a strictly convex obstacle has been removed. Let the outer boundary be the set of tomodraphy \( E \) and the inner boundary reflective (the set \( \bar{R} \)). Suppose \( f \in C^2(\bar{\Omega}) \) has vanishing broken ray transform. Our goal is now to show that \( f = 0 \). The results in [V] hold for Riemannian surfaces, but for the sake of simplicity we assume here \( \Omega \) to be Euclidean.

Let \( S\Omega \) denote the unit sphere bundle of \( \Omega \), which in the Euclidean case is simply \( S\Omega = \Omega \times S^1 \). A point in \( S\Omega \) defines a point in \( \Omega \) and a direction at that point. For \( (x, v) \in S\Omega \) we have a unique broken ray \( \gamma_{x,v} : [0, \tau_{x,v}] \to \Omega \) satisfying \( \gamma_{x,v}(0) = x \) and \( \dot{\gamma}_{x,v}(0) = v \). Here \( \tau_{x,v} \) is the time it takes for the broken ray to reach the set \( E \).

For a sufficiently smooth function \( u : S\Omega \to \mathbb{R} \) we define the derivatives (vector fields) \( X \) and \( V \) by letting \( Xu(x, v) = v \cdot \nabla_x u(x, v) \) and \( Vu(x, v_\theta) = \partial_\theta u(x, v_\theta) \), where we have written \( v_\theta = (\cos \theta, \sin \theta) \).

Now we define a function \( u^f \) on \( S\Omega \) by letting

\[
u^f(x, v) = \int_0^{\tau_{x,v}} f(\gamma_{x,v}(t)) \, dt.
\]
By the fundamental theorem of calculus this function satisfies $Xu^f(x, v) = -f(x)$. But since $f(x)$ does not depend on the direction $v$, we get $V Xu^f(x, v) = 0$ for all $(x, v) \in S\Omega$.

We would like the equation $V Xu = 0$ to have a unique solution, but this obviously requires that we fix some boundary conditions. The boundary of $S\Omega$ is $E \times S^1 \cup R \times S^1$. The function $f$ has vanishing broken ray transform, so $u^f(x, v) = 0$ for all $(x, v) \in S\Omega$.

We want the equation $V Xu = 0$ to have a unique solution, but this obviously requires that we fix some boundary conditions. The boundary of $S\Omega$ is $E \times S^1 \cup R \times S^1$. The function $f$ has vanishing broken ray transform, so $u^f(x, v) = 0$ for all $(x, v) \in S\Omega$.

At every $x \in R$ we have the outer unit normal $\nu(x)$. We define the map $\rho_x : S^1 \to S^1$ by $\rho_x v = v - 2 \langle v, \nu(x) \rangle \nu(x)$; this map reflects $v$ across the tangent of $R$ at $x$. Since broken rays reflect so that the incoming and outgoing directions $v_i$ and $v_o$ are related via $v_o = \rho_x v_i$ (for a reflection at $x$), we have that $u^f(x, v) = u^f(x, \rho_x v)$ for all $x \in R$. This is the boundary condition on $R$.

It thus remains to show that $u = 0$ is the only solution to

$$
\begin{cases}
V Xu = 0 & \text{in } S\Omega \\
u = 0 & \text{on } E \times S^1 \\
u(x, v) = u(x, \rho_x v) & \text{on } R \times S^1.
\end{cases}
$$

We cannot assume that $u \in C^2(S\Omega)$ even if $f \in C^2(\bar{\Omega})$, since broken rays do not depend smoothly on their initial point and direction close to tangential reflections. We will return to this regularity issue shortly.

As it turns out, for $u \in C^2(S\Omega)$ we have [V, Lemma 8]

$$
\|V Xu\|_{L^2(S\Omega)}^2 = \|X Vu\|_{L^2(S\Omega)}^2 + \|X u\|_{L^2(S\Omega)}^2 - (K u, u)_{L^2(\partial SM)} - (\nabla_T u, u)_{L^2(\partial SM)},
$$

where $\nabla_T = \langle \nu^\perp, \nabla \rangle$ is the tangential derivative and $K$ is the Gaussian curvature of $\Omega$. Of course $K = 0$ for an Euclidean domain $\Omega$, but the same identity holds true on Riemannian surfaces where $K$ need not vanish.

One can further manipulate the boundary term of equation (1). If $\kappa$ denotes the signed curvature of $\partial M$ and $u_e$ and $u_o$ are the even and odd part of $u$ on $\partial M$ with respect to the reflection $\rho$, we have [V, Lemma 9]

$$
(\nabla_T u, u)_{L^2(\partial SM)} = (\nabla_T u_e, u_o)_{L^2(\partial SM)} + (\nabla_T u_o, u_e)_{L^2(\partial SM)} - (\kappa u_e, u_o)_{L^2(\partial SM)}.
$$

Combining the identities (1) and (2) with the boundary conditions $u = 0$ on $E \times S^1$ and $u_o = 0$ on $R \times S^1$, we find

$$
\|V Xu\|_{L^2(S\Omega)}^2 = \|X Vu\|_{L^2(S\Omega)}^2 + \|X u\|_{L^2(S\Omega)}^2 - (K u, u)_{L^2(\partial SM)} - (\kappa u, u)_{L^2(\partial SM)}.
$$

Energy estimates of this type are known as Pestov identities.
We need $u$ to have enough regularity so that the calculations leading to the identity (3) can be justified. Proving this regularity is rather involved, and we omit it here. The outcome of the regularity analysis is that if $f \in C^2(\Omega)$ has vanishing broken ray transform, then $u^f$ has enough regularity for the identity (3) to hold.

The obstacle is convex, so $\kappa \leq 0$. In Euclidean domains $K = 0$, and more generally on nonpositively curved Riemannian surfaces $K \leq 0$. These conditions give a positive sign to the last two terms in (3). As observed above, $VXu^f = 0$ and $Xu^f = -f$, so (3) yields

$$0 \geq \|f\|^2_{L^2(\Omega)}.$$  

This implies that $f = 0$, which we set out to prove.

If we could prove the Pestov identity (3) for $u^f$ without assuming that the broken ray transform of $f \in C^2(\Omega)$ vanishes, we would obtain

$$C \left\| Vu^f \right\|^2_{L^2(E \times S^1)} \geq \|f\|^2_{L^2(\Omega)},$$

where the constant $C$ depends on the maximum of the curvature $\kappa$ at $E$. Since with natural identifications $u^f|_{E \times S^1} = Gf$, this would be a stability estimate for the broken ray transform.

In the Euclidean case this result actually follows easily from Helgason’s support theorem [10]. The support theorem states that if a continuous function in the bounded domain $\Omega$ integrates to zero over all lines that avoid the convex obstacle, then it has to vanish outside the obstacle. Similar support theorems are also available on manifolds [15,33], but they cannot be used on two dimensional manifolds with nonanalytic metric.

Thus the result in [V] only gives new information on nonanalytic surfaces. The approach used is suitable for tackling the same problem with several convex obstacles, whereas support theorems can only give information of the exterior of the convex hull of the union of the obstacles.

The broken ray transform was shown to be injective in the case of several reflecting obstacles in the Euclidean plane by Eskin [8]. That result, however, relies on stringent geometrical assumptions on the obstacles, which for example prevent them from being smooth. Showing injectivity of the broken ray transform in a planar Euclidean domain with two smooth, convex obstacles remains an open problem. Actually, if one only assumes the two obstacles to be convex, there is an example (see figure 5) of noninjective broken ray transform, so at least one of the obstacles should be strictly convex.

Pestov identities have been widely used in the study of the X-ray transform. Such energy identities have proven to be useful when there is not enough structure or symmetry for explicit calculations. For a review of the use of Pestov identities, see [27, Section 4]. Mukhometov has also used Pestov identities to study the broken ray transform [20,21,22,23,24,25].
3. Examples and counterexamples

To make the results more concrete, we list examples of domains $\Omega$ (with the corresponding set of tomography $E$ or its complement $R = \partial \Omega \setminus E$) where the broken ray transform has been proven to be injective:

- Euclidean ball with arbitrarily small open $E$ for functions that are analytic in the angular variable in a suitable sense [I],
- any Euclidean cone for piecewise continuous functions where $R$ is contained in the surfaces of the cone [II],
- quarter of $S^n$, $n \geq 3$, where $E$ is half of the boundary for compactly supported $L^2$ functions [II],
- two dimensional hemisphere where $E$ is slightly larger than half of the boundary for compactly supported smooth functions [II], and
- nonpositively curved compact Riemannian surface with strictly convex boundary $E$ with a convex obstacle removed (boundary of the obstacle is $R$) for $C^2$ functions [V].

Suppose we know the broken ray transform of a function in the Euclidean ball (of dimension three or higher) which is smooth in a neighborhood of the boundary. Assuming that $E$ is open and contains a hemisphere, we may reconstruct the Taylor polynomial of the function at all boundary points from its broken ray transform. [III, after Remark 5]

There are also examples for the periodic broken ray transform, where all of the boundary is reflective. These examples are an octant of $S^2$ for compactly supported smooth functions [II] and the cube $[0,1]^n$, $n \geq 2$, for a sum of an $L^1$ function and a compactly supported distribution [IV].
In addition to examples, we also have counterexamples. If a manifold with boundary contains a (generalized) reflecting tubular part, then the broken ray transform (periodic or regular) is not injective [II]. For a precise statement see [II, Proposition 29]; we only remark here that this includes the counterexample of figure 5. There is also a counterexample for the periodic broken ray transform; namely, the periodic broken ray transform is not injective on compactly supported smooth functions in the Euclidean unit disc [II].

Many of these examples and counterexamples are described in more detail in [II, Sections 6–7].

4. Applications

The X-ray transform has important applications in medical imaging. Inferring the three dimensional structure of an object from X-ray images taken from all directions relies on our ability to invert the X-ray transform. This mathematical model underlies CT, PET and SPECT imaging. For details of these applications, see [26].

The X-ray transform also arises as a linearization of other inverse problems. For example, if one linearizes the boundary rigidity problem (see section 1.1) with respect to a conformal variation of the metric, one ends up with the X-ray transform. If one has a suitable form of the inverse function theorem (the spaces in question are infinite dimensional, so this is not obvious), solving the linearized problem shows that the original problem can be solved locally. For an example of a boundary rigidity problem solved via linearization, see [30].

Applications of the broken ray transform are described in [V, Section 6], and we give a brief summary of that discussion here.

Imaging with broken rays is relevant for seismology. To give a hint of this vast area of research, we refer to [9, 6, 7].

Eskin [8] reduced an inverse boundary value problem for the electromagnetic Schrödinger operator to the injectivity of the broken ray transform. Similarly, Kenig and Salo [14] reduced a partial data problem for Calderón’s problem in a tubular domain to the injectivity of the broken ray transform in the transversal domain. These results are based on constructing solutions to the corresponding PDEs which concentrate near broken rays. Therefore new results for the broken ray transform imply results for these problems; an example is given in [II, Theorem 3].

Like the X-ray transform, the broken ray transform also arises as a linearization. Let \((M, g)\) be a Riemannian manifold with boundary and let its boundary be partitioned in two sets \(E\) and \(R\) as above, and suppose \(E\) is open. Let \(x \in E\) be a boundary point and \(v \in S_x M\) a unit vector at it. Assuming that \(v\) points into the interior of \(M\), there is a unique broken ray \(\gamma_{x,v}\) with initial point and velocity \((x, v)\). Suppose \(\gamma_{x,v}\) is has no tangential reflections. Let then \(f \in C^\infty(M)\) be any function and define new metrics \(g_s = (1 + sf)g\) for all \(s \in (-\varepsilon, \varepsilon)\). If \(\varepsilon\) is small enough, the tensors \(g_s\) really are Riemannian metrics. Then let \(\gamma_{x,v}^s\) be the unique broken ray
with respect to the metric $g_s$ starting at the point $(x, v)$. By making $\varepsilon$ smaller if necessary we may assume that none of these broken rays has tangential reflections.

Let $\tau_s$ denote the length of $\gamma_{x,v}^s$ with respect to the metric $g_s$. If the broken rays $\gamma_{x,v}^s$ have the same endpoint, then the derivative $\frac{d}{ds} \tau_s|_{s=0}$ is the integral of the function $2f$ over the broken ray $\gamma_{x,v}$. In this sense the broken ray transform is the linearization of the boundary distance function with reflections. This result holds in more generality; see [V, Theorem 17].

Ray transforms are also related to spectral geometry. Spectral properties of a domain or a manifold can often be related to the lengths of periodic geodesics or broken rays (billiard trajectories) on it; see [5] for a review of the topic. Since lengths of broken rays correspond to the broken ray transform via linearization as described above, it is natural to expect the periodic broken ray transform to have applications in spectral geometry. As far as we know, the precise connection between the periodic broken ray transform and spectral geometry of manifolds with boundary is yet to be discovered.
References

25


A. A beginner’s introduction to inverse problems

One constantly encounters situations where something has to be measured: speed of a car, volume of flour, power of a lamp, or something else. It is of utmost importance that suitable tools are available for measurement so that one may successfully measure – and live.

Measurement requires two kinds of tools, technical and mental. If we want to measure the speed of a car, we drive it on a road, measure the spent time by a clock, and count the number of times a wheel has gone around to figure out the distance. We also need to know the circumference of the wheel, which we can easily measure with a tape measure. Thus, using three technical tools (clock, counter and tape measure), we have found three quantities, but none of them is the speed.

Next we need mental tools, that is, methods for reasoning and computing. Multiplying the circumference by the number of rotations gives the total length. Dividing this by the time gives the average speed. We have now finally measured the speed, and both technical and mental tools have been of use.

Sometimes a measurement cannot be done directly. If a physician wants to measure the density of a patient’s thighbone, he should not remove the bone from the thigh and measure its volume and mass to obtain the density as their quotient. Care for the patient’s health poses a limit for the method, and the measurement is to be done indirectly. One has to use some other way to figure out the density of the bone. Because this reasoning is not always easy, the importance of mental tools is great in comparison to technical ones in indirect measurements.

A.1. Direct and inverse problems. The goal of inverse problems research is to produce tools of both kinds for indirect measurements. My own research is focused on the mathematical side of inverse problems, so for me inverse problems are somewhat synonymous with the mathematics of indirect measurements. Mental tools become central, and multiplication and division no longer suffice for mathematical tools.

The nature of a direct problems is the following: when the properties of an object are known, one has to determine its behaviour. If for example the length, thickness, density and tension of a guitar string are known, one can figure out the sound it makes. In this case the sound is composed of a fundamental frequency and its multiples. From the mentioned properties one can determine all the possible frequencies, but not their strengths.

An inverse problem asks the same question in a different direction: when the behaviour is known, one has to determine the properties. In the guitar example we would hear a guitar string, and from the sound we should figure out the properties of the string (length, thickness, density and tension). But this cannot be done, since if for example the diameter is halved and length doubled, the sound stays the same. The inverse problem does not have a clear solution, but something is known about the properties of the string.
If three of the four properties of the string are known and the frequencies have been measured, the fourth one can be figured out. If we can measure the length and thickness and we know the material (we can check a table book for its density), we can work out the tension. When formulated in this way, the inverse problem does indeed have a unique solution.

What happens if the guitar is replaced with a drum? If the properties of the membrane of the drum are known, one can calculate the sound it makes (direct problem). But if we hear a drum and we know all about the membrane but its shape, can we deduce the shape? No one can give full answer to this question, but something is known.

If the sound sounds like it could come from a perfectly round drum, then it certainly comes from a round drum. In other words, one can always recognize a round drum by its sound. But there are examples of different drums with corners which make an identical sound, so there are also drums that cannot be recognized.

What drums can then be recognized by their sound? This problem is under intense study, and the problems are in the mental tools. This problem can be formulated mathematically, but solving it is very difficult. Eagerness to this research does not stem from a widespread drum hobby, but the same mathematical model has also other applications.

The direct and inverse problems described above represent a field called integral geometry. As its name suggests, it studies the relation between spectrum (frequencies of a sound) and geometry (length or shape). There are also other kinds of inverse problems, but these make a good example.

An example of an indirect measurement resembling the instrument examples is known to many. Watermelon farmers can tell whether a melon is ripe by thumping it and listening, and similar reasoning has been used for ages.

Inverse problems are laborious to study. As the examples demonstrate, a problem does not always have a unique solution (the same behaviour can be caused by many different properties) and a problem is more difficult than the corresponding direct problem. It is very interesting – and useful for applications – only to find out, which inverse problems have unique solutions, or in other words, which methods of indirect measurement could work even in principle.

A.2. Weighing a rope. Let us look for a better understanding of indirect measurements by means of an example which is related to the previous ones. We set out to weigh a rope which is fixed to a wall from both ends like a slack clothesline. If we can detach the rope and put it on a scale, the problem is easy, but it becomes more difficult if we disallow detaching it. Such prohibition from breaking the object under study is more natural in medical applications, because a broken patient cannot be rebuilt, but let us not be distracted by this.

One possibility is to use the rope as a guitar string and use earlier ideas, and so we shall do. We place a bar of some kind next to one fastening point and let the
rope lay over it. (If we cannot find a suitable support for the bar, we ask a friend to hold it still.) In the short part of the rope between this support and the wall we place a weight whose mass is known and much greater than that of the rope.

Now the longer part of the rope is tense and can be played as a guitar string. Pulling the rope slightly down in the middle and letting go, we can make it oscillate beautifully. We measure the time needed for ten oscillations by a clock to find out the frequency. If all of the rope goes up and down in the same rhythm, this frequency is the fundamental frequency.

The length and thickness of the rope are easy to measure. In addition we need to know the tension. That we can find out, when we know the mass of the weight and measure the angle of the rope on which it hangs.

Now we are in the situation described earlier: we know the length, thickness and tension of the rope and its fundamental frequency. From these we can calculate is density, and multiplying it with the volume gives the mass. This is of course only the mass of the tense part, but the mass of the entire rope is easily figured out since we can see what portion of the rope is tense. We have thus finally found the mass of the rope.

A.3. X-ray tomography. The solution of one inverse problem has made the life of many people easier and deserves a mention. But let us begin with the direct problem and description of the situation.

X-rays are very similar to light rays. X-ray radiation is light whose color the human eye cannot see – it is in a way too blue. Such light can be produced and measured with a suitable lamp and camera, so a properly equipped human can see with X-rays. Unlike usual light it mostly goes through a human body, so it can be used to see inside a human.

A broken bone is easy to see in an X-ray image, but some structures are difficult to see in one image. And even if there were multiple images, it can be difficult to see something inside a bone, for example, because something else is always in the way.

If one knows the precise structure of a human body, that is, what is in inside and where exactly it is, one can fairly easily reason what an X-ray image taken from any direction should look like. The corresponding inverse problem asks whether one could figure out the exact three dimensional structure of the body from X-ray images from all directions. It turns out that this is indeed possible, but it is not easy. One needs the help of a computer because the procedure requires complicated calculations, but this is easy for a modern computer. Due to this dependency on computers this method is often called computerized tomography (CT).

This indirect measurement method is very useful. Every day in hospitals around the world doctors use it to see inside their patients. It is therefore no wonder that it earned its discoverers Cormack and Hounsfield the Nobel prize in medicine in 1979.
My own research is related to inverse problems similar to X-ray tomography, but describing in more detail does not fit here. However, inverse problems of different kinds have surprising connections, and for example the inverse problem of X-ray tomography is closely related to spectral geometry which was described above. This is the salt of mathematical research; hardly anyone could have guessed that hearing the shape of a drum is eventually very similar to finding a fracture in a bone!
B. Aloittelijan johdatus inversio-ongelmiin


Mittaanmisessa tarvitaan kahdenlaisia välineitä, teknisiä ja henkisiä. Jos haluamme mitata auton vauhdin, ajamme jonkin tienpätkän ja mittaamme ajamiseen kuluneen ajan kellolla ja ajomatkan laskemalla montako kierrosta auton rengas on matkan aikana pyörähtänyt. Lisäksi tarvitsemme tiedon renkaan ympärysmitasta, ja sen voimme tehdä mittanauhalla. Näm teknisiä apuvälineitä (kello, kierroslaskuri ja mittanauhu) käyttäen olemme saaneet kolme mittaustulosta, mutta mikään näistä ei ole auton vauhti.


Käänteinen ongela sen sijaan kysyy saman kysymyksen toiseen suuntaan: kun tiedetään jonkin esineen käytös, täytyy päätellä, millaiset sen ominaisuudet ovat. Kitaraesimerkissä meille siis soitettaisiin kitaran kieltä, ja kuullusta äänestä pitäisi
päättää kielen ominaisuudet (pituus, paksuus, tiheys ja jännitys). Tämä ei kuitenkaan onnistu, sillä jos vaikkapa kielen halkaisija puolittuu ja pituus tuplaantuu, pysyvää ääni samanlaisena. Käänteisongelmalla ei siis olekaan selkeää ratkaisua, mutta kielen ominaisuuksista tiedetään silti jotain.

Jos kielen neljästä ominaisuudesta tiedetään kolme ja taajuudet ovat selvinneet mittaamalla, voidaan neljässä aina päättää. Jos siis saamme mitata kielen pituuden ja paksuuden ja tiedämme sen materiaalinkin (voimme katsoa jostain taulukkokirjasta sen tiheyden), saamme selville kielen jännityksen. Tällä tavalla muotoiltuna käänteiseen ongelmasta löytyy yksikäsitteinen ratkaisu.


Jos ääni kuulostaa siltä, että se tulee täysin pyöreästi rumpusta, silloin se varmasti tulee pyöreästi rumpusta. Pyöreän rumpun äänestä ei siis voi erehtyä. Sen sijaan on olemassa lukuisia esimerkkejä erinäköisistä kulmikkaista rumpuista, joista lähtee tänään samanlainen ääni, joten on myös sellaisia rumpuja, joiden äänestä voi erehtyä.

Mitä rumpuja äänen perusteella voi sitten tunnistaa? Tätä kysymystä tutkitaan kuumeisesti, ja ongelmia ovat nimenomaisesti henkisissä työkaluissa. Tämä ongelma osataan muotoilla tarkasti matemaattisesti, mutta sen ratkaiseminen on hyvin vaikeaa. Into tähän tutkimukseen ei kuitenkaan johdu laajasta rumpuharrastuksesta, vaan samalla matemaattisella mallilla on muitakin sovelluksia.

Esitelty suorat ja käänteiset ongelmat edustavat spektraaligeometriaksi kutsuttua alaa. Nimisä mukaisesti se tutkii spektrin (äänen taajuudet) ja geometrian (pituus tai muoto) välistä suhdetta. Käänteisongelmia on moninkinlaisia, mutta näitä käyvät hyvin esimerkiksi.

Soitinaihaisia esimerkkejä muistuttava äänen perustuva epäsuora mittaus on monille tuttu. Vesimelonin viljelijä osaa vaatia melonia kumauttamalla ja kuuntelemalla päättää, onko se kypsä, ja vastaavanlaita päätelyä on osattu käyttää jo kauan.

Käänteisongelmat ovat työläitä tutkittavia. Kuten annetut esimerkit osoittavat, ei ongelmalla välttämättä ole yksikäsitteistä ratkaisua (sama käytös vo johtua monenlaisista ominaisuuksista) ja ongelma on vaikeampi kuin vastaava suora ongela. Varsin kiinnostavaa – ja sovellusten kannalta hyödyllistäkin – on jo pelkästään selvittää, millä käänteisongelmailla on yksikäsitteinen ratkaisu eli millaiset epäsuorat mittausmenetelmät voivat toimia edes periaatteessa.

B.2. Narun punnitus. Haetaan yhden edellisiin liittyvän esimerkin avulla vielä lisäämmärrystä epäsuoraan mittaamiseen. Otetaan tehtäväksi punnita naru,
joka on molemmista päistään seinässä kiinni löysän pyykkinarun tapaan. Jos voimmee irrottaa narun ja laittaa sen vaa’alle, on ongelma helppo, mutta siitä tulee vaikeampi, jos kiellämme sen irrottamisen. Tällainen rikkoksielto on luontevampi lääketieteellisissä sovelluksissa, koska rikottua potilasta ei voi yleensä enää koota, mutta äkkiämme antako sen häiritä.

Yksi mahdollisuus on käyttää narua kitaran kielenä ja soveltaa edellä esitettyjä oppajeja, ja näin me teemme. Sijoittamme ensin lähelle toista kiinnityskohtaa jonkinlaisen tangon, jonka yli narun annetaan kulkea. (Jos sopivaa telinettä ei löydy, pyydämme ystävämme pitelemään tankoa paikoillaan.) Tämän tuen ja seinän väliselle lyhyelle naruosuudelle ripustamme parin kilon punnuksen, jonka massan tiedämme ja joka on selvästi narua painavampi.

Nyt pidempi naruosuus on pingottunut, ja sitä voi soittaa kuin kitaran kieltä. Vetämällä narua keskeltä vähän alas ja päästämällä liikkeelle saamme narun väärähtelemään kauniisti. Mittaamme sekuntikellolla, kauanko narun kestäää tehdämme kyseisellä kyseessä, ja näin saamme selville kymmenen vähähdystä, ja näin saamme selville samassa tähdisissä ylös ja alas, on tämä taajuus narun perustaajuus.

Narun pituus ja paksumuus on helppo mitata. Lisäksi tarvitsemme tiedon jännitysvoimasta. Sen saamme selville, kun tiedämme pumukan massan ja mitaaamme, missä kulmassa punnusta kannatteleva naru on.


**B.3. Röntgen-tomografia.** Erään käänteisongelman ratkaisu on helpottanut monen ihmisen elämää, ja se ansaitsee maininnan. Aloitetaan kuitenkin suorasta ongelmasta ja tilanteen kuvailusta.

Röntgen-säteet ovat hyvin samanlaisia kuin valonsäteet. Röntgen-säteily on valoa, jonka värä ihmiskehän ei erota – se on tavallaan liian sinistä. Tällaisia valoja voidaan tuottaa ja mitata samaan tapaan kuin lampulla ja kameralla, joten sopivien laitteiden avustamana ihmisen kykenee näkemään Röntgen-säteiden avulla. Toisin kuin tavallinen valo se kulkee suurimmaksi osaksi ihmisen läpi, joten sen avulla voi nähdä ihmisen sisäänsä.

Röntgen-kuvasta näkee helposti murtuneen luun, mutta joitain rakenteita on vaikea nähdä yhdessä kuvasta. Ja vaikka kuvia olisi montakin, voi olla vaikea nähdä, millainen vaario esimerkiksi luun sisällä on, sillä joka suunnasta katsottuna tiellä on jotaan muuta.

Jos tiedätään ihmisen tarkka rakente, eli mitä missäkin kohdassa kehon sisällä on, voidaan tästä melko helposti päätellä, mitä mistä tahansa suunnasta otettu Röntgen-kuva näyttää. Vastaava käänteisongelman kysy, voidaanko kaikista eri
suunnista otetuista Röntgen-kuvista päätellä ihmisen tarkka kolmiulotteinen rakenne. Tämä päätely on kuin onkin mahdollista, mutta se ei ole helppoa. Päätellyyn tarvitaan avuksi tietokone, koska se edellyttää monimutkaisia laskuja, mutta nykyisiltä tietokoneilta tämä käy helposti. Tietokoneavusteisuuksen vuoksi tämä menetelmä kutsutaan usein tietokonetomografiaksi (lyhenne CT englanninkielisestä nimestä computerized tomography).


Oma tutkimukseni liittyy Röntgen-tomografian tapaisiin käänteisongelmiin, mutta sen esittely ei tähän kirjoitukseen sovi. Eri alojen käänteisongelmat kuitenkin liittyvät toisiinsa yllättävällä tavoilla, ja esimerkiksi Röntgen-tomografian käänteisongelma liittyy läheisesti edellä kuvattuun spektraaligeometriaan. Juuri tämä on matematiikan tutkimuksen suola; tuskin kukaan olisi arvannut, että rummun muodon ja luiden murtumien selvittäminen on lopulta hyvin samanlaista!
C. Introductio tironis in problemata inversa

Continenter occurrunt rerum condiciones, in quibus aliquid mensurandum est: velocitas autoraedae, capacitas farinae, effectus lucernae aut aliquid aliud. Quam maxime necesse est instrumenta mensurando apta habere, quo facilius mensurare – et vivere – possimus.

In mensurando instrumenta duorum generum habere oportet, technica et mentalia. Si volumus mensurare velocitatem autoraedae, vehimus ea et mensuramus tempus adhibitum horologio et longitudinem itineris circuitibus canthi numerandis. Scire volumus et longitudinem unius circuitus, quem funiculo metrico parare possimus. Sic instrumentis technicis (horologio, numeratro circuituum, funiculo metrico) tres mensurationes factae sunt, sed nulla horum velocitas est.


Interdum recte mensurare non licet. Si medicus densitatem ossis femoris aegri mensurare vult, non licet os e femore removere et deinde massam et capacitatem mensurare de eisque dividendo densitatem calculare. Salus aegri fovenda limitem ponit rationi mensurandi, et indirecte mensurandum est. Debet aliquo alio modo densitatem ratiocinari. Quia ratiocinatio talis haud facilis est, grave est pondus instrumentorum rationalium nec solum technicorum.

C.1. Problema rectum et inversum. Scopus investigationis problematum inversorum est instrumenta utriusque generis parare, quae in mensurationibus indirectis adhiberi possint. Ipse latus mathematicum problematum inversorum scrutor, quare haec problemata artem mathematicam indirecte mensurandi habeo. In foco sunt instrumenta mentalia, nec, quod ad mathematicen attinet, multiplicationem cum divisione satisfacit.

Tale est natura problema rectum: proprietatibus aliquiuii rei cognitis ratiocinandum est, quo modo illa se gerat. Si exempli causa notae sunt longitudo, crassitudo, densitas et tension chordae citharae, vox chorda edita facile calculatur. Voci enim inest frequentia fundamentalis necnon ea quoque numerorum naturalium multiplicata. His proprietatibus datis cognosci possunt omnes frequentiae, quas chorda edere potest, sed magnitudo cuiusque frequentiae ignota manet.

Problema inversum autem hanc quaestionem invertit: gestu alicuiuii rei noto eius proprietates ratiocinandae sunt. In exemplo nostro citharico vocem audimus et proprietates (longitudo, crassitudo, densitas et tension) nobis explorandae sunt. Hoc enim fieri non potest, quia si diametros chordae in dimidium minuitur et longitudo duplicatur, vox eadem manet. Itaque problema hoc inversum clara solutione caret, sed aliquid de indolibus chordae cognosci potest.
Si tres ex quattuor chordae proprietatibus notae sunt explorataeque frequentiae, quarta semper calculari potest. Si igitur longitudinem et crassitudinem mensurare licet ac materia nota est (densitas in libro tabulario legi potest), noscitur tensio. Problema inversum sic formatum solutionem unicam habet.

Quid si citharam in tympanum mutamus? Si proprietates membranae tympani notae sunt, vox facilis est calculatu (problema rectum). Sed si vocem tympani audimus et omnes proprietates praeter formam scimus, potestne forma eius explorari? Nemo responsum perfectum dare quit, sed nonnil iam notum est.

Si vox tympani talis est, ut eam membrana rotunda dare possit, certe membrana rotunda data est. Vox tympani rotundi numquam nos fallit. Cum autem multae formae angulosae, quae eandem vocem dent, inventae sint, possimus etiam errare de aliqibus formis tympanorum.

Quales tympanos possimus eorum vocibus auditis recognoscere? Haec quaestio acerrime investigatur, et problemata sunt in instrumentis mentalibus. Haec quaestio mathematica exacte formulari potest, sed eam solvere difficilimum est. Studium huius quaestionis non ab oblecturemus musicalibus oritur, sed idem exemplar mathematicum alias quoque usui est.

Problemata recta et inversa supra exposita pertinent ad aream mathematicae, quae geometria spectralis appellatur. Secundum nomen suum suum nexus inter spectrum (frequentiae vocis) et geometriam (longitudo sive forma) scrutatur. Alia quoque sunt problemata inversa, sed haec bonum exemplum praebent.

Mensuratio indirecta exemplis musicalibus similis nota est. Cultores melonis melonem quatiendo et audiendo ratiocinari possumus, an maturus sit, et talis ratiocinatio iam diu adhibita est.

Problemata inversa investiganda laboriosa sunt. Ut exempla demonstrant, interdum solutio unica problemati deest (variae proprietates eosdem gestus efficere possunt) et problema est difficilium quam versio eius recta. Interest – et usui est – investigare, quae problemata inversa solutionem unicam habeant sive quae methodi indirecte mensurandi omnino adhiberi possint.

C.2. Restis pensanda. Conemur uno exemplo mensurationes indirectas melius intelligere. Metam nobis ponamus restem pensandam, quae utroque limite parieti affixa est sicut restis linearia. Si restem abripere licet et ad libram ponere, problema facile est, sed difficilium fit, si eam non licet solvere. Prohibitio talis naturalior est in arte medicina, cum homo aegrotus fractus restitui non possit, sed ne hoc distrahamus.

Possimus restem sicut chordam citharae adhibere et meminisse supra scripta, et sic facimus. Ponimus prope alterum locum affixionis axem aliquiis generis, super quam restis eat. (Si nil est, in quo axem affigeremus, rogamus amicum, ut axem teneat.) In parte breviore restis suspendimus pondus duorum fere chiliogrammatum, cuius massa est multo maior quam restis et nobis nota.

Longitudo et crassitudo sunt faciles mensuratu. Volumus scire etiam tensionem. Eam cognoscimus, quia massa ponderis nota est et angulus partis restis, de qua pendet, facile mensuratur.


Radii Röntgeniani persimiles sunt radiis lucis. Radiatio Röntgeniana est lux, cuius colorem oculus humanus non capit – est, ut ita dicam, nimis caeruleus. Lux talis parari et mensurari potest sicut lux visibilis lampade et machina photographica, quare instrumentis aptis ornatus homo röntgenianus videt et potest sicut solitis radiis lucis. Dissimiliter luci hodiernae haec lux magnopere per corpus humanum it, quod efficit, ut interiora hominis videre liceat.

In imagine Röntgeniana os fractum facile videtur, sed non omnes structurae tam facile in una imagine percipiuntur. Etsi multae sint imagines, est defectus in medulla ossis plerumque difficilis percutu, cum ossibus aliosque structuris a quaque parte umbretur.

Si structura hominis accurate nota est, id est, quid ubique sit, facile est ratiocinari, quales imaginques Röntgenianae a quaque parte factae videantur. Problema inversum ad hoc attinens rogit, possitne structura hominis tridimensionalis ex imaginibus Röntgenianis ab omnibus partibus factis inveniri. Accidit, ut vero inveniri possit, sed haud facile est. In ratiocinando auxilio computatri opus est, quia computationes sunt complicatae, sed computatris hodiernis hoc facillime efficitur. Propri usum computaturae haec methodus plerumque tomographia computatralis appellatur (breviter CT ex verbis Anglicis computerized tomography).


Investigationes meae ad problema inversa tomographiae Röntgenianae similia attinent, sed expositio earum hic non apta est. Attamen diversa problema inversa nexus improvisos inter se habent, et exempli causa problema inversum tomographiae Röntgenianae nexus intimum habet cum geometria spectrali supra exposita. Hoc
est sal mathematices; haud quisquam coniectaverit ratiocinationem formae tympani et fracturae ossis simillimam esse!