SPHERICAL RADON TRANSFORMS AND MATHEMATICAL PROBLEMS OF THERMOACOUSTIC TOMOGRAPHY

A Dissertation

by

GAIK AMBARTSOUMIAN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2006

Major Subject: Mathematics

SPHERICAL RADON TRANSFORMS AND MATHEMATICAL PROBLEMS OF THERMOACOUSTIC TOMOGRAPHY

A Dissertation

by

GAIK AMBARTSOUMIAN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Peter Kuchment
Committee Members,	Francis J. Narcowich
	Joseph E. Pasciak
	Lihong Wang
Head of Department,	Albert Boggess

August 2006

Major Subject: Mathematics

ABSTRACT

Spherical Radon Transforms and

Mathematical Problems of Thermoacoustic Tomography. (August 2006)

Gaik Ambartsoumian, Dipl., Obninsk Institute of Nuclear Power Engineering, Russia

Chair of Advisory Committee: Dr. Peter Kuchment

The spherical Radon transform (SRT) integrates a function over the set of all spheres with a given set of centers. Such transforms play an important role in some newly developing types of tomography as well as in several areas of mathematics including approximation theory, integral geometry, inverse problems for PDEs, etc.

In Chapter I we give a brief description of thermoacoustic tomography (TAT or TCT) and introduce the SRT.

In Chapter II we consider the injectivity problem for SRT. A major breakthrough in the $2D$ case was made several years ago by M. Agranovsky and E. T. Quinto. Their techniques involved microlocal analysis and known geometric properties of zeros of harmonic polynomials in the plane. Since then there has been an active search for alternative methods, which would be less restrictive in more general situations. We provide some new results obtained by PDE techniques that essentially involve only the finite speed of propagation and domain dependence for the wave equation.

In Chapter III we consider the transform that integrates a function supported in the unit disk on the plane over circles centered at the boundary of this disk. As is common for transforms of the Radon type, its range has an infinite co-dimension in standard function spaces. Range descriptions for such transforms are known to be very important for computed tomography, for instance when dealing with incomplete data, error correction, and other issues. A complete range description for the circular

Radon transform is obtained.

In Chapter IV we investigate implementation of the recently discovered exact backprojection type inversion formulas for the case of spherical acquisition in 3D and approximate inversion formulas in 2D. A numerical simulation of the data acquisition with subsequent reconstructions is made for the Defrise phantom as well as for some other phantoms. Both full and partial scan situations are considered.

To my Parents.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor, Peter Kuchment, for his guidance and constant support throughout the years of research and life as a graduate student. Without his advice and help this dissertation would never be written.

I would like to thank the members of my advisory committee, Dr. Francis J. Narcowich, Dr. Joseph E. Pasciak and Dr. Lihong Wang, for their assistance.

The dissertation is partly based upon work supported by the NSF under Grants DMS 0296150, 9971674, 0002195, and 0072248. Chapter IV is partly based on the work that has been done in Summer 2004 at GE Healthcare Technologies, Milwaukee, WI as part of a summer internship program. I would like to thank NSF and GE Healthcare Technologies for their support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

CHAPTER I

INTRODUCTION

A. Thermoacoustic tomography

Most tomographic methods of medical imaging (as well as industrial non-destructive evaluation, geological imaging, sonar, and radar) are based on the following procedure: one sends towards a non-transparent body some kind of a signal (acoustic or electromagnetic wave, X-ray or visual light photons, etc.) and measures the wave after it passes through or reflects back from the body. Then the problem becomes to use the measured information to recover the internal structure of the object of study. The common feature of most traditional methods of tomography is that the sent and received signals have the same physical nature. Although the development of tomography during past several decades has brought many remarkable successes [55, 57], each of the methods has its own shortfalls. For instance, when imaging biological tissues, microwaves and optical imaging often provide good contrasts between different types of tissues, but are inferior in terms of resolution in comparison with ultrasound or X-rays. This, in particular, is responsible for practical impossibility of getting any good resolution in optical or electrical impedance tomography, unless one wants to image only skin-deep areas. On the other hand, ultrasound, while giving good resolution, does not do a good job in terms of contrast. It is amazing that the idea of combining different types of radiation for triggering the signal and for the measured signal had to wait for such a long time to appear. By now, thermoacoustic tomography and its sibling photoacoustic tomography (PAT) have already made significant advances (e.g., [37]-[40], [79]-[82]), while some others are still in a

---

This dissertation follows the style of *Inverse Problems*.

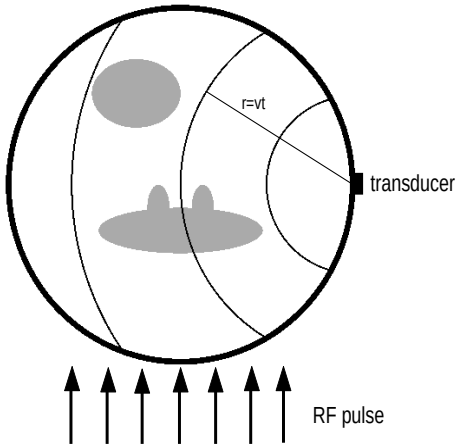


Fig. 1. Sketch of a TAT system

development stage. Since PAT in terms of the relevant mathematics is identical to TAT, we will describe briefly only the latter. In TAT, a short radiofrequency (RF) electromagnetic pulse is sent through the biological object (see Fig. 1). At each internal location x certain energy $f(x)$ will be absorbed. It is known (see [36] and the references above), that cancerous cells absorb several times more RF energy than the normal ones, which means that significant increases of the values of $f(x)$ are expected at tumorous locations. It is believed that this contrast is due to the increased water and sodium content in tumors, which is partly due to extra blood vessel growth there. The absorbed energy, due to resulting heating, causes thermoelastic expansion, which in turn creates a pressure wave. This wave can be detected by ultrasound transducers placed outside the object. Since the delivered pulses are very short, the thermal diffusion during the experiment can be neglected. Now the former weakness of ultrasound (low contrast) becomes an advantage. Indeed, due to that low contrast in many cases (e.g., for mammography) one can assume the sound speed v to be constant. Hence, the sound waves detected at any moment t of time are coming from the locations

at a constant distance (depending on time and sound speed) from the transducer. The strength of the signal coming from a location x reflects the energy absorption $f(x)$. Thus, one effectively measures the integrals of $f(x)$ over all spheres centered at the transducers' locations. In other words, to recover the internal structure of the object one needs to invert a generalized Radon transform of f ("generalized," since integration is done over spheres). Exact implementation of this idea involves simple handling of the wave equation ([38], [79]-[82]).

This method combines advantages of two types of radiation used (contrast for microwaves and resolution for ultrasound), while avoiding their deficiencies.

B. Spherical Radon transform

The transform integrating a function over the set of all spheres with a given set of centers is usually called spherical mean or circular Radon transform. We will use the two terms interchangeably throughout the text. Let $f(x)$ be a continuous function on $\mathbb{R}^n$, $n \geq 2$.

Definition 1. *The circular Radon transform of f is defined as*

$$Rf(p,r) = \int_{|y-p|=r} f(y) d\sigma(y),$$

where $d\sigma(y)$ is the surface area on the sphere $|y-p| = r$ centered at $p \in \mathbb{R}^n$.

In this definition we do not restrict the set of centers p or radii r. It is clear, however, that this mapping is overdetermined, since the dimension of pairs (p,r) is $n+1$, while the function f depends on n variables only. This and the tomographic consideration above suggest to restrict the set of centers to a set (hypersurface) $S \subset \mathbb{R}^n$, while not imposing any restrictions on the radii. We denote this restricted transform

by R_S:

$$R_S f(p,r) = Rf(p,r)|_{p\in S}.$$

The most popular geometries of these surfaces (curves) that have been implemented in TAT are spheres (circles), planes (lines), and cylinders [79]-[81].

Such transforms have been studied over the years in relation to many problems of approximation theory, integral geometry, PDEs, sonar and radar imaging, and other applications (e.g., [2, 3, 17, 20, 35, 43, 49, 50, 51, 59, 62, 63]). The central problems that arise in these studies are:

- Uniqueness of reconstruction: is the information collected sufficient for the unique determination of the energy deposition function f?
- Reconstruction formulas and algorithms, and stability of the reconstruction.
- Description of the range of the transform: what conditions should ideal data satisfy?
- Incomplete data problems: what happens to the reconstruction if only a part of transducers' locations can be (or are) used?

All these questions have been essentially answered for the classical Radon transform that arises in X-ray CT, Positron Emission Tomography (PET), and Magnetic Resonance Imaging (MRI) [55, 57]. However, they are much more complex and not that well understood for the circular Radon transform that arises in TAT. Although significant progress has been achieved, some related analytic problems have proven to be rather hard and remain unresolved till now.

CHAPTER II

UNIQUENESS OF RECONSTRUCTION[1]

A. Formulation of the problem

The main problem addressed in this chapter is motivated by the question whether the information collected in a TAT procedure is sufficient for the unique determination of the energy deposition function f. However, the mathematical formulation can be more general and involve questions of unique reconstruction of functions in much larger spaces than those arising in medical imaging (e.g. non-compactly supported functions).

Definition 2. *The transform R is said to be* **injective** *on a set S (S is a* **set of injectivity***) if for any $f \in C_c(\mathbb{R}^n)$ the condition $Rf(p,r) = 0$ for all $r \in \mathbb{R}$ and all $p \in S$ implies $f \equiv 0$.*

In other words, S is a set of injectivity, if the mapping R_S is injective on $C_c(\mathbb{R}^n)$.

Here we use the standard notation $C_c(\mathbb{R}^n)$ for the space of compactly supported continuous functions on $\mathbb{R}^n$. The situation can be significantly different and harder to study without compactness of support (or at least some decay) condition [2, 3]. Fortunately, tomographic problems usually yield compactly supported functions.

One now arrives to the

Problem 3. *Describe all sets of injectivity for the circular Radon transform R on $C_c(\mathbb{R}^n)$.*

---

[1] Part of this chapter is reprinted with permission from "On the injectivity of the circular Radon transform arising in thermoacoustic tomography", by G. Ambartsoumian and P. Kuchment, Inverse Problems **21** (2005), 473–485. Copyright ©2005 by IOP Publishing LTD and individual contributors.

This problem has been around in different guises for quite a while [3, 20, 49, 50]. The paper [3] contains a survey of some other problems that lead to the injectivity question for R_S.

B. Algebraicity of non-injectivity hypersurfaces

The first important observations concerning non-injectivity sets were made by V. Lin and A. Pincus [49, 50] and by N. Zobin [83]. For the completeness of exposition we repeat some of their results in this section. Even though the statements below are proved for compactly supported functions, the results remain valid also for non-compactly supported functions decaying sufficiently fast (e.g. exponentially) at infinity.

For every $f \in C_c(\mathbb{R}^n)$ define

$$\mathcal{S}[f] = \{x \in \mathbb{R}^n \mid Rf(x,r) = 0 \ \ \forall r \in \mathbb{R}_+\}.$$

For some functions, e.g. if the integral of f over the whole space is not zero, $\mathcal{S}[f] = \varnothing$.

Let $k = 0, 1, \ldots$. Then

$$Q_k(x) = Q_k[f](x) = r^{2k} * f \ \ (x) = \int_{\mathbb{R}^n} |x-\xi|^{2k} \, f(\xi) \, d\xi, \tag{2.1}$$

where $r = x_1^2 + \ldots + x_n^2$, is a polynomial of degree deg $Q_k \leq 2k$.

With each polynomial Q let us associate the algebraic variety

$$V[Q] = \{x \in \mathbb{R}^n \mid Q(x) = 0\}.$$

Then we have the following statements.

Lemma 4. $\mathcal{S}[f] = \bigcap_{k=0}^{\infty} V[Q_k]$.

Proof. The following conditions are equivalent:

(1) $Rf(x,r) = 0$ for all $r \in \mathbb{R}_+$

(2) $\int_{\mathbb{R}^n} \alpha(|x-\xi|^2) f(\xi)\, d\xi = 0$ for every $\alpha \in C_c([0,\infty))$

Indeed, (1)⇒(2) simply by passing to polar coordinates centered at x in the integral in (2). For (2)⇒(1) take a sequence of functions $\alpha_n(t)$ converging to the Dirac δ-function $\delta(t-r)$.

Consider now $x \in \mathcal{S}[f]$. Then $Rf(x,r) = 0$ for all $r \in \mathbb{R}_+$, and by equivalence proved above $\int_{\mathbb{R}^n} \alpha(|x-\xi|^2) f(\xi)\, d\xi = 0$ for every $\alpha \in C_c([0,\infty))$. This means in particular, $Q_k(x) = 0$ for every k, hence $x \in \bigcap_{k=0}^{\infty} V[Q_k]$. We can conclude now that $\mathcal{S}[f] \subset \bigcap_{k=0}^{\infty} V[Q_k]$.

Suppose now that $x \in \bigcap_{k=0}^{\infty} V[Q_k]$. Then $Q_k(x) = 0$ for every k. By Weierstrass' theorem about the denseness of polynomials and equation (2.1) we obtain $\int_{\mathbb{R}^n} \alpha(|x-\xi|^2) f(\xi)\, d\xi = 0$ for every $\alpha \in C_c([0,\infty))$. From the equivalence shown above $Rf(x,r) = 0$ for all $r \in \mathbb{R}_+$, i.e. $x \in \mathcal{S}[f]$. Hence $\bigcap_{k=0}^{\infty} V[Q_k] \subset \mathcal{S}[f]$. But we know that $\mathcal{S}[f] \subset \bigcap_{k=0}^{\infty} V[Q_k]$. So $\mathcal{S}[f] = \bigcap_{k=0}^{\infty} V[Q_k]$. ■

Proposition 5.

1. *$f \equiv 0$ if and only if $Q_k[f] \equiv 0$ for every $k = 0, 1, \ldots$.*
2. *For $f \not\equiv 0$ let $Q_{k_0}[f]$ be the smallest degree non-zero polynomial among Q_k's. Then Q_{k_0} is harmonic.*

Proof. 1. The condition $Q_k[f] \equiv 0$ for every $k = 0, 1, \ldots$ means $\bigcap_{k=0}^{\infty} V[Q_k] = \mathbb{R}^n$. From Lemma 4 it follows, that $\mathcal{S}[f] = \mathbb{R}^n$, in other words all integrals of f over all spheres in $\mathbb{R}^n$ are vanishing, i.e. $f \equiv 0$.

2. Applying Laplacian to the convolution in (2.1) and using the easily verifiable identity

$$\triangle |x|^{2k} = 2k(2k+n-2)|x|^{2k-2}$$

one arrives to the following relation

$$\triangle Q_k = 2k(2k+n-2)Q_{k-1}.$$

If $Q_{k_0}[f]$ is the nontrivial polynomial of minimal degree, then $Q_{k_0-1}[f] \equiv 0$. Hence $\triangle Q_{k_0} \equiv 0$ ■.

Lemma 4 and Proposition 5 imply that, if R is not injective on S, then S is contained in the zero set of a harmonic polynomial. Therefore we get a sufficient condition for injectivity:

Corollary 6. *Any set $S \subset \mathbb{R}^n$ of uniqueness for harmonic polynomials is a set of injectivity for the transform R.*

In particular, this implies

Corollary 7. *If $U \subset \mathbb{R}^n$ is a bounded domain, then $S = \partial U$ is an injectivity set for R.*

Proof. Indeed, if $f(x)$ is a harmonic function on U s.t. $f|_{\partial U} = 0$, then the maximum principle for harmonic functions implies that $f \equiv 0$. Now Corollary 6 completes the proof. ■

We will see later a different proof of this fact that does not use harmonicity.

C. Solution in the plane and conjecture for higher dimensions

So, what are possible non-injectivity sets? Any hyperplane S is such a set. Indeed, for any function f that is odd with respect to S, one gets $R_S f \equiv 0$. There are other options as well. In order to describe them in $2D$, let us first introduce the following definition.

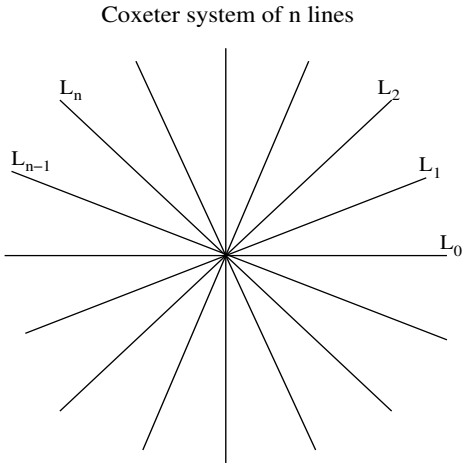


Fig. 2. Coxeter set

Definition 8. *For any $N \in \mathbb{N}$ denote by Σ_N the Coxeter system of N lines $L_0, \dots, L_{n-1}$ in the plane*[2] *(see Fig. 2):*

$$L_k = \{te^{i\pi k/n} | -\infty < t < \infty\}.$$

In other words, Σ_N is a "cross" of N lines passing through the origin and forming equal angles π/N. It is rather easy to construct a non-zero compactly supported function that is simultaneously odd with respect to all lines of a given Coxeter set (e.g., $f(\rho, \phi) = h(\rho)\sin(2n\phi)$ for arbitrary $h(\rho) \in C_0^\infty[0, \infty]$, where (ρ, ϕ) are polar coordinates). Hence, Σ_N is a non-injectivity set as well. Applying any rigid motion ω, one preserves non-injectivity property. It has been also discovered (e.g., [3]) that one can add any finite set F preserving non-injectivity. Thus, all sets $\omega\Sigma_N \cup F$ are non-injectivity sets. It was conjectured by V. Lin and A. Pincus that these are the only non-injectivity sets for compactly supported functions on the plane. Proving this conjecture, M. Agranovsky and E. Quinto established the following result:

---

[2] In the formula below we identify the plane with the complex plane $\mathbb{C}$.

Theorem 9. *[3] The following condition is necessary and sufficient for a set $S \subset \mathbb{R}^2$ to be a set of injectivity for the circular Radon transform on $C_c(\mathbb{R}^2)$:*

S is not contained in any set of the form $\omega(\Sigma_N) \bigcup F$, where ω is a rigid motion in the plane and F is a finite set.

The (unproven) conjecture below describes non-injectivity sets in higher dimensions.

Conjecture 10. *[3] The following condition is necessary and sufficient for S to be a set of injectivity for the circular Radon transform on $C_c(\mathbb{R}^n)$:*

S is not contained in any set of the form $\omega(\Sigma) \bigcup F$, where ω is a rigid motion of $\mathbb{R}^n$, Σ is the zero set of a **homogeneous** *harmonic polynomial, and F is an algebraic subset in $\mathbb{R}^n$ of co-dimension at least 2.*

For $n = 2$ this boils down to Theorem 9. Indeed, let $u_n(x, y)$ be a harmonic homogenous polynomial of order n. Then it is the real (or imaginary) part of the complex valued function $g(z) = z_0 z^n$, where $z_0 = Ce^{in\delta}$ is some constant. Using the polar representation of complex numbers, we get $u_n(r, \theta) = Cr^n \cos(n(\theta - \delta))$ (or in case of imaginary part $u_n(r, \theta) = Cr^n \sin(n(\theta - \delta))$). Thus $u_n = 0$ on the lines $\theta = \delta + \pi/2n + k\pi/n$ (or on the rays $\theta = \delta + k\pi/n$), for $0 \leq k \leq 2n - 1$. In both cases the zeros set of u_n is a Coxeter set.

D. Alternative approach

The beautiful proof of Theorem 9 by M. Agranovsky and E. Quinto is built upon the following tools: microlocal analysis (Fourier Integral Operators technique) that guarantees existence of certain analytic wave front sets at the boundary of the support of a function located on one side of a smooth surface (Theorem 8.5.6 in [34]),

and known geometric structure of level sets of harmonic polynomials in $2D$ (e.g., [24]). These methods, unfortunately, restrict wider applicability of the proof. The microlocal tool works at an edge of the support and hence is not applicable for non-compactly-supported functions. On the other hand, the geometry of level sets of harmonic polynomials does not work well in dimensions higher than 2 or on more general Riemannian manifolds (e.g., on the hyperbolic plane). Thus, the quest has been active for alternative approaches since [3] has appeared.

It is instructive to look at alternative reformulations of the problem (which there are plenty [3]). There is a revealing reformulation [3, 41] that stems from known relations between spherical integrals and the wave equation (e.g., [17, 35]). Namely, consider the initial value problem for the wave equation in $\mathbb{R}^n$:

$$u_{tt} - \triangle u = 0, \quad x \in \mathbb{R}^n, t \in \mathbb{R} \tag{2.2}$$

$$u|_{t=0} = 0, \quad u_t|_{t=0} = f. \tag{2.3}$$

Then

$$u(x,t) = \frac{1}{(n-2)!} \frac{\partial^{n-2}}{\partial t^{n-2}} \int_0^t r(t^2 - r^2)^{(n-3)/2} (Rf)(x,r) dr, \quad t \geq 0.$$

Hence, it is not hard to show [3] that the original problem is equivalent to the problem of recovering $u_t(x,0)$ from the value of $u(x,t)$ on subsets of $S \times (-\infty, \infty)$.

Lemma 11. *[3, 41] A set S is a non-injectivity set for $C_c(\mathbb{R}^n)$ if and only if there exists a non-zero compactly supported continuous function f such that the solution $u(x,t)$ of the problem (2.2)-(2.3) vanishes for any $x \in S$ and any t.*

Hence, non-injectivity sets are exactly the nodal sets of oscillating free infinite membranes. In other words, injectivity sets are those that observing the motion of the membrane over S gives complete information about the motion of the whole membrane.

One can now try to understand the geometry of non-injectivity sets in terms of wave propagation. The first example of such a consideration was the original proof [41] of Corollary 7 that did not use harmonicity (not known at the time). Let $S = \partial U$ be a non-injectivity (and hence nodal for wave equation) set, where U is a bounded domain. Then on one hand, the membrane is free and hence the energy of the initial compactly supported perturbation must move away. Thus, its portion inside U should decay to zero. On the other hand, one can think that S is a fixed boundary and hence the energy inside must stay constant. This contradiction allows one to conclude that in fact $f = 0$. The same PDE idea, with many more technical details, was implemented in [2] to prove the following statement:

Theorem 12. *[2] If U is a bounded domain in $\mathbb{R}^n$, then $S = \partial U$ is an injectivity set for R in the space $L^q(\mathbb{R}^n)$ if $q \leq 2n/(n-1)$. This statement fails when $q > 2n/(n-1)$, in which case spheres fail to be injectivity sets.*

In spite of these limited results, it still had remained unclear what distinguishes in terms of wave propagation the "bad" flat lines S in Theorem 9 that can be nodal for all times, from any truly curved S that according to this theorem cannot stay nodal. An approach to this question was found in the recent paper [23] by D. Finch, Rakesh, and S. Patch, where in particular some parts of the injectivity results due to [3] were re-proven by simple PDE means without using microlocal tools and harmonicity of non-injectivity hypersurfaces:

Theorem 13. *[23] Let D be a bounded, open, subset of $\mathbb{R}^n$, $n \geq 2$, with a strictly convex smooth boundary S. Let Γ be any relatively open subset of S. If f is a smooth function on $\mathbb{R}^n$ supported in $\bar{D}$, u is the solution of the initial value problem (1), (2) and $u(p,t) = 0$ for all t and $p \in \Gamma$, then $f = 0$.*

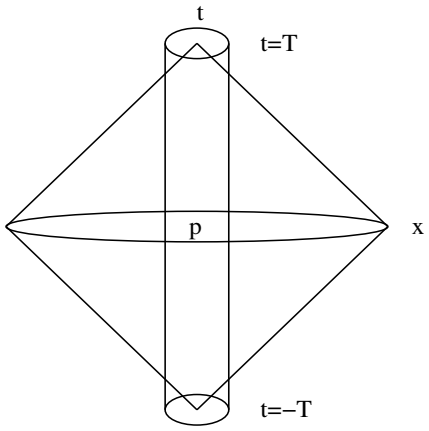


Fig. 3. Cones of unique continuation

Although this theorem follows from microlocal results in [3][3], its significance lies in the proof provided in [23] (that paper contains important results concerning inversion as well, which we do not touch here).

The following two standard statements concerning the unique continuation and finite speed of propagation for the wave equation were the basis of the proof of the Theorem 13 in [23]. They will be relevant for our purpose as well.

Proposition 14. *[23] Let* $B_\epsilon(p) = \{x \in \mathbb{R}^n \,|\, |x-p| < \epsilon\}$. *If* u *is a distribution and satisfies (2.2) and* u *is zero on* $B_\epsilon(p) \times (-T, T)$ *for some* $\epsilon > 0$, *and* $p \in \mathbb{R}^n$, *then* u *is zero on*

$$\{(x,t) : |x-p| + |t| < T\},$$

and in particular on

$$\{(x,0) : |x-p| < T\}$$

(see Fig. 3).

---

[3] Results of [3] make the situation described in Theorem 13 impossible, since the support of f lies on one side of a tangent plane to Γ. See also Theorem 21 and [51].

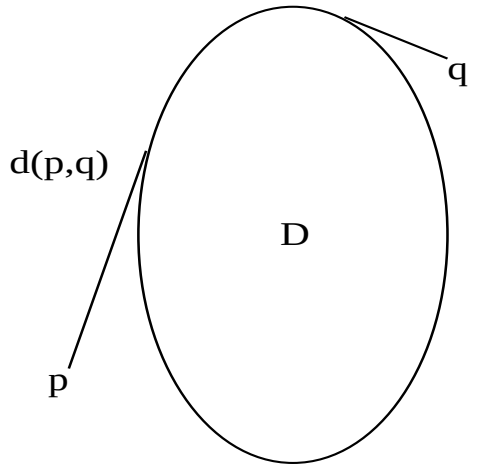


Fig. 4. "Curvy distance"

Let now D be a bounded, open subset of $\mathbb{R}^n$ with the boundary S. For points p, q outside D, let $d(p,q)$ denote the infimum of the lengths of all the piecewise C^1 paths in $\mathbb{R}^n \setminus D$ joining p to q. Then $d(p,q)$ is a metric on $\mathbb{R}^n \setminus D$ (see Fig. 4). For any point p in $\mathbb{R}^n \setminus D$ and any positive number r, define $E_r(p)$ to be the ball of radius r and center at p in $\mathbb{R}^n \setminus D$ with respect to this metric, i.e.

$$E_r(p) = \{x \in \mathbb{R}^n \setminus D : d(x,p) < r\}.$$

Proposition 15. *[23] Suppose D is a bounded, open, connected subset of $\mathbb{R}^n$, with a smooth boundary S. Let u be a smooth solution of the exterior problem*

$$u_{tt} - \triangle u = 0, \;\; x \in \mathbb{R}^n \setminus D, \;\; t \in R$$

$$u = h \;\; on \;\; S \times R.$$

Suppose p is not in D, and $t_0 < t_1$ are real numbers. If $u(.,t_0)$ and $u_t(.,t_0)$ are zero

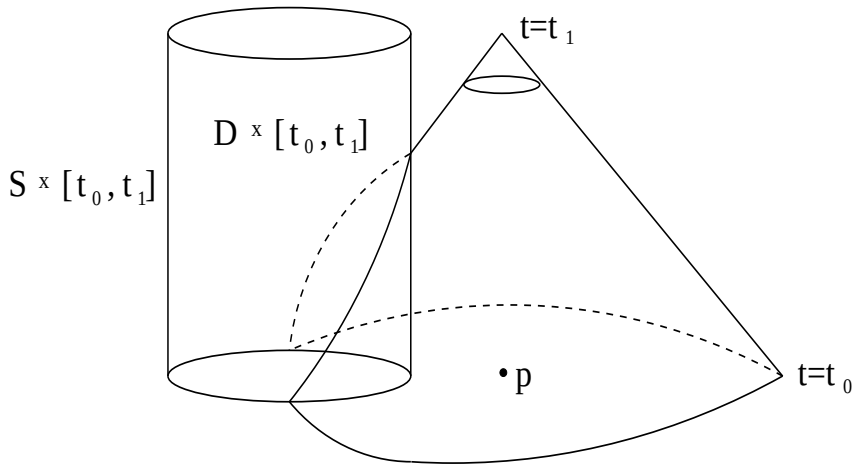


Fig. 5. "Curvy-distance" cone

on $E_{t_1-t_0}(p)$ *and* h *is zero on*

$$\{(x,t) : x \in S, t_0 \le t \le t_1, d(x,p) \le t_1 - t\},$$

then $u(p,t)$ *and* $u_t(p,t)$ *are zero for all* $t \in [t_0, t_1)$ *(see Fig. 5).*

E. Further injectivity results by PDE means

We will show now how simple tools similar to the Propositions 14 and 15, namely finite speed of propagation and domain of dependence for the wave equation allow one to obtain more results concerning geometry of non-injectivity sets, as well as to re-prove some known results with much simpler means. Our final goals with this approach were to recover the full result of [3] in $2D$ and to prove its analogs in higher dimensions and for other geometries (e.g., hyperbolic one) using these simple means. Albeit this goal has not been completely achieved yet, we can report some progress in all these directions.

Let us start with some initial remarks that will narrow the cases we need to consider. First of all, one can assume functions f are as smooth as we wish, since convolution with smooth radial mollifiers does not change the fact that $R_S f = 0$ (e.g., [3]). Secondly, according to the results mentioned before, any non-injectivity set S in the class of compactly supported functions is contained in an algebraic surface that is also a non-injectivity set. It is rather straightforward to show that the same is true for functions that decay exponentially. Thus, **considering only exponentially decaying functions, one does not restrict generality by assuming from the start algebraicity of** S. It is known [1] that algebraic surfaces of co-dimension higher than 1 are automatically non-injectivity sets. Thus, we can restrict our attention to algebraic hypersurfaces S of $\mathbb{R}^n$ only. Any set that is not algebraic (or rather, is not a part of such an algebraic surface) is automatically an injectivity set. So, when trying to obtain necessary conditions for non-injectivity, confining ourselves to the case of algebraic hypersurfaces solely we do not lose any generality. One can also assume irreducibility of that surface, if this helps. When needed, one can also exclude the case of closed hypersurfaces, since according to Corollary 7 those are all injectivity sets.

Our goal now is to exclude some pairs (S, f), where S is an algebraic surface and f is a non-zero function as possible candidates for satisfying the non-injectivity condition $R_S f = 0$. We will do this in terms of geometry of the support of function f. Notice that Theorem 13 does exactly that when S contains an open part of the boundary of a smooth strictly convex domain where f is supported. Theorem 9, on the other hand excludes all compactly supported f's in $\mathbb{R}^2$, unless $S = \omega\Sigma_N$. Similarly, Theorem 12 excludes boundaries S of bounded domains when f is in an appropriate space $L_p(\mathbb{R}^n)$.

Let S be an algebraic hypersurface (which can be assumed to be irreducible if

needed) that splits $\mathbb{R}^n$ into connected parts H^j, $j = 1, ..., m$. One can define the interior metric in H^j as follows:

$$d^j(p, q) = inf\{\text{length of}\,\gamma\}, \tag{2.4}$$

where the infimum is taken over all C^1-curves γ in H^j joining points $p, q \in H^j$.

Theorem 16. *Let S and H^j be as above and $f \in C(\mathbb{R}^n)$ be such that $R_S f = 0$. Let also $x \in \bar{H^j}$, where $\bar{H^j}$ is the closure of H^j. Then*

$$\begin{aligned} dist(x, supp\ f \cap H^j) &= dist^j(x, supp\ f \cap H^j) \\ &\leq dist(x, supp\ f \cap H^k),\ k \neq j, \end{aligned} \tag{2.5}$$

where distances $dist^j$ are computed with respect to the metrics d^j, while dist is computed with respect to the Euclidean metric in $\mathbb{R}^n$.

In particular, for $x \in S$ and any j

$$dist(x, supp\ f \cap H^j) = dist^j(x, supp\ f \cap H^j) = dist(x, supp\ f). \tag{2.6}$$

Thus, the expressions in (2.6) in fact do not depend on $j = 1, ..., m$.

Remark 17. *Notice that under the condition of algebraicity of S the theorem does not require the function f to be compactly supported and in fact imposes no condition on behavior of f at infinity. On the other hand, as it has been mentioned before, if f decays exponentially, then the algebraicity assumption does not restrict the generality of consideration.*

Proof of the theorem. Notice first of all, that the function $d^j(p, x)$ has gradient $|\nabla_x d^j(p, x)| \leq 1$ a.e.[4]

[4] In order to justify legality of the calculation presented below, one can either use geometric measure theory tools, as in [23], or just notice that due to algebraicity of S, the function $d^j(p, x)$ is piece-wise analytic.

Let us prove now the equality

$$dist(x, \text{supp } f \cap H^j) = dist^j(x, \text{supp } f \cap H^j). \tag{2.7}$$

Since $d^j(p,q) \geq |p-q|$, it is sufficient to prove that the left hand side expression cannot be strictly smaller than the one on the right. Assume the opposite, that

$$dist(x, \text{supp } f \cap H^j) = d_1 < d_2 = dist^j(x, \text{supp } f \cap H^j). \tag{2.8}$$

Pick a smaller segment $[d_3, d_4] \subset (d_1, d_2)$. Then, by continuity, for any point p in a small ball $B \subset H^j$ near x (not necessarily containing x, for instance when $x \in S$) one has

$$dist(p, \text{supp } f \cap H^j) \leq d_3 < d_4 \leq dist^j(p, \text{supp } f \cap H^j). \tag{2.9}$$

For such a point p, consider the volume V in the space-time region $H^j \times \mathbb{R}$ bounded by the space-like surfaces Σ_1 given by $t = 0$ and Σ_2 described as $t = \phi(x) = \tau - d^j(p,x)$, and the "vertical" boundary $S \times \mathbb{R}$. Here $\tau \leq (d_3 + d_4)/2$. Consider the solution $u(x,t)$ of the wave equation problem (2.2)–(2.3) with the initial velocity f. Then, by construction, this solution and its time derivative are equal to zero at the lower boundary $t = 0$ and on the lateral boundary $S \times \mathbb{R}$. Hence, by the standard energy computation (integrating the equality $u \Box u = 0$, see, e.g., Section 2.7, Ch. 1 in [14]) we conclude that $u = 0$ in V. For the reader's convenience, let us provide brief details of the corresponding calculations from [14]: Since $\Box u = 0$, $u = u_t = 0$ on Σ_1, and $u|_S = 0$ for all times, we get by integration by parts

$$\begin{aligned} 0 = \int\limits_V u_t \Box u dx dt = \int\limits_{t=\phi(x)} \tfrac{1}{2} \left(|\nabla u|^2 + u_t^2 + 2 u_t \nabla \phi \cdot \nabla u \right) dx \\ = \tfrac{1}{2} \int\limits_{\phi(x) \geq 0} \left(|\nabla (u(x, \phi(x))|^2 + (1 - |\nabla \phi|^2) u_t(x, \phi(x))^2 \right) dx. \end{aligned} \tag{2.10}$$

Since $|\nabla\phi| \leq 1$, we conclude that

$$\int\limits_{\phi(x)\geq 0} \left(|\nabla(u(x,\phi(x))|^2\right) dx = 0$$

and hence u is constant on Σ_2. Taking into the account the zero conditions on S and Σ_1, one concludes that $u = 0$ on Σ_2, and hence in V.

In particular, $u(p,t) = 0$ for all $p \in B$ and $|t| \leq (d_3 + d_4)/2$. Notice that $(d_3 + d_4)/2 > d_3$. Now applying Proposition 14 to the wave equation in the whole space, we conclude that

$$dist(p, \text{supp } f) > d_3, \tag{2.11}$$

and hence

$$dist(p, \text{supp } f \cap H^j) > d_3, \tag{2.12}$$

which is a contradiction. This proves (2.7).

It is now sufficient to prove

$$dist(x, \text{supp } f \cap H^j) \leq dist(x, \text{supp } f \cap H^k) \tag{2.13}$$

for $k \neq j$. This in fact is an immediate consequence of (2.11). Alternatively, we can repeat the same consideration as above in a simplified version. Namely, suppose that

$$dist(x, \text{supp } f \cap H^j) > d_2 > d_1 > dist(x, \text{supp } f \cap H^k) \tag{2.14}$$

for a point $x \in H^j \cap S$, and hence for all points p in a small ball in H^j. Consider the volume V in the space-time region $H^j \times \mathbb{R}$ bounded by the space-like surfaces $t = 0$ and $t = d_2 - |x - p|$ (p fixed in the small ball) and the boundary $S \times \mathbb{R}$. Consider the solution $u(x,t)$ of the wave equation problem (2.2)-(2.3) with the initial velocity f. Then, by construction, this solution and its time derivative are equal to zero at the lower boundary $t = 0$ and on the lateral boundary $S \times \mathbb{R}$. Hence, by the same

standard domain of dependence argument (see, e.g., Section 2.7, Ch. 1 in [14]) we conclude that $u = 0$ in V. In particular, $u(p,t) = 0$ for all $p \in B$ and $|t| \leq d_2$. Now applying Proposition 14 to the wave equation in the whole space, we conclude that

$$dist(p, \text{supp } f) > d_2,$$

and hence

$$dist(p, \text{supp } f \cap H^k) > d_2, \tag{2.15}$$

which is a contradiction. ■

We will now show several corollaries that can be extracted from Theorem 16.

Corollary 18. *Let f be continuous and $S \subset \mathbb{R}^n$ be an algebraic hypersurface such that $R_S f = 0$. Let L be any hyperplane such that $L \cap$ supp $f \neq \emptyset$ and such that supp f lies on one side of L. Let $x \in L \cap$ supp f and r_x be the open ray starting at x, perpendicular to L, and going into the direction opposite to the support of f. Then either $r_x \subset S$ (and hence, the whole line containing r_x belongs to S), or r_x does not intersect S.*

Proof. Assuming otherwise, let $p \in r_x \cap S$ and H^j be the connected components of $\mathbb{R}^n \backslash S$ such that p belongs to their closures. Since x is the only closest point to p in the support of f, Theorem 16 implies that for any j there exist paths t_ϵ joining x and p through H^j and such that the length of t_ϵ tends to $|x - p|$ when $\epsilon \to 0$. This means that these paths converge to the linear segment $[x, p]$. Hence, this segment belongs to H^j for any j, and thus to $\bigcap\limits_j H^j$, which is a part of S. We conclude that the segment $[x, p]$, and then, due to algebraicity of S, the whole its line belongs to S. This proves the statement of the corollary. ■

One notices that a similar proof establishes the following

Corollary 19. *Let f be continuous and $S \subset \mathbb{R}^n$ be an algebraic hypersurface such that $R_S f = 0$. Suppose $p \in S$ is such that p does not belong to supp f and there exists unique point x in supp f closest to p. Then S contains the whole line passing through the points x and p.*

Let $S \subset \mathbb{R}^n$. For any points $p, q \in \mathbb{R}^n - S$ we define the distance $d_S(p, q)$ as the infimum of lengths of C^1 paths in $\mathbb{R}^n - S$ connecting these points. Clearly $d_S(p, q) \geq |p - q|$. Using this metric, we can define the corresponding distances $dist_S$ from points to sets.

Theorem 20. *Let a set $S \subset \mathbb{R}^n$ and a non-zero function $f \in C(\mathbb{R}^n)$ exponentially decaying at infinity be such that $R_S f = 0$. Then for any point $p \in \mathbb{R}^n - S$*

$$dist_S(p, suppf) = dist(p, suppf). \tag{2.16}$$

The same conclusion holds for any continuous function, if one assumes that S is an algebraic hypersurface.

Proof. Assume that (2.16) does not hold, i.e.

$$dist_S(p, \mathrm{supp} f) > dist(p, \mathrm{supp} f).$$

As it has been mentioned before, under the conditions of the theorem, we can assume S to be a part of an algebraic surface Σ for which $R_\Sigma f = 0$. Let Σ divide the space into parts H^j. Then, in notations of the previous theorem, we have

$$dist^j(p, \mathrm{supp}\ f \cap H^j) \geq dist_S(p, \mathrm{supp} f) \tag{2.17}$$

and hence

$$dist^j(p, \mathrm{supp}\ f \cap H^j) > dist(p, \mathrm{supp} f). \tag{2.18}$$

This, however, contradicts Theorem 16. ■

Let us formulate another example of a geometric constraint on pairs S, f such that $R_S f = 0$.[5]

Theorem 21. *Let $S \subset \mathbb{R}^n$ be a relatively open piece of a C^1-hypersurface and $f \in C_c(\mathbb{R}^n)$ be such that $R_S f = 0$. If there is a point $p_0 \in S$ such that the support of f lies strictly on one side of the tangent plane $T_{p_0}S$ to S at p_0, then $f = 0$.*[6]

Proof of the theorem. Let us denote by $K_p(\text{supp } f)$ the convex cone with the vertex p consisting of all the rays starting at p and passing through the convex hull of the support of f. Then $K_{p_0}(\text{supp } f)$, due to the condition of the theorem, lies on one side of $T_{p_0}S$, touching it only at the point p_0. Let us pull the point p_0 to the other side of the tangent plane along the normal to a nearby position p. Then it is easy to see that for p sufficiently close to p_0, all rays of the cone $K_p(\text{supp } f)$ will intersect S. This means in particular, that for this point p we have $dist_S(p, \text{supp } f) > dist(p, \text{supp } f)$. According to Theorem 20, this implies that $f = 0$. ■

Corollary 22. *Let $S \subset \mathbb{R}^n$ be an algebraic hypersurface and $f \in C_c(\mathbb{R}^n)$. If $R_S f = 0$, then every tangent plane to S intersects the convex hull of the support of f.*

The above results present significant restrictions on the geometry of the non-injectivity sets S and of the supports of functions f in the kernel of R_S. One can draw more specific conclusions about these sets. The statement below was proven in [3] by using the geometry of zeros of harmonic polynomials, which we avoid.

Proposition 23. *Let $S \subset \mathbb{R}^2$ be an algebraic curve such that $R_S f = 0$ for some non-zero compactly supported continuous function f. Then S has no compact components, and each its component has asymptotes at infinity.*

---

[5] A similar statement in the case of analytic surfaces S was announced in [51] for distributions f. The proof is claimed to be based upon microlocal analysis.

[6] This implies, in particular, Theorem 13.

Proof. Corollary 7 excludes bounded components. So, we can think that S is an irreducible unbounded algebraic curve. Existence of its asymptotes can be shown as follows. Let us take a point $p \in S$ and send it to one of the infinite ends of S. According to Corollary 22, every tangent line T_pS intersects the convex hull of the support of f, which is a fixed compact in $\mathbb{R}^2$. This makes this set of lines on the plane compact. Hence, we can choose a sequence of points p_j such that the lines $T_{p_j}S$ converge to a line T in the natural topology of the space of lines (e.g., one can use normal coordinates of lines to introduce such topology). This line T is in fact the required asymptote. Indeed, let us choose the coordinate system where T is the x-axis. Then the slopes of the sequence $T_{p_j}S$ converge to zero. Due to algebraicity, for a tail of this sequence, the convergence is monotonic, and in particular holds for all $p \in S$ far in the tail of S. Let us for instance assume that these slopes are negative. Then the tail of S is the graph of a monotonically decreasing positive function. This means that S has a horizontal asymptote. This asymptote must be the x-axis T, otherwise the y-intercepts of $T_{p_j}S$ would not converge to zero, which would contradict the convergence of $T_{p_j}S$ to T. ■

The next statement proves the Agranovsky-Quinto Theorem 9 in some particular cases. In order to formulate it, we need to introduce the following condition:

Condition A. Let K be a compact subset of $\mathbb{R}^n$. We will say that the boundary of K satisfies **condition A**[7], if there exists a positive number r_0 such that for any $r < r_0$ and any point x in the infinite connected component of $\mathbb{R}^n \setminus K$ such that dist $(x, K) = r$ there exists a unique point k on K such that $|x - k| = r$.

Examples of such sets are convex sets (where $r_0 > 0$ can be chosen arbitrarily) and sets with a C^2 boundary (where r_0 should be sufficiently small).

---

[7] This condition essentially restricts the curvature of the boundary from below.

Theorem 24. *Let $S \subset \mathbb{R}^2$ and $f(\neq 0) \in C_c(\mathbb{R}^2)$ be such that $R_S f = 0$. If the external boundary of the support of f (i.e., the boundary of the infinite component of the complement of the support) is connected and satisfies Condition A, then $S \subset \omega\Sigma_N \cup F$ in notations of Theorem 9.*

The conditions of the theorem are satisfied for instance when the support of f contains the boundary of its convex hull, or when the support's external boundary is connected and of the class C^2.

Proof. First of all, up to a finite set, we can assume that S is an algebraic curve. Since the external boundary of the support is assumed to be connected, Theorem 16 implies that any irreducible component of S must meet any neighborhood of the support of f. If we take the neighborhood of radius $r < r_0$, then each point on S in this neighborhood will have a unique closest point on supp f. Applying now Corollary 19, we conclude that S consists of straight lines L_j intersecting the support. It is known that any straight line L is a non-injectivity set, but the only functions annihilated by R_L are the ones odd with respect to L (e.g., [3, 17, 35]). Hence, f is odd with respect to all lines L_j. In particular, every of these lines passes through the center of mass of the support of f. Hence, lines L_j form a "cross"[8]. It remains now to show that the angles between the lines are commensurate with π. This can also be shown in several different ways. For instance, this follows immediately from existence of a **harmonic** polynomial vanishing on S. Another simple option is to notice that if this is not the case, then there is no non-zero function that is odd simultaneously with respect to all the lines. ■

[8] One can prove that all these lines pass through a joint point also in a different manner. Indeed, due to oddness of f, each line is a symmetry axis for the support of f. Then, considering the group generated by reflections through these lines, one can easily conclude that if they did not pass through a joint point, then the support of f must have been non-compact.

Exactly the same consideration as above shows that in higher dimensions the following statement holds:

Proposition 25. *Let $S \subset \mathbb{R}^n$ and $f(\neq 0) \in C_c(\mathbb{R}^n)$ be such that $R_S f = 0$. If the external boundary of the support of f (i.e., the boundary of the infinite component of the complement of the support) is connected and satisfies Condition A, then S is ruled* [9].

The conditions of the theorem are satisfied for instance when the support of f contains the boundary of its convex hull, or when the support's external boundary is connected and of the class C^2.

Remark 26. *If we could also show that all these lines pass through the same point, then this would immediately imply, as in the previous proof, the validity of Conjecture 10 for this particular case.*

F. Remarks

1. M. Agranovsky and E. T. Quinto have written besides [3], several other papers devoted to the problem considered here. They consider some partial cases (e.g., distributions f supported on a finite set) and variations of the problem (e.g., in bounded domains rather than the whole space). See [1, 4, 5, 6] for details.

2. One of our goals was to obtain the complete Theorem 9, the main result of [3] by simple PDE tools, avoiding using the geometry of zeros of harmonic polynomials and microlocal analysis (or at least one of those), as well as to prove

[9] A **ruled surface** is a union of a family of lines (e.g., [77])

its analogs in higher dimensions and for other geometries (e.g., hyperbolic one). Although we have not completely succeeded in this yet, the results presented (e.g., Propositions 23 and 25 and Theorem 24) are moving in this direction.

3. The PDE methods presented here in principle bear a potential for considering non-compactly-supported functions. In order to achieve this, one needs to have qualitative versions of statements like Proposition 15 and Theorem 20, where instead of just noticing whether a wave has come to certain point at a certain moment (which was our only tool), one controls the amount of energy it carries.

4. One of the motivations for studying the injectivity problem was the thermoacoustic tomography. One wonders then if considerations of $2D$ problems (rather than $3D$ ones) bear any relevance for TAT. In fact, they do. If either the scanned sample is very thin, or the transducers are collimated in such a way that they register the signals only coming parallel to a given plane, one arrives to a $2D$ problem.

5. A closer inspection of the results of the previous section shows that most of them have their local versions, where it is not required that the whole transform R_S of a function vanishes, but rather only for radii up to a certain value. One can see an example of a local uniqueness theorem for the circular transform in [51].

6. As J. Boman notified us during the April 2004 AMS meeting in Lawrenceville, he jointly with J. Sjostrand, being unaware of our work, had recently independently obtained some results analogous to some of those presented here (e.g., to Theorem 20).

7. We have not touched the problem of finding explicit inversion formulas for the circular transforms. Such formulas are known for the spherical, planar, and

cylindrical sets of centers [12, 19, 21, 23, 58, 65, 79, 80, 81]. They come in two kinds: the ones involving expansions into special functions, and the ones of backprojection type. We use an expansion type inversion formula in Chapter III to derive range conditions in spherical geometry. Exact backprojection type formulas are known for the planar geometry [19, 65] and recently for the spherical geometry in odd dimensions [23] if the function to be reconstructed is supported inside the sphere of transducers. The latter are used in Chapter IV for some numerical reconstructions.

CHAPTER III

RANGE CONDITIONS[1]

A. Introduction

As it is common for transforms of Radon type, the range of the circular Radon transform has infinite co-dimension in standard function spaces. Range descriptions for such transforms are known to be very important for computed tomography, for instance when dealing with incomplete data, error correction, and other issues. In this chapter we will give the complete range description for the circular Radon transform and will be dealing with the planar case only. Due to tomographic applications, where S is the set of locations of transducers [38, 79, 80, 82], we will be from now on looking at the specific case when S is the unit circle $|x| = 1$ in the plane. Moreover, we will be dealing with functions supported inside the circle S only. The properties of the operator R_S (e.g., stability of the inversion, its FIO properties, etc.) deteriorate on functions with supports extending outside S (e.g., [3, 23, 82]). However, in tomographic applications one normally deals with functions supported inside S only [38, 68, 79, 82].

As it has already been mentioned, the range of R_S has infinite co-dimension (e.g., in spaces of smooth functions, see details below) and thus infinitely many range conditions appear. It seems to be a rather standard situation for various types of Radon transforms that range conditions split into two types, one of which is usually easier to discover, while another "half" is harder to come by. For instance, it took

[1] Part of this chapter is reprinted with permission from "A range description for the planar circular Radon transform", by G. Ambartsoumian and P. Kuchment, to appear in the SIAM Journal on Mathematical Analysis. Copyright ©2006 by Society for Industrial and Applied Mathematics.

about a decade to find the complete range description for the so called exponential Radon transform arising in SPECT (single photon emission computed tomography) [7, 8, 45, 46, 78]. For a more general attenuated transform arising in SPECT, it took twice as much time to move from a partial set of range conditions [55, 56] to the complete set [64]. In the circular case, a partial set of such conditions was discovered recently [68]. It happens to be incomplete, and the goal of this chapter is to find the complete one.

One might ask why is it important to know the range conditions. These conditions have been used extensively in tomography (as well as in radiation therapy planning, e.g. [15, 16, 42, 71]) for various purposes: completing incomplete data, detecting and correcting measurement errors and hardware imperfections, recovering unknown attenuation, etc. [33, 52, 53, 54, 55, 60, 61, 69, 75, 76]. Thus, as soon as a new Radon type transform arises in an application, a quest for the range description begins.

In order to explain our approach, we start in the next section with treating a toy example of the standard Radon transform on the plane, where the range conditions are well known (e.g., [20, 25, 26, 27, 32, 55, 57]). Our approach, however, is different from the standard ones and naturally leads to the considerations of the circular transform in the rest of the paper.

B. The case of the planar Radon transform

In this section we will approach in a somewhat non-standard way the issue of the range description for the standard Radon transform on the plane. Consider a compactly

supported smooth function $f(x)$ on the plane and its Radon transform

$$(\mathcal{R}f)(\omega, s) = g(\omega, s) := \int\limits_{x\cdot\omega=s} f(x)dl, \tag{3.1}$$

where $s \in \mathbb{R}$, $\omega \in S^1$ is a unit vector in $\mathbb{R}^2$, and dl is the arc length measure on the line $x \cdot \omega = s$. We want to describe the range of this transform, say on the space $C_0^\infty(\mathbb{R}^2)$. Such a description is well known (e.g., [20, 25, 26, 27, 32, 55, 57], or any other book or survey on Radon transforms or computed tomography):

Theorem 27. *A function g belongs to the range of the Radon transform on C_0^∞ if and only if the following conditions are satisfied:*

1. $g \in C_0^\infty(S^1 \times \mathbb{R})$,
2. *for any $k \in \mathbb{Z}^+$ the k-th moment $G_k(\omega) = \int\limits_{-\infty}^{\infty} s^k g(\omega, s)ds$ is the restriction to the unit circle S^1 of a homogeneous polynomial of ω of degree k,*
3. $g(\omega, s) = g(-\omega, -s)$.

We would like to look at this result from a little bit different prospective, which will allow us to do a similar thing in the case of the circular Radon transform.

In order to do so, let us expand $g(\omega, s)$ into the Fourier series with respect to the polar angle ψ (i.e., $\omega = (\cos\psi, \sin\psi)$)

$$g(\omega, s) = \sum_{n=-\infty}^{\infty} g_n(s)e^{in\psi}. \tag{3.2}$$

We can now reformulate the last theorem in the following a little bit strange way:

Theorem 28. *A function g belongs to the range of the Radon transform on C_0^∞ if and only if the following conditions are satisfied:*

1. $g \in C_0^\infty(S^1 \times \mathbb{R})$,

2. *for any n, the Mellin transform $Mg_n(\sigma) = \int_0^\infty s^{\sigma-1} g_n(s)ds$ of the n-th Fourier coefficient g_n of g vanishes at any pole σ of the function $\Gamma(\frac{\sigma+1-|n|}{2})$,*

3. *$g(\omega, s) = g(-\omega, -s)$.*

Since the only difference in the statements of these two theorems is in the conditions 2, let us check that these conditions mean the same thing in both cases. Indeed, let us expand $g(\omega, s)$ into Fourier series (3.2) with respect to ψ. Representing $e^{in\psi}$ as the homogeneous polynomial $(\omega_1 + i(sign\, n)\omega_2)^{|n|}$ of ω of degree $|n|$, and noticing that $\omega_1^2 + \omega_2^2 = 1$ on the unit circle, one easily concludes that the condition 2 in Theorem 27 is equivalent to the following: the k-th moment $\int_{\mathbb{R}} s^k g_n(s)ds$ of the n-th Fourier coefficient vanishes for integers $0 \leq k < |n|$ such that $k - n$ is even.

Let us now look at the condition 2 in Theorem 28, still using the same Fourier expansion. Notice that when $k - |n|$ is a negative even integer, $Mg_n(\sigma)$ is one-half of the moment of order $k = \sigma - 1$ of g_n. Taking into account that $\Gamma(\frac{\sigma+1-|n|}{2}) = \Gamma(\frac{k+2-|n|}{2})$ has poles exactly when $k - |n|$ is a negative even integer, we see that conditions 2 in both theorems are in fact saying the same thing.

One can now ask the question, why should one disguise in the statement of Theorem 28 negative integers as poles of Gamma-function and usual moments as values of Mellin transforms? The answer is that in the less invariant and thus more complex situation of the circular Radon transform, one can formulate a range description in the spirit of Theorem 28, albeit it is unclear how to get an analog of the version given in Theorem 27.

As a warm-up, let is derive the condition 2 in Theorem 28 directly, without relying on the version given in the preceding theorem. This is in fact an easy by-product of the A. Cormack's inversion procedure, see e.g. [57, Section II.2]. Indeed, if we write down the original function $f(x)$ in polar coordinates $r(\cos\phi, \sin\phi)$ and

expand into the Fourier series with respect to the polar angle ϕ

$$f(r(\cos\phi, \sin\phi)) = \sum_{n=-\infty}^{\infty} f_n(r)e^{in\phi}, \tag{3.3}$$

then the Fourier coefficients f_n and g_n of the original and of its Radon transform are related as follows [55, formula (2.17) and further]:

$$M(rf_n(r))(s) = \frac{(Mg_n)(s)}{B_n(s)}, \tag{3.4}$$

where

$$B_n(s) = const\frac{\Gamma(s)2^{-s}}{\Gamma((s+1+|n|)/2)\Gamma((s+1-|n|)/2)} \tag{3.5}$$

Thus, condition 2 of Theorem 28 guarantees that the function $M(rf_n(r))(s)$ does not develop singularities (which it cannot do for a C_0^∞-function f) at zeros of $B_n(s)$. It is not that hard now to prove also sufficiency in the theorem, applying Cormack's inversion procedure to g satisfying conditions 1 - 3. However, we are not going to do so, since in the next sections we will devote ourselves to doing similar thing in the more complicated situation of the circular Radon transform.

C. The circular Radon transform. Formulation of the main result

Let us recall the notion of Hankel transform (e.g., [18]). For a function $h(r)$ on $\mathbb{R}^+$, one defines its Hankel transform of an integer order n as follows:

$$(\mathcal{H}_n h)(\sigma) = \int_0^\infty J_n(\sigma r)h(r)r\,dr, \tag{3.6}$$

where the standard notation J_n is used for Bessel functions.

Let, as in the Introduction, R_S be the circular Radon transform on the plane that integrates functions compactly supported inside the unit disk D over all circles

$|x-p| = \rho$ with centers p located on the unit circle $S = \{p \,|\, |p| = 1\}$. Since this transform commutes with rotations about the origin, the Fourier series expansion with respect to the polar angle partially diagonalizes the operator, and thus the n-th Fourier coefficient $g_n(\rho)$ of $g = R_S f$ will depend on the n-th coefficient f_n of the original f only. It was shown in [62] that the following relation between these coefficients holds:

$$g_n(\rho) = 2\pi\rho\mathcal{H}_0\{J_n\mathcal{H}_n\{f_n\}\}. \tag{3.7}$$

For the reader's convenience, we will provide the brief derivation from [62]. Considering a single harmonic $f = f_n(r)e^{in\phi}$ and using polar coordinates, one obtains

$$g_n(\rho) = \int_0^\infty r f_n(r) dr \int_0^{2\pi} \delta\left[(r^2+1-2r\cos\phi)^{1/2} - \rho\right] e^{-in\phi} d\phi. \tag{3.8}$$

Thus, the computation boils down to evaluating the integral

$$I = \int_0^{2\pi} \delta\left[(r^2+1-2r\cos\phi)^{1/2} - \rho\right] e^{-in\phi} d\phi.$$

Using the standard identity

$$\delta(\rho' - \rho) = \rho \int_0^\infty J_0(\rho' z) J_0(\rho z) z dz$$

and

$$2\pi J_n(az) J_n(bz) = \int_0^{2\pi} J_0[z(a^2+b^2-2ab\cos\phi)^{1/2}] e^{-in\phi} d\phi,$$

one arrives from (3.8) to (3.7).

Since Hankel transforms are involutive, it is easy to invert (3.7) and get Norton's inversion formulae [62]

$$f_n = \frac{1}{2\pi}\mathcal{H}_n\{\frac{\mathcal{H}_0\{g_n(\rho)/\rho\}}{J_n}\}. \tag{3.9}$$

Now one can clearly see analogies with the case of the Radon transform, where zeros of Bessel functions should probably introduce some range conditions. This happens to be correct and leads to the main result of this article:

Theorem 29. *In order for the function $g(p,\rho)$ on $S^1\times\mathbb{R}$ to be representable as $R_S f$ with $f\in C_0^\infty(D)$, it is necessary and sufficient that the following conditions are satisfied:*

1. *$g\in C_0^\infty(S^1\times(0,2))$.*
2. *For any n, the $2k$-th moment $\int\limits_0^\infty \rho^{2k} g_n(\rho)d\rho$ of the n-th Fourier coefficient of g vanishes for integers $0\le k<|n|$. (Equivalently, the $2k$-th moment $\int\limits_0^\infty \rho^{2k} g(p,\rho)d\rho$ is the restriction to the unit circle S of a (non-homogeneous) polynomial of p of degree at most k.)*
3. *For any $n\in\mathbb{Z}$, function $\mathcal{H}_0\{g_n(\rho)/\rho\}(\sigma)=\int\limits_0^\infty J_0(\sigma\rho)g_n(\rho)d\rho$ vanishes at any zero $\sigma\neq 0$ of Bessel function J_n. (Equivalently, the nth Fourier coefficient with respect to $p\in S^1$ of the "Bessel moment" $G_\sigma(p)=\int\limits_0^\infty J_0(\sigma\rho)g(p,\rho)d\rho$ vanishes if $\sigma\neq 0$ is a zero of Bessel function J_n.)*

D. Proof of the main result

Let us start with proving necessity, which is rather straightforward. Indeed, the necessity of condition 1 is obvious. Let us prove the second condition. In fact, it has already been established in [68]. Let us repeat for completeness its simple proof. Let k be an integer. Consider the moment of order $2k$ of g:

$$\int\limits_0^\infty \rho^{2k} g(p,\rho)d\rho=\int\limits_{\mathbb{R}^2}|x-p|^{2k}f(x)dx=\int\limits_{\mathbb{R}^2}(|x|^2-2x\cdot p+1)^k f(x)dx \tag{3.10}$$

(we have taken into account that $|p| = 1$). We see that the resulting expression is the restriction to S^1 of a (non-homogeneous) polynomial of degree k in variable p. Expanding into Fourier series with respect to the polar angle of p, we see that the nth harmonic g_n contributes the following homogeneous polynomial of degree $|n|$ in the variable p:

$$\left(\int_0^\infty \rho^{2k} g_n(\rho) d\rho \right) e^{in\phi}.$$

Here as before $p = (\cos\phi, \sin\phi)$. Thus, for $|n| > k$, this term must vanish, which gives necessity of condition 2. We will return to a discussion of this condition below to add a new twist to it.

Necessity of condition 3 follows immediately from Norton's formula (3.9), which implies in particular that

$$\mathcal{H}_0\{g_n(\rho)/\rho\} = 2\pi J_n \mathcal{H}_n\{f_n\}.$$

Since both functions J_n and $\mathcal{H}_n\{f_n\}$ are entire, $\mathcal{H}_0\{g_n(\rho)/\rho\}$ vanishes whenever J_n does.

Remark 30. *The reader might ask why in the third condition of the Theorem we do not take into account the zero root of J_n, which in fact has order n, while non-zero roots are all simple. The reason is that the condition 2 already guarantees that $\sigma = 0$ is zero of order $2n$ of $\mathcal{H}_0\{g_n(\rho)/\rho\}$ (twice higher than that of J_n). Indeed, due to evenness of J_0, function $\mathcal{H}_0\{g_n(\rho)/\rho\}(\sigma)$ is also even. Thus, all odd order derivatives at $\sigma = 0$ vanish. The known Taylor expansion of J_0 at zero leads to the formula*

$$\mathcal{H}_0\{g_n(\rho)/\rho\}(\sigma) = \sum_m \frac{(-1)^m}{(m!)^2} \left(\frac{\sigma}{2}\right)^{2m} \int_0^\infty r^{2m} g_n(r) dr.$$

We see now that the moment condition 2 guarantees that $\sigma = 0$ is zero of order $2n$ of $\mathcal{H}_0\{g_n/\rho\}(\sigma)$.

Let us move to the harder part, proving sufficiency. Assume a function g satisfies conditions of the theorem and is supported in $S \times (\epsilon, 2-\epsilon)$ for some positive ϵ. We will show that then $g = R_S f$ for some $f \in C_0^\infty(D_\epsilon)$, where D_ϵ is the disk $|x| < 1-\epsilon$ in the plane.

Due to Norton's formulas, it is natural to expect the proof to go along the following lines: expand g into the Fourier series $g = \sum_m g_m(\rho) e^{im\psi}$ with respect to the angle variable ψ, then use inversion formula (3.9) to construct a function f and then show that f is of an appropriate function class and that its circular Radon transform is equal to g. This is what we are going to do, with a small caveat that instead of constructing f itself, we will construct its two-dimensional Fourier transform. Besides, we will start considering the partial sums of the series $h_n = \sum_{|m|\le n} g_m(\rho) e^{im\phi}$. But first, we need to get some simple estimates from below for the Bessel function of the first kind J_n.

Lemma 31. *On the entire complex plane except for a disk S_0 centered at the origin and a countable number of disks S_k of radii $\pi/6$ centered at points $\pi(k + \frac{2n+3}{4})$, one has*

$$|J_n(z)| \ge \frac{Ce^{|Im\, z|}}{\sqrt{|z|}}, \qquad C > 0 \tag{3.11}$$

Proof: Let us split the complex plane into three parts by a circle S_0 of a radius R (to be chosen later) centered at the origin and a planar strip $\{\sigma = x + iy \,|\, |y| < a\}$, as shown in Fig. 6 below. We will prove the estimate (3.11) separately outside and inside the strip (i.e., in the first and second parts shown in the picture). Using the parity property of function J_n, it is sufficient to consider only the right half plane $Re\, z \ge 0$.

The Bessel function of the first kind $J_n(z)$ has the following known asymptotic

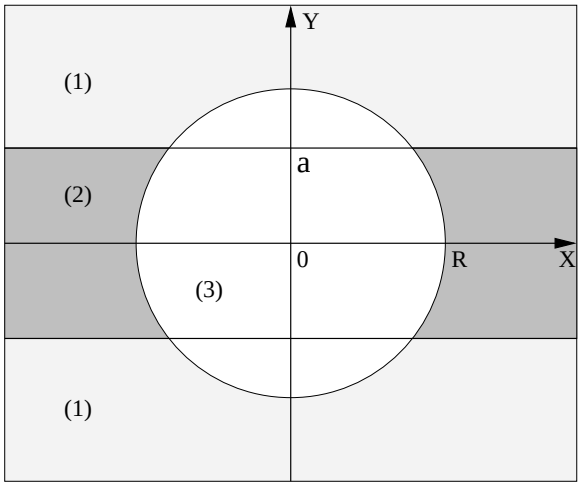


Fig. 6. Bessel estimate

representation in the sector $|\arg z| \le \pi - \delta$ (e.g., [13, formula (4.8.5)] and [48, formula (5.11.6)]):

$$\begin{aligned} J_n(z) = \sqrt{\frac{2}{\pi z}}\cos(z - \frac{\pi n}{2} - \frac{\pi}{4})(1 + O(|z|^{-2})) \\ -\sqrt{\frac{2}{\pi z}}\sin(z - \frac{\pi n}{2} - \frac{\pi}{4})\left(\frac{4n^2-1}{8z} + O(|z|^{-3})\right) \end{aligned} \tag{3.12}$$

Let us start estimating in the first part of the complex plane, i.e. where $|Im\, z| > a$ and $|z| > R$ for sufficiently large a and R (and, as we have agreed, $Re\, z \ge 0$). There, due to boundedness of $\tan z$ in this region, one concludes that $\frac{\sin z}{z} = \cos z\ (O(|z|^{-1}))$, and thus (3.12) implies

$$J_n(z) = \sqrt{\frac{2}{\pi z}}\cos(z - \frac{\pi n}{2} - \frac{\pi}{4})(1 + O(|z|^{-1})),$$

which in turn for sufficiently large a, R leads to

$$|J_n(z)| \ge \frac{Ce^{|Im\, z|}}{\sqrt{|z|}} \tag{3.13}$$

In the second part of the plane (right half of the strip), due to boundedness of

$\sin z$ we have

$$J_n(z) = \sqrt{\frac{2}{\pi z}}\left[\cos(z - \frac{\pi n}{2} - \frac{\pi}{4})(1 + O(|z|^{-2})) + O(|z|^{-1})\right].$$

Consider the system of non-intersecting circles S_k with centers at $z_k = \frac{\pi}{2} + k\pi + \frac{\pi n}{2} + \frac{\pi}{4}$ and radii equal to $\frac{\pi}{6}$. Then outside these circles $|\cos(z - \frac{\pi n}{2} - \frac{\pi}{4})| \geq C$ and

$$|J_n(z)| \geq \frac{C}{\sqrt{|z|}}(1 + O(|z|^{-1})).$$

This implies that for a suitably chosen and sufficiently large R, inside of the strip and outside the circles S_k, we have

$$|J_n(z)| \geq \frac{Ce^{|Im\, z|}}{\sqrt{|z|}} \tag{3.14}$$

for $|z| > R$. This proves the statement of the lemma. ■

Let us now return to our task: consider the function g and the partial sums h_n of its Fourier series.

Lemma 32. *1. If $g(\phi, \rho) = \sum_m g_m(\rho)e^{im\psi}$ satisfies conditions of Theorem 29 and is supported in $S \times (\epsilon, 2-\epsilon)$, then each partial sum $h_n = \sum_{|m|<n} g_m(\rho)e^{im\psi}$ does so.*

2. For any n, h_n is representable as $R_S f_n$ for a function $f_n \in C_0^\infty(D_\epsilon)$.

Proof of the lemma. The first statement of the lemma is obvious.

Due to 1), it is sufficient to prove the second statement for a single term $g = g_n(\rho)e^{in\psi}$. As it was just mentioned, we will reconstruct the Fourier transform F of the function f. In order to do this, we will use the standard relation between Fourier and Hankel transforms. Let as before $f(x) = f_n(r)e^{in\phi}$, where $r = |x|$ and ϕ are polar coordinates on $\mathbb{R}^2$. Then the Fourier transform $F(\xi)$ of f at points of the form $\xi = \sigma\omega$, where $\sigma \in \mathbb{C}$ and $\omega = (\cos\psi, \sin\psi) \in \mathbb{R}^2$ can be written up to a constant

factor as follows:

$$F(\sigma\omega) = \mathcal{H}_n(f_n)(\sigma)e^{in\psi}. \tag{3.15}$$

If we knew that $f = R_S g$, the according to (3.7) this would mean that

$$F(\sigma\omega) = F(\sigma)e^{in\psi} = \frac{1}{2\pi}\frac{\mathcal{H}_0(g_n(\rho)/\rho)(\sigma)}{J_n(\sigma)}e^{in\psi}. \tag{3.16}$$

Let us now take this formula (3.16) as the definition of $F(\sigma\omega)$. Due to the standard parity property of Bessel functions, such F is a correctly defined function of $\sigma\omega$ for $\sigma \neq 0$ (i.e., $F(\sigma\omega) = F((-\sigma)(-\omega))$). We would like to show that it is the Fourier transform of a function $f \in C_0^\infty(D_\epsilon)$. Let us prove first that F belongs to the Schwartz space $\mathcal{S}(\mathbb{R}^2)$. In order to do so, we need to show its smoothness with respect to the angular variable ψ, smoothness and fast decay with all derivatives in the radial variable σ, as well as that no singularity arises at the origin, which in principle could, due to usage of polar coordinates. Smoothness with respect to the angular variable is obvious, due to (3.16). Let us deal with the more complex issue of smoothness and decay with respect to σ. First of all, taking into account that $g_n(\rho)$ is supported inside $(0, 2)$, and due to the standard $2D$ Paley-Wiener theorem, we conclude that $u(\sigma) = \mathcal{H}_0(g_n(\rho)/\rho)$ is an entire function that satisfies for any N the estimate

$$|u(\sigma)| \leq C_N(1 + |\sigma|)^{-N}e^{(2-\epsilon)|Im\,\sigma|}. \tag{3.17}$$

According to the range conditions 2 and 3 of the Theorem, this function vanishes at all zeros of Bessel function $J_n(\sigma)$ at least to the order of the corresponding zero. This means, that function $F(\sigma) = \dfrac{u(\sigma)}{2\pi J_n(\sigma)}$ is entire. Let us show that it belongs to a Paley-Wiener class.

Indeed, $\mathcal{H}(g_n(\rho)/\rho)$ is an entire function with Paley-Wiener estimate (3.17). Due to the estimate from below for J_n (3.11) given in Lemma 31, we conclude that $F(\sigma\omega)$

is an entire function of Paley-Wiener class in the radial directions, uniformly with respect to the polar angle. Namely,

$$|F(\sigma)| \leq C_N(1+|\sigma|)^{-N} e^{(1-\epsilon)|Im\,\sigma|}. \tag{3.18}$$

Indeed, outside the family of circles S_k the estimate (3.11) together with (3.17) give the Paley-Wiener estimate we need. Inside these circles, application of the maximum principle finishes the job. Smoothness with respect to the polar angle is obvious. Thus, the only thing one needs to establish to verify that F belongs to the Schwartz class is that F is smooth at the origin. This, however, is the standard question in the Radon transform theory, the answer to which is well known [25, 26, 27, 32]. Namely, one needs to establish that for any non-negative integer k, the kth radial (i.e., with respect to σ) derivative of $F(\sigma\omega)$ at the origin is a homogeneous polynomial of order k with respect to ω. So, let us check that this condition is satisfied in our situation. First of all, the parity of the function F is he same as of n. Thus, we do not need to worry about the derivatives $F_\sigma^{(k)}|_{\sigma=0}$ with $k-n$ odd, since they are zero automatically. Due to the special single-harmonic form of F, we only need to check that $F_\sigma^{(k)}|_{\sigma=0} = 0$ for $k < |n|$ with $k-n$ even.This, however, as we have discussed already in Remark 30, follows from the moment conditions 2 of the Theorem.

Due to the smoothness that we have just established and Paley-Wiener estimates, $F \in \mathcal{S}(\mathbb{R}^2)$. Thus, $F = \hat{f}$ for some $f \in \mathcal{S}(\mathbb{R}^2)$. It remains to show that f is supported inside the disk D_ϵ. Consider the usual Radon transform $\mathcal{R}f(s,\phi)$ of f. According to the standard Fourier-slice theorem [20, 25, 26, 27, 32, 55], the one-dimensional Fourier transform (denoted by a "hat") from the variable s to σ gives (up to a fixed constant factor) the values $\widehat{\mathcal{R}f}(\sigma,\psi) = F(\sigma\omega)$, if as before $\omega = (\cos\psi, \sin\psi)$. Here $\mathcal{R}$, as before, denotes the standard Radon transform in the plane. Since functions $F(\sigma\omega)$

of σ, as we have just discussed, are uniformly with respect to ω of a Paley-Wiener class, this implies that $\mathcal{R}f(s,\omega)$ has uniformly with respect to ω bounded support in $|s| < 1-\epsilon$. Now the "hole theorem" [32, 55] (which is applicable to functions of the Schwartz class), implies that f has compact support.

The last step is to show that $R_S f = g = g_n(r)e^{in\phi}$. This, however, immediately follows from comparing formulas (3.16) and (3.7), which finishes the proof of the main Lemma 32. ■

Let us now return to the proof of Theorem 29. We have proven so far that any partial sum h_n of the Fourier series for g belongs to the range of the operator R_S acting on smooth functions supported inside the disk D_ϵ. The function g itself is the limit of h_n in $C_0^\infty(S\times(\epsilon,2-\epsilon))$. The only thing that remains to be proven is that the range is closed in an appropriate topology. Microlocal analysis can help with this.

Consider R_S as an operator acting from functions defined on the open unit disk D to functions defined on the open cylinder $\Omega = S\times(0,2)$. As such, it is a Fourier integral operator [28, 29, 70]. If R_S^t is the dual operator, then $E = R_S^t R_S$ is an elliptic pseudo-differential operator of order -1 [28, Theorem 1][2].

Lemma 33. *The continuous linear operator $E : H_0^2(D_\epsilon) \mapsto H_{loc}^3(D)$ has zero kernel and closed image.*

Proof of the lemma. Since $E = R_S^t R_S$, the kernel of this operator coincides with the kernel of R_S acting on $H_0^2(D_\epsilon)$. Since S is closed, it is known that R_S has no compactly supported functions in its kernel [2, 3] (this also follows from analytic ellipticity of E and Theorem 8.5.6 of [34], see also Lemma 4.4 in [3]). Thus, the statement about the kernel is proven and we only need to prove the range closedness.

[2]Bolker's injective immersion condition [28] is satisfied here, as shown in the proof of Lemma 4.3 in [3].

Let P be a properly supported pseudo-differential parametrix of order 1 for E [74]. Then $PE = I + B$, where B is an infinitely smoothing operator on D. Consider the operator Π that acts as the composition of restriction to D_ϵ and then orthogonal projection onto $H_0^2(D_\epsilon)$ in $H^2(D_\epsilon)$. On $H_0^2(D_\epsilon)$ one has $\Pi PE = I + K$, where K is a compact operator on $H_0^2(D_\epsilon)$. Notice that the operator ΠP is continuous from the Frechet space $H^3_{loc}(D)$ to $H_0^2(D_\epsilon)$. Due to the Fredholm structure of the operator $\Pi PE = I + K$ acting on $H_0^2(D_\epsilon)$, its kernel is finite-dimensional. Let $M \subset H_0^2(D_\epsilon)$ be a closed subspace of finite codimension complementary to the kernel, so $I + K$ is injective on M and has closed range. Then one can find a bounded operator A in $H_0^2(D_\epsilon)$ such that $A(I + K)$ acts as identity on M. Thus, the operator $A\Pi P$ provides a continuous left inverse to $E : M \mapsto H^3_{loc}(D)$. This shows that the range of E on M is closed in $H^3_{loc}(D)$. On the other hand, the total range of E differs only by a finite dimension from the one on M. Thus, it is also closed. ■

We can now finish the proof of the theorem. Indeed, the last lemma shows that the function $R_S^t g$, being in the closure of the range, is in fact in the range, and thus can be represented as Ef with some $f \in H_0^2(D_\epsilon)$. In other words, $R_S^t(R_S f - g) = 0$. Since the kernel of R_S^t on compactly supported functions is orthogonal to the range of R_S, we conclude that $R_S f - g = 0$. Since $Ef = R_S^t g$ is smooth, due to ellipticity of E we conclude that f is smooth as well. This finishes the proof. ■

E. Remarks

We would like to finish with some remarks.

- It should be possible to prove that the operator R_S in the situation considered in the text is semi-Fredholm between appropriate Sobolev spaces (analogously

to the properties of the standard and of the attenuated Radon transform, e.g. [31, 55]). This would eliminate the necessity of the closedness of the range discussion provided in the end of the proof of Theorem 29.

- We considered the important for tomographic imaging situation when the functions to reconstruct are supported inside the aperture S. What happens when the supports of functions extend outside the circle S? It is known that compactly supported (or even belonging to L_p with sufficiently small p) functions can still be uniquely reconstructed [2, 3]. However, due to standard microlocal reasons [44, 47, 51, 72, 82], then some parts of the wave front set of the function outside S will not be stably recoverable. For instance, nice backprojection type inversion formulas available in odd dimensions [23] fail for such functions.

- Our result is stated and proven in $2D$ only. D. Finch and Rakesh have recently obtained by different methods some range descriptions in odd dimensions [22]. Their method does not apply to even dimensions though.

CHAPTER IV

NUMERICAL RECONSTRUCTIONS

A. Introduction

It is a known fact in tomography that the existence of inversion formulae does not guarantee the stable reconstruction in practice. That is why numerical validation is required for any new inversion formula or algorithm when it becomes available.

In this chapter we will deal only with the case of spherical geometry (i.e. the transducers are located on a unit sphere) and from now on we will assume that $|p| = 1$.

Two different approaches have been used to derive exact inversion formulae for this case. Fourier-Bessel and spherical harmonic expansions result in solutions written as infinite series for two and three dimensions respectively [62, 63]. For 3D the TCT analog of ρ-filtered backprojection inversion was derived in [23]

$$f(x) = -\frac{1}{8\pi^2}\triangle_x \left(\int_{|p|=1} \frac{1}{|x-p|}\, Rf(p, |x-p|)dp \right) \tag{4.1}$$

as well as a filtered backprojection (FBP) type version

$$f(x) = -\frac{1}{8\pi^2} \left(\int_{|p|=1} \frac{1}{|x-p|}\ \ \frac{\partial^2}{\partial r^2} Rf(p, |x-p|)dp \right) \tag{4.2}$$

Both formulas can be generalized to higher odd dimensions [23]. Notice that, as one can expect for a codimension 1 Radon transform in 3D, the formulas are local.

In Section B we describe the numerical simulation of the data acquisition in 3D. The reconstruction algorithms based on the ρ-filtered backprojection formula (4.1) and the filtered backprojection one (4.2) are discussed in Section C.

Unfortunately backprojection type formulas are not known for spherical acquisition geometry in even dimensions. However different approximations of such formulas

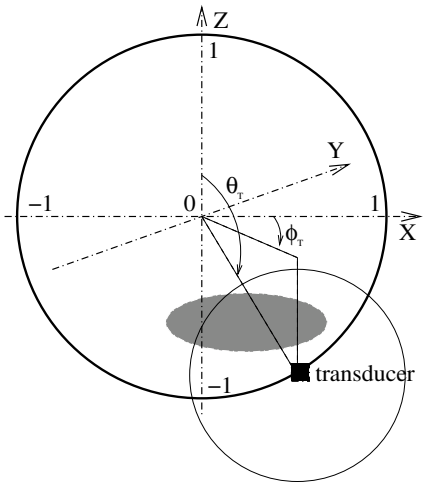


Fig. 7. Phantom setup

can be used. They happen to work well under most circumstances and can be improved in conjunction with post-processing by iterative methods. In Section D we present reconstructions in limited view 2D TAT using approximate backprojection type formulas. The comparison of reconstructions using approximate and exact formulas in 3D is discussed in Subsection 3 of Section C.

B. Data simulation in 3D

The region of reconstruction is the unit ball centered at the origin (see Fig. 7). All phantoms considered here are sums of indicator functions supported in ellipsoids completely contained inside the unit ball. The transducers are located on the surface of the unit sphere.

We parameterize the transducer location by two angles (ϕ_T, θ_T), where $\phi_T \in [0, 2\pi)$ is the the azimuthal angle in the xy-plane and $\theta_T \in [0, \pi]$ is the polar angle measured from the z-axis.

Since the spherical Radon transform is linear, it is enough to create projections for phantoms with a single ellipsoid and then superimpose the projections. For a single ellipsoid the data measured at a fixed transducer location at a given moment (i.e. for fixed (ϕ_T, θ_T, r)) is the surface area of a part of the sphere of integration cut by the intersecting ellipsoid. It can be expressed as a finite sum with terms of the form

$$\int_0^{2\pi} \int_{\theta_1(\phi)}^{\theta_2(\phi)} \sin\theta \; d\theta \; d\phi = \int_0^{2\pi} \left[\cos\theta_1(\phi) - \cos\theta_2(\phi)\right] d\phi \tag{4.3}$$

where each such term corresponds to a connected component of the intersection. Here ϕ and θ parameterize the sphere of integration and are independent of ϕ_T and θ_T, which parameterize the transducer location. The angles $\theta_1(\phi)$ and $\theta_2(\phi)$ are defined by the intersection of the integration sphere and the phantom's ellipsoid. The cosines of these angles can be found from the solution of a quartic equation describing that intersection.

In the numerical results presented below, the quartic equation is solved using the MATLAB built-in function "roots". By adding up these roots in an appropriate way, we obtain the inner integral with respect to the polar angle θ in equation (4.3). The result is a function of azimuthal angle ϕ, which we will denote $F(\phi)$. Depending on the location and parameters of the ellipsoid, $F(\phi)$ might be either a smooth π-periodic function of ϕ, or a piecewise smooth one (see Fig. 8). In the first case we compute its values at uniformly discretized locations on the interval $[0, \pi]$ and use the trapezoidal rule to compute the integral. For $F(\phi) \in C^2$, numerical integration using the trapezoidal rule is accurate to $O(h^4)$. If, however, $F(\phi)$ is only piecewise smooth on $[0, \pi]$, then we locate the pieces of *supp* $F(\phi)$ where it is smooth and use Gaussian quadrature to integrate over each piece.

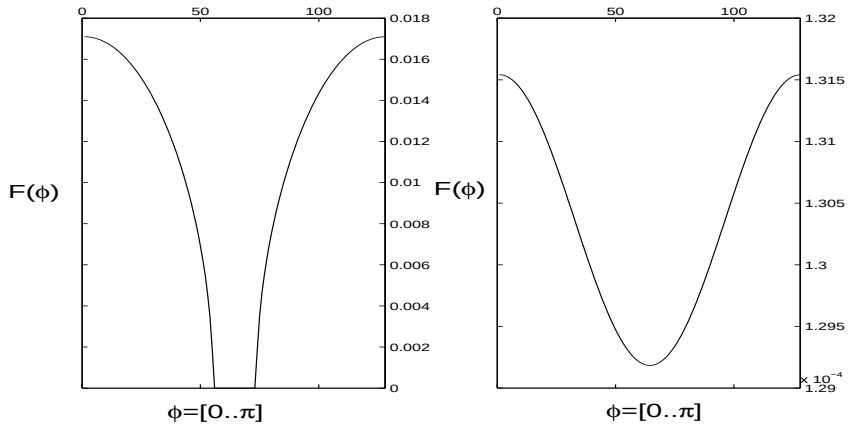


Fig. 8. $F(\phi)$ for two different locations of an ellipsoid

C. Reconstruction in 3D

Once we have generated the projection data, we reconstruct the original indicator functions of the phantoms. The reconstruction algorithms are based on the ρ-filtered backprojection (4.1) or the filtered backprojection (4.2).

The integrals over the unit sphere in (4.1) and (4.2) are computed as double integrals with respect to the azimuthal angle ϕ_T and the polar angle θ_T. The function to be integrated is periodic with respect to ϕ_T, making the trapezoidal rule an appealing quadrature choice. Integration with respect to θ_T is done by Gaussian quadrature. The Laplace operator is implemented through the Matlab built-in function "del2". The reconstructions were generated using Matlab 5.0.

In the results below, the resolution is $256 \times 256 \times 256$ over a $2 \times 2 \times 2$ volume, resulting in isotropic pixel dimension of $1/128$.

The algorithm is tested on the Defrise phantom which consists of five thin ellipsoids symmetrically centered along the z-axis (see Fig. 9). We numerate them from 1 to 5 starting with the lowest.

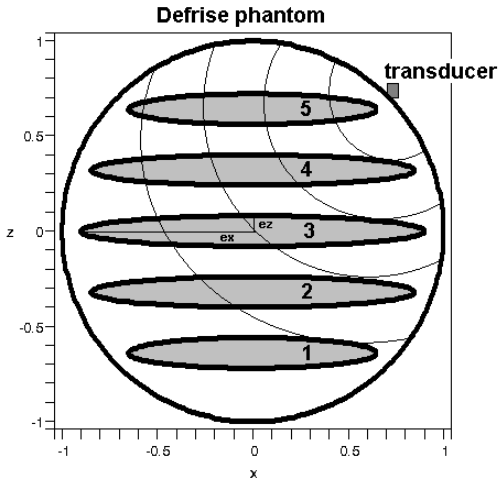
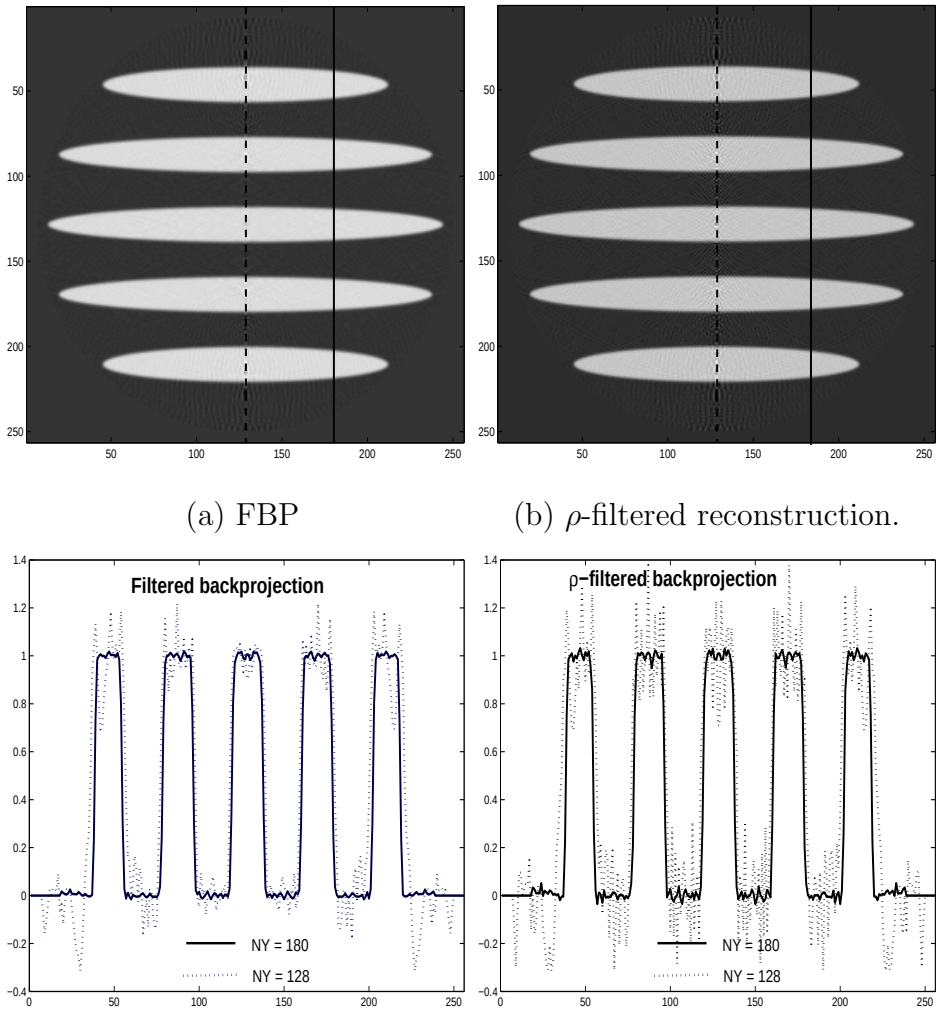


Fig. 9. The Defrise phantom slice along the plane y=0

Table I. Ellipsoids in the Defrise phantom

ellipse number	center $= (x_0, y_0, z_0)$	semiaxes lengths $= (e_x, e_y, e_z)$
1	$(0, 0, -0.64)$	$(0.65, 0.65, 0.08)$
2	$(0, 0, -0.32)$	$(0.85, 0.85, 0.08)$
3	$(0, 0, 0)$	$(0.9, 0.9, 0.08)$
4	$(0, 0, 0.32)$	$(0.85, 0.85, 0.08)$
5	$(0, 0, 0.64)$	$(0.65, 0.65, 0.08)$



(a) FBP (b) ρ-filtered reconstruction.

Fig. 10. Reconstructions and profiles of the Defrise phantom along the center $x = 0$ slice. Dashed lines correspond to the center $x = 0 = y$ profile; solid lines correspond to $x = 0,\ y = 0.4$

1. Full scan data

The data was acquired from the transducers located discretely over the sphere in the following way. The azimuthal angles of the transducer locations were uniformly

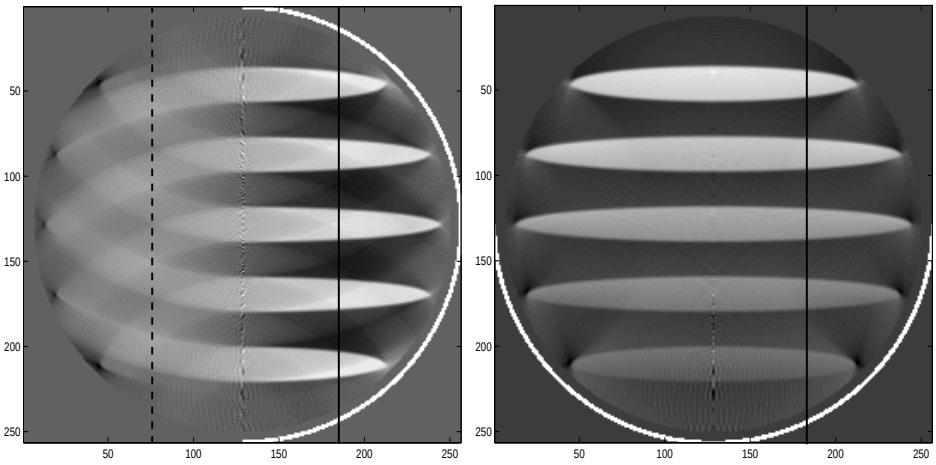
discretized to $N_\phi = 400$ points between 0 and 2π. The polar angles of the transducer locations corresponded to $N_\theta = 200$ Gaussian nodes on the interval form 0 to π, as described in the previous section. The radii of the integration spheres were uniformly discretized to $N_r = 200$ points from 0 to 2. The reconstruction was done by both methods: filtered backprojection and ρ-filtered backprojection.

The obtained results validate reconstruction formulas (4.1) and (4.2) (see Fig. 10). In both cases the Defrise phantom has a good reconstruction everywhere except along the z-axis ($x = y = 0$), where some noise is present, which, while not always noticeable on reconstructions, is visible on the graphs. The reason for appearance of that noise is the correlation of numerical errors along that axis of phantom's symmetry and is discussed in Subsection 4.

2. Partial scan data

Half-scan reconstructions were done using data from only the eastern hemisphere ($N_\phi = 200$, $N_\theta = 200$) or the southern hemisphere ($N_\phi = 400$, $N_\theta = 100$). These hemispheres are highlighted in Figure 11. The rest of the data has been zero-filled.

It is known [72, 44, 51, 65, 66, 82] that in case of incomplete data one can expect to recover stably only certain parts of the image the rest of it being blurred out. Namely, some parts of the wavefront set of the image will be lost. For our phantom, the singularities are jump discontinuities (edges) of imaged value f across an interface I (a surface of an ellipsoid in 3D). The wavefront $WF(f)$ of f in this situation is the set of pairs (x, n), where x is a point on I, and n is a vector conormal to I at x. As it was shown in [51, 82] using microlocal analysis, a point $(x, n) \in WF(f)$ can be stably detected from the Radon data, if and only if Rf includes data obtained from a sphere passing through x and conormal to n.



(a) ρ-filtered BP; 1/2-scan with respect to ϕ. The profiles along the lines are plotted below.

(b) FBP; 1/2-scan with respect to θ. The profile along the solid line is plotted below.

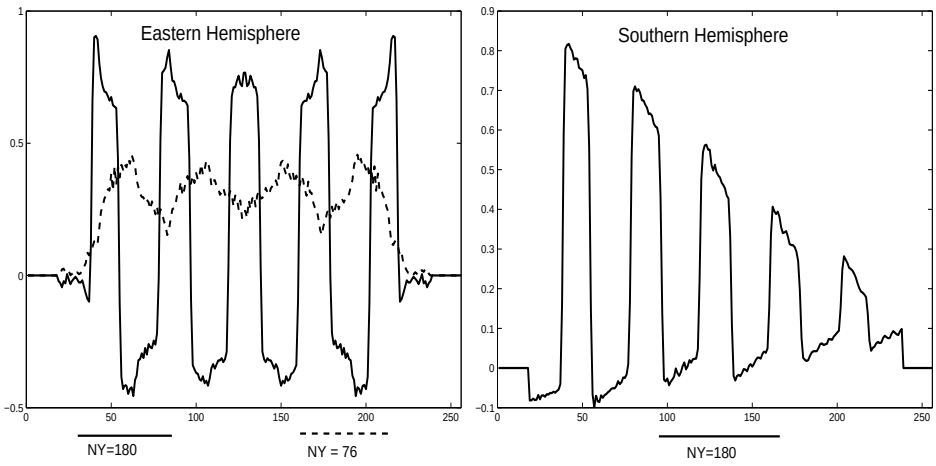


Fig. 11. Partial scan reconstructions of the Defrise phantom

In other words, one can see only those parts of an interface, that can be tangentially touched by spheres of integration centered at available transducer locations. The rest of the interface will be blurred.

Edges in the Defrise phantom were reconstructed (see Fig. 11) as expected. When the data is collected from the eastern hemisphere, there are enough spheres to touch tangentially all edges in the eastern hemisphere (see Fig. 9), but none to do it in the

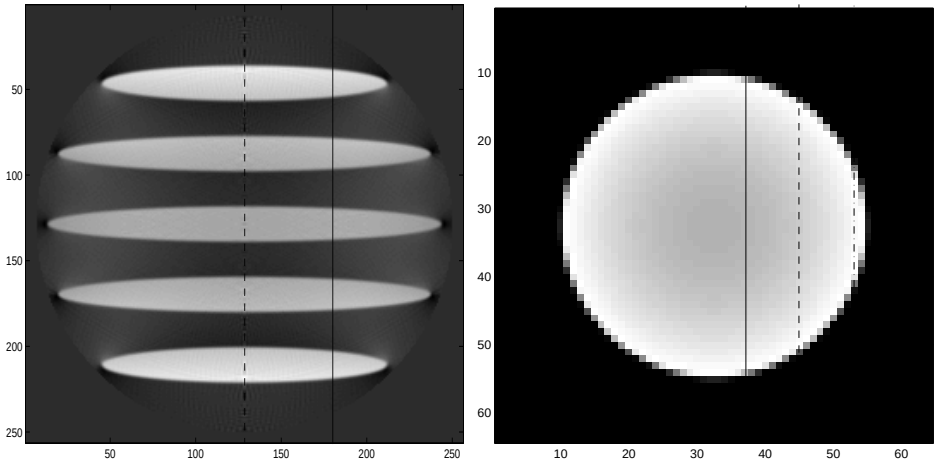
western hemisphere. That is why the locations of the edges in the eastern hemisphere were correctly reconstructed, while those in the western part were blurred. When the data is collected from the southern hemisphere, there are enough spheres to touch tangentially all edges in the Defrise phantom, hence all of them were resolved.

From the geometric description above, it is not hard to see that there may exist certain regions of reconstruction (locations of x, sometimes called audible zones) where any possible pair (x, n) belonging to $WF(f)$ is recognizable from Rf. In our examples, when the data is collected from the eastern or southern hemisphere, these regions are the eastern and southern half of the unit ball correspondingly.

Notice that the image values were not reconstructed correctly, since part of the data was missing. However, certain iterative techniques allow one to improve substantially the image values in the audible zone [66, 82, 67].

3. Comparison with an approximate backprojection

In early experimental work on thermoacoustic tomography, an approximate backprojection formula was used. It was written in analogy with the backprojection of regular Radon transform and looked similar to equation (4.1), except the missing weight factor $\frac{1}{|\mathbf{x}-\mathbf{p}|}$. The composition of this operator with the direct Radon transform is an elliptic pseudo-differential operator of order zero (see, e.g., [51, 28, 44].) Thus the locations and "strengths" of image singularities should be recovered correctly. However the values of the image function will not be accurate. The obtained reconstructions (see Fig. 12) validate the predictions correctly recovering locations of edges. The values of image functions are accurate near the center where $r \sim 1$, but degrade slowly with distance from the origin, as expected.



(a) Reconstruction of Defrise phantom at resolution 256×256 along $x = 0$

(b) Reconstruction of a sphere at resolution 64×64 along the slice $x = 0$

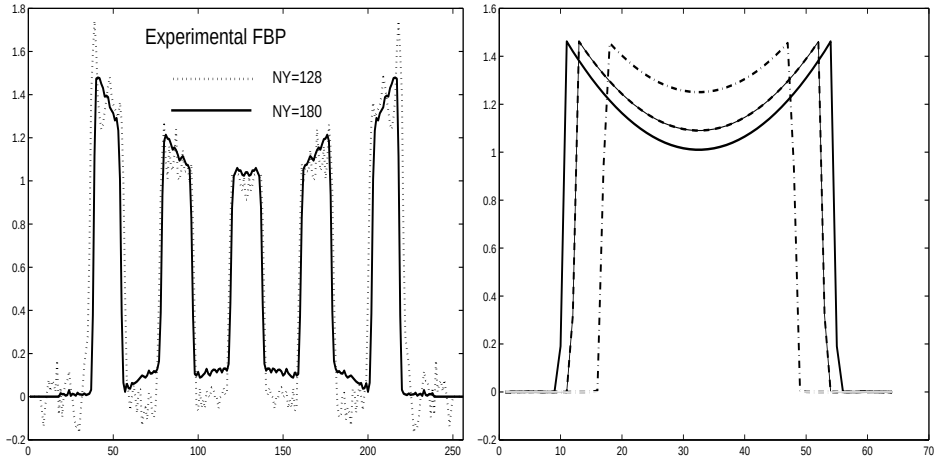


Fig. 12. Approximate FBP shows low-frequency shading

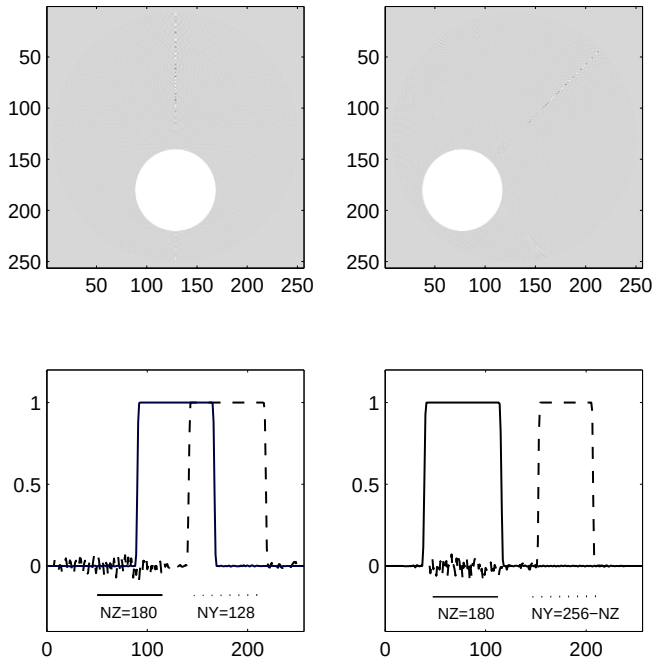


Fig. 13. FBP errors along the axis of symmetry

4. Errors in reconstruction

As it was mentioned before, the reconstructions of Defrise phantom have some noise along the axis of phantom's symmetry $x = y = 0$ (see Figs. 10, 11, 13). To discuss the reasons of appearance of that noise we consider reconstructions of some simpler phantoms consisting of indicator functions of a perfect ball. This allows us to compute the Radon transform analytically, hence to exclude the errors in the data simulation. For every fixed p_0, the function $Rf(p_0, r)$ is a third order polynomial with respect to r for $0 < r_1 \leq r \leq r_2 < 1$ and is zero for every other r. Filtered backprojection requires differentiating with respect to the radial variable r. We used centered finite differences to estimate the second order derivative $d^2/dr^2 Rf(p, r)$, which is exact on the third

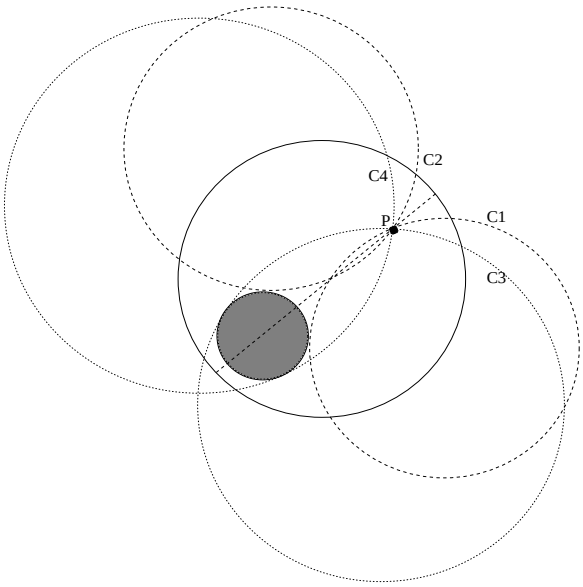


Fig. 14. The spheres C3 and C4 contain the phantom ball, while C1 and C2 do not

degree polynomials. Therefore, we compute $d^2/dr^2 Rf(p,r)$ exactly for all radii r, at least $2\Delta r$ away from r_1 and r_2. Hence the only errors in numerical differentiation that spread into the backprojection come with the data from spheres close to the ones touching tangentially the phantom ball. None of these spheres passes inside the phantom ball, hence backprojection at those points is free of errors from numerical differentiation (see Fig. 13).

Now let us consider a point p_1 on the axis of symmetry of the ball phantoms (the line connecting the center of the phantom ball and the origin). There are two sets of spheres that pass through that point and touch the phantom ball tangentially. The spheres in the first set contain the phantom ball, while the spheres in the second set do not. A 2D slice of this scenario is presented in Fig. 14.

Notice that all spheres in the same set have the same radius. So the errors from the numerical differentiation that they will bring into the backprojection algorithm are absolutely the same. The axis of symmetry is the only location in the reconstruc-

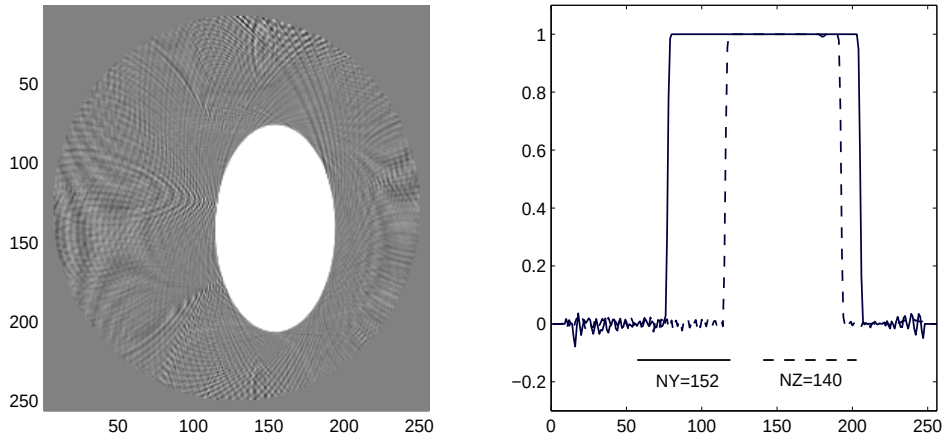


Fig. 15. An ellipsoidal phantom with center at (0, 0.2, -0.1) and semiaxes lengths equal to (0.4, 0.3, 0.5)

tion region where these errors are perfectly correlated. This resonance increases the magnitude of errors resulting in the noise along the symmetry axis on reconstructed images (see Fig. 13).

In case of ellipsoids in the Defrise phantom everything said above holds. In fact, the magnitude of errors is five times bigger since there are five ellipsoids with the same axis of symmetry there. At the same time, the reconstruction of an ellipsoidal phantom without any rotational symmetry has no axis of emphasized errors (see Fig. 15). (By rotational symmetry here we mean symmetry with respect to rotations around an axis passing through the origin.)

Another important observation is that the artifacts due to numerical errors are more severe in FBP than ρ-filtered images. However they might be reduced by mollification techniques [30].

D. Reconstructions in 2D limited view TAT[1]

In many applications of TAT, the signals cannot be collected from all directions. For example, in mammography the solid angle of detection is at most 2π steradians for a breast (π radians for a 2D scan). So, one faces here an incomplete data problem. Although (as we discussed in Chapter II) theoretically an arbitrarily small scanning arc (i.e., the arc of a circle over which the detectors move) suffices for the uniqueness of recovery, in practical implementations the limited-view problems usually lead to losing some parts of the high-frequency information and hence blurring of some sharp details. This is due to the fact that solving incomplete data problems usually leads to operations like Fourier filtrations with fast growing filters (e.g., Sect. 2.5.3 in [57]), which implies high sensitivity to errors in data. This in turn requires cutting off high frequencies and hence blurring the images. Sacrifices in high frequencies naturally lead to destroying sharp details (interfaces between different tissues) in the reconstruction.

The question of what parts of the singularities of the image can be stably reconstructed depending on the scanning geometry was already addressed in Subsection 2 of Section C. In short, the discussion there showed that in TAT one can see without blurring only those parts of the interfaces that can be touched tangentially by circles (spheres) centered at available detector positions.

As it has already been mentioned before, exact inversion procedures are known for circular and spherical Radon transforms in some special detection configurations. However, for the circular trajectories of detectors in 2D only special-function-

[1] Part of this section is reprinted with permission from "Reconstructions in limited view thermoacoustic tomography", by Y. Xu, L. Wang, G. Ambartsoumian and P. Kuchment, Medical Physics 31(4) April 2004, 724-733. Copyright ©2004, by American Association of Physicists in Medicine.

expansion methods are known. Our approach is to use an approximate FBP formula, which happens to work well under most circumstances and can be improved in conjunction with post-processing by an iterative method. Namely, for objects not too close to the detectors, one can think of projection lines as close to straight lines, and hence the circular Radon transform as being close to the standard Radon transform.

In this approach, the center p of the projection circle and its radius r (which is proportional to time) are analogs of the normal coordinates (ω, s) of a line $x \cdot \omega = s$ in the standard Radon transform, where ω is a unit vector normal to the line. FBP inversion of the standard Radon transform on the plane consists (up to a constant factor) in applying the first derivative with respect to s, then Hilbert transform with respect to s, and finally the backprojection operator, which averages over lines passing through a given point(see formula (2.4), Sect. II.2, [55]):

$$f(x) = \frac{1}{4\pi}\, R^\sharp\, H\, \frac{d}{ds}\, Rf(\omega, s),$$

where $R^\sharp$ is the dual Radon transform or backprojection:

$$(R^\sharp g)(x) = \int_{S^1} g(\omega, x \cdot \omega)\, d\omega,$$

and the Hilbert transform is defined through its Fourier transform as

$$(Hh)^\wedge(\sigma) = -i\, sgn(\sigma)\, \hat{h}(\sigma).$$

We implement a similar procedure in the circular Radon transform. This amounts to a differentiation with respect to the radius, a Hilbert transform with respect to the radius, and then a circular backprojection, i.e., averaging over the circles passing through a given point. One should also make sure that during the backprojection the tangent lines (or the normal vectors) to the projection curves at the given point rather than the centers of the projection curves (which coincide with detector positions),

rotate at a constant speed. In the case of incomplete data, one just replaces the missing data with zeros (possibly gradually phasing off the existing data closely to the missing data region to reduce the artifacts caused by the missing data) and then applies the formula. Although this is not an exact inversion, one can show using microlocal analysis that it preserves all "visible" singularities the numerical and experimental results presented below agree with this conclusion).

Another reconstruction method is to apply an additional (second) differentiation with respect to time (the radius) without applying a Hilbert transform. This leads to a local tomography type formula. The result of this procedure produces an expression of the form Λf where Λ is a pseudo-differential operator of positive order, which means that all the "visible" interfaces and other sharp details not only have correct locations, but also are emphasized (e.g., [44, 51]). This effect is well known in image processing, where for instance the Laplace operator is sometimes used to emphasize the edges.

In Fig.16 we present the numerical reconstructions of several phantoms for different partial scan situations using the approximate formula with Hilbert transform and local tomography formula emphasizing the edges.

Fig. 17 depicts the results of experimental measurements conducted by L. Wang and Y. Xu on physical phantoms. They applied the reconstruction methods described above to the obtained data and then iteratively improved the reconstruction using a truncated conjugate gradient method. The experimental data, due to the shape of the transducers impulse response function and electromagnetic pulse shape, already carries a filtration that makes the reconstruction similar to the local one. Then, unless an appropriate deconvolution is applied to the data during pre-processing, the interfaces are accentuated in the reconstruction. One can notice this in the actual reconstructions from experimental data.

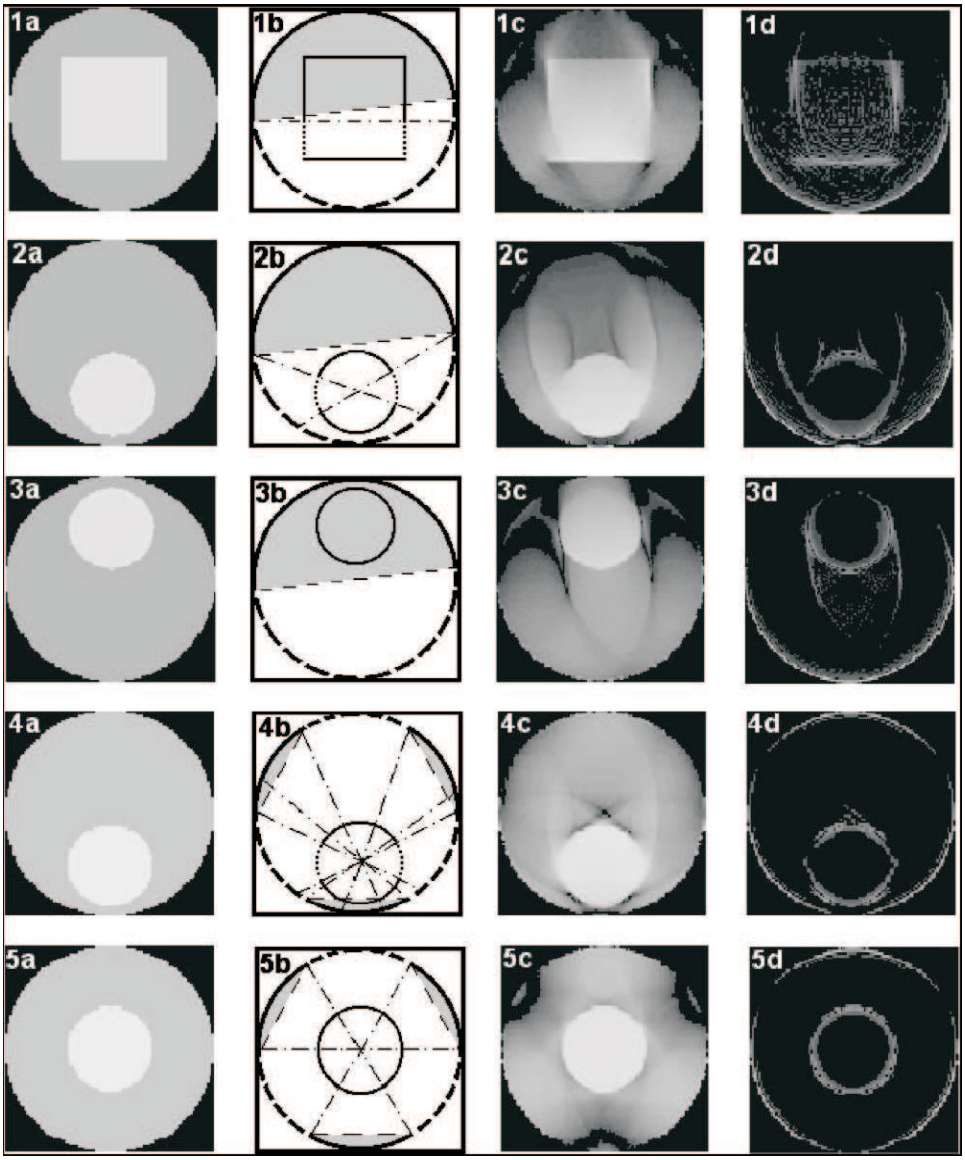


Fig. 16. (1a) A square phantom inside a circular detection curve. (1b) The diagram showing the detection curve (solid part of the outer circle), the "visible" (solid) and "invisible" (dashed) boundaries of the object predicted by theory, and the audible zone (shaded). (1c) FBP reconstruction. (1d) Local tomography reconstruction, where the boundary is emphasized. (2a-2d) A disk phantom outside the audible zone. (3a-3d) A disk phantom inside the audible zone. (4a-4d) An off-center disk phantom and a detection curve consisting of three arcs. (5a-5d) A centered disk phantom and a detection curve consisting of three arcs.

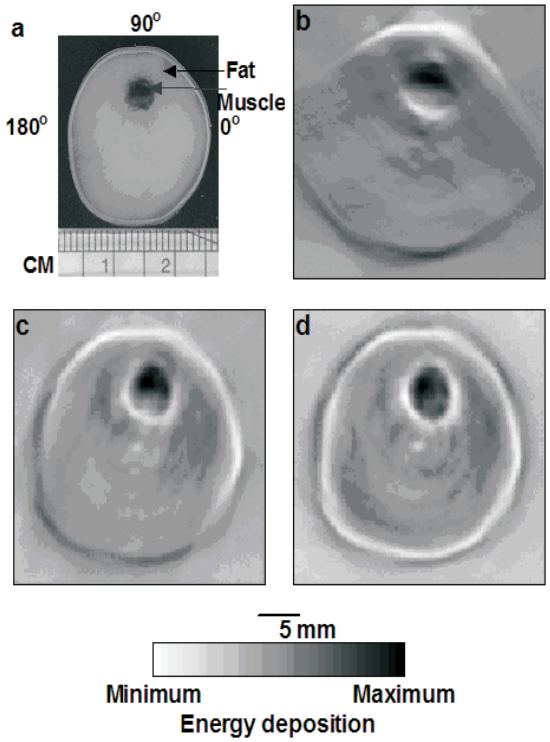


Fig. 17. (a) A photograph of the experimental sample. (b)-(d) TAT reconstructions using detection arcs of 92 degrees (from 50° to 142°), 202 degrees (from -18° to 184°), and 360 degrees, respectively. The blurred parts of the boundaries in (b) due to the limited view agree with the theoretical predictions. In (c) all the boundaries are resolved, since the object fits into the audible zone.

CHAPTER V

SUMMARY

The spherical Radon transform puts into correspondence to a given function its integrals over the set of all spheres with a given set of centers. SRT arises in several contemporary imaging techniques, including the newly developing thermoacoustic tomography and its sibling optoacoustic tomography, as well as radar, sonar and other applications. It has also been considered in relation to some problems of approximation theory, mathematical physics, and other areas of mathematics.

An important question arising in the study of these kinds of transforms is the uniqueness of reconstruction of an unknown function from its transform, or the problem of the transform's injectivity. A major result in this area is due to M. Agranovsky and E. T. Quinto who answered the question in dimension 2 for the class of compactly supported functions. However, the techniques used in their proof do not alow simple generalizations to higher dimensions or to the classes of non-compactly supported but rapidly decreasing functions. We have provided some new results concerning geometry of non-injectivity sets, as well as re-proved some known results with much simpler means. Most of these results hold in any dimension. The main theorem, which significantly restricts the geometry of the non-injectivity sets, does not require the unknown function to be compactly supported and in fact imposes no condition on its behavior at infinity.

Another important question is the description of the range for SRT. The range conditions are used extensively in tomography for various purposes: completing incomplete data, detecting and correcting measurement errors and hardware imperfections, recovering unknown attenuation, etc. Thus, as soon as a new Radon type

transform arises in an application, a quest for the range description begins. We have obtained a complete range description for the circular Radon transform in 2D with spherical acquisition geometry.

As in other types of tomography, the existence of exact or approximate inversion formulae and algorithms in TAT does not ensure the possibility of stable reconstruction in practice. A careful numerical validation is required for any new inversion formula or algorithm when it becomes available. We numerically validated recently discovered FBP and ρ-filtered BP inversion formulae for TAT data in 3D spherical geometry on the high-frequency Defrise phantom and some other ellipsoidal phantoms. Artifacts due to numerical errors were analyzed. We have also implemented partial scan reconstructions which agreed with the theoretical predictions of the types and locations of singularities that can be stably recovered. Similar work has been done for some approximate inversion algorithms in 2D spherical geometry.

REFERENCES

[1] Agranovsky M 2000 On a problem of injectivity for the Randon transform on a paraboloid *Contemp. Math.* **251** 1–14

[2] Agranovsky M, Berenstein C A and Kuchment P 1996 Approximation by spherical waves in L^p-spaces *J. Geom. Anal.* **6** 365–83

[3] Agranovsky M L and Quinto E T 1996 Injectivity sets for the Radon transform over circles and complete systems of radial functions *J. Funct. Anal.* **139** 383–413

[4] Agranovsky M L and Quinto E T 2001 Geometry of stationary sets for the wave equation in R^n, the case of finitely supported initial data *Duke Math. J.* **107** 57–84

[5] Agranovsky M L and Quinto E T 2003 Stationary sets for the wave equation on crystallographic domains *Trans. Am. Math. Soc.* **355** 2439–51

[6] Agranovsky M, Volchkov V and Zalcman L 1999 Conical uniqueness sets for the spherical Radon transform *Bull. Lond. Math. Soc.* **31** 231–6

[7] Aguilar V, Ehrenpreis L and Kuchment P 1996 Range conditions for the exponential Radon transform *J. d'Analyse Mathematique* **68** 1–13

[8] Aguilar V and Kuchment P 1995 Range conditions for the multidimensional exponential X-ray transform *Inverse Problems* **11** 977–82

[9] Ambartsoumian G and Kuchment P 2005 On the injectivity of the circular Radon transform arising in thermoacoustic tomography *Inverse Problems* **21** 473–85

[10] Ambartsoumian G and Kuchment P A range description for the planar circular Radon transform *SIAM J. Math. Anal.* To appear

[11] Ambartsoumian G and Patch S K Thermoacoustic tomography - implementation of exact backprojection formulas *arXiv:math.NA/0510638* Preprint

[12] Andersson L-E 1988 On the determination of a function from spherical averages *SIAM J. Math. Anal.* **19** 214–32

[13] Andrews G E, Askey R and Roy R 1999 *Special Functions (Encyclopedia of Mathematics and its Applications, vol. 71)* (Cambridge: Cambridge University Press)

[14] Bers L, John F and Schechter M 1964 *Partial Differential Equations* (New York: Wiley)

[15] Cormack A and Quinto E T 1989 A problem in radiotherapy: questions of non-negativity *Internat. J. Imaging Systems and Technology* **1** 120–4

[16] Cormack A and Quinto E T 1990 The mathematics and physics of radiation dose planning *Contemp. Math.* **113** 41–55

[17] Courant R and Hilbert D 1962 *Methods of Mathematical Physics, Volume II Partial Differential Equations* (New York: Interscience)

[18] Davies B 2001 *Integral Transforms and their Applications* (New York: Springer)

[19] Denisjuk A 1999 Integral geometry on the family of semi-spheres *Fract. Calc. Appl. Anal.* **2** 31–46

[20] Ehrenpreis L 2003 *The Universality of the Radon Transform* (Oxford: Oxford University Press)

[21] Fawcett J A 1985 Inversion of n-dimensional spherical averages *SIAM J. Appl. Math.* **45** 336–41

[22] Finch D and Rakesh 2006 The range of the spherical mean value operator *Inverse Problems* **22** 923–38

[23] Finch D, Rakesh and Patch S 2004 Determining a function from its mean values over a family of spheres *SIAM J. Math. Anal.* **35** 1213–40

[24] Flatto L, Newmann D J and Shapiro H S 1966 The level curves of harmonic polynomials *Trans. Amer. Math. Soc.* **123** 425–36

[25] Gelfand I, Gindikin S and Graev M 1980 Integral geometry in affine and projective spaces *J. Sov. Math.* **18** 39–167

[26] Gelfand I, Gindikin S and Graev M 2003 *Selected Topics in Integral Geometry (Transl. Math. Monogr., vol 220)* (Providence, RI: American Mathematical Society)

[27] Gelfand I, Graev M and Vilenkin N 1965 *Generalized Functions, Volume V Integral Geometry and Representation Theory* (New York: Academic Press)

[28] Guillemin V 1985 On some results of Gelfand in integral geometry *Proc. Symp. Pure Math.* **43** 149–55

[29] Guillemin V and Sternberg S 1979 Some problems of integral geometry and some related problems in microlocal analysis *Amer. J. Math.* **101** 915–55

[30] Haltmeier M, Schuster T and Scherzer O 2005 Filtered backprojection for thermoacoustic computed tomography in spherical geometry *Math. Methods in the Applied Sciences* **28** 1919–37

[31] Heike U 1986 Single-photon emission computed tomography by inverting the attenuated Radon transform with least-squares collocation *Inverse Problems* **2** 307–30

[32] Helgason S 1980 *The Radon Transform* (Basel: Birkhäuser)

[33] Hertle A 1988 The identification problem for the constantly attenuated Radon transform *Math. Z.* **197** 13–9

[34] Hörmander L 1983 *The Analysis of Linear Partial Differential Operators*, vol 1 (New York: Springer)

[35] John F 1971 *Plane Waves and Spherical Means, Applied to Partial Differential Equations* (New York: Dover)

[36] Joines W, Zhang Y, Li C and Jirtle R 1994 The measured electrical properties of normal and malignant human tissues from 50 to 900 MHz *Med. Phys.* **21** 547–50

[37] Kruger R A, Kopecky K K, Aisen A M, Reinecke D R, Kruger G A and Kiser Jr W L 1999 Thermoacoustic CT with radio waves: a medical imaging paradigm *Radiology* **211** 275–8

[38] Kruger R A, Liu P, Fang Y R and Appledorn C R 1995 Photoacoustic ultrasound (PAUS) reconstruction tomography *Med. Phys.* **22** 1605-9

[39] Kruger R A, Miller K D, Reynolds H E, Kiser Jr W L, Reinecke D R and Kruger G A 2000 Contrast enhancement of breast cancer in vivo using thermoacoustic CT at 434 MHz *Radiology* **216** 279–83

[40] Kruger R A, Reinecke D R and Kruger G A 1999 Thermoacoustic computed tomography *Med. Phys.* **26** 1832–7

[41] Kuchment P 1993 Unpublished results on spherical means, Dept. of Mathematics, Wichita State University, Wichita, KS

[42] Kuchment P 1993 On positivity problems for the Radon transform and some related transforms *Contemp. Math.* **140** 87–95

[43] Kuchment P 2006 Generalized transforms of Radon type and their applications *The Radon Transform, Inverse Problems, and Tomography (Proc. Symp. Appl. Math., vol 63)* ed G Olafsson and E T Quinto (Providence, RI: American Mathematical Society) pp 67 – 91

[44] Kuchment P, Lancaster K and Mogilevskaya L 1995 On the local tomography *Inverse Problems* **11** 571–89

[45] Kuchment P and Lvin S 1990 Paley-Wiener theorem for exponential Radon transform *Acta Appl. Math.* **18** 251–60

[46] Kuchment P and Lvin S 1991 The range of the exponential Radon transform *Soviet Math. Dokl.* **42** 183–4

[47] Kuchment P and Quinto E T 2003 Some problems of integral geometry arising in tomography *Chapter XI in [20]*

[48] Lebedev N N 1972 *Special Functions and Their Applications* (New York: Dover)

[49] Lin V Ya and Pinkus A 1993 Fundamentality of ridge functions *J. Approx. Theory* **75** 295–311

[50] Lin V Ya and Pinkus A 1994 Approximation of multivariable functions *Advances in Computational Mathematics* ed H P Dikshit and C A Micchelli (Singapore: World Scientific) pp 1–9

[51] Louis A K and Quinto E T 2000 Local tomographic methods in Sonar *Surveys on Solution Methods for Inverse Problems* (Vienna: Springer) pp 147–54

[52] Lvin S 1994 Data correction and restoration in emission tomography, pp 149–55 in [73]

[53] Mennesier C, Noo F, Clackdoyle R, Bal G and Desbat L 1999 Attenuation correction in SPECT using consistency conditions for the exponential ray transform *Phys. Med. Biol.* **44** 2483–510

[54] Natterer F 1983 Exploiting the range of Radon transform in tomography *Numerical Treatment of Inverse Problems in Differential and Integral Equations* ed P Deuflhard and E Hairer (Basel: Birkhäuser)

[55] Natterer F 1986 *The Mathematics of Computerized Tomography* (New York: Wiley)

[56] Natterer F 2001 Inversion of the attenuated Radon transform *Inverse Problems* **17** 113–119

[57] Natterer F and Wübbeling F 2001 Mathematical methods in image reconstruction *Monographs on Mathematical Modeling and Computation*, vol 5 (Philadelphia, PA: SIAM)

[58] Nilsson S 1997 Application of fast backprojection techniques for some inverse problems of integral geometry *Linkoeping Studies in Science and Technology* Dissertation 499, Dept. of Mathematics, Linkoeping University, Linkoeping, Sweden

[59] Nolan C J and Cheney M 2002 Synthetic aperture inversion *Inverse Problems* **18** 221–35

[60] Noo F, Clackdoyle R and Wagner J-M 2002 Inversion of the $3D$ exponential X-ray transform for a half equatorial band and other semi-circular geometries *Phys. Med. Biol.* **47** 2727–35

[61] Noo F and Wagner J-M 2001 Image reconstruction in $2D$ SPECT with 180^{o} acquisition *Inverse Problems* **17** 1357–71

[62] Norton S J 1980 Reconstruction of a two-dimensional reflecting medium over a circular domain: exact solution *J. Acoust. Soc. Am.* **67** 1266–73

[63] Norton S J and Linzer M 1981 Ultrasonic reflectivity imaging in three dimensions: exact inverse scattering solutions for plane, cylindrical, and spherical apertures *IEEE Trans. Biomed. Eng.* **28** 202–20

[64] Novikov R 2002 On the range characterization for the two–dimensional attenuated X–ray transform *Inverse Problems* **18** 677–700

[65] Palamodov V P 2000 Reconstruction from limited data of arc means *J. Fourier Anal. Appl.* **6** 25–42

[66] Pan X, Zou Y and Anastasio M 2003 Data redundancy and reduced-scan reconstruction in reflectivity tomography *IEEE Trans. Imag. Proc.* **12** 784–95

[67] Paultauf G, Viator J A, Prahl S A and Jacques S L 2002 Iterative reconstruction algorithm for optoacoustic imaging *J. Acoust. Soc. Am.* **112** 1536–44

[68] Patch S K 2004 Thermoacoustic tomography - consistency conditions and the partial scan problem *Phys. Med. Biol.* **49** 1–11

[69] Ponomarev I 1995 Correction of emission tomography data. Effects of detector displacement and non-constant sensitivity *Inverse Problems* **10** 1–8

[70] Quinto E T 1980 The dependence of the generalized Radon transform on defining measures *Trans. Amer. Math. Soc.* **257** 331–46

[71] Quinto E T 1982 Null spaces and ranges for the classical and spherical Radon transforms *J. Math. Anal. Appl.* **90** 408–20

[72] Quinto E T 1993 Singularities of the X-ray transform and limited data tomography in $\mathbf{R}^2$ and $\mathbf{R}^3$ *SIAM J. Math. Anal.* **24** 1215–25

[73] Quinto E T, Cheney M and Kuchment P (ed) 1994 *Tomography, Impedance Imaging, and Integral Geometry (Lectures in Appl. Math., vol 30)* (Providence, RI: American Mathematical Society)

[74] Shubin M 1987 *Pseudodifferential Operators and Spectral Theory* (Berlin - New York: Springer)

[75] Solmon D 1990 Two inverse problems for the exponential Radon transform *Inverse Problems in Action* ed P S Sabatier (Berlin: Springer) pp 46–53

[76] Solmon D 1995 The identification problem for the exponential Radon transform *Math. Methods in the Applied Sciences* **18** 687–95

[77] Spivak M 1999 *A Comprehensive Introduction to Differential Geometry* vol 3 (Houston: Publish or Perish)

[78] Tretiak O J and Metz C 1980 The exponential Radon transform *SIAM J. Appl. Math.* **39** 341–54

[79] Xu M and Wang L V 2002 Time-domain reconstruction for thermoacoustic tomography in a spherical geometry *IEEE Trans. Med. Imag.* **21** 814-22

[80] Xu Y, Feng D and Wang L V 2002 Exact frequency-domain reconstruction for thermoacoustic tomography: I. Planar geometry *IEEE Trans. Med. Imag.* **21** 823-8

[81] Xu Y, Xu M and Wang L V 2002 Exact frequency-domain reconstruction for thermoacoustic tomography: II. Cylindrical geometry *IEEE Trans. Med. Imag.* **21** 829-33

[82] Xu Y, Wang L V, Ambartsoumian G and Kuchment P 2004 Reconstructions in limited view thermoacoustic tomography *Med. Phys.* **31** 724–33

[83] Zobin N 1993 Private communication

VITA

Gaik Ambartsoumian received his Diploma in Applied Mathematics from the Obninsk Institute of Nuclear Power Engineering (Russia) in February 2001. He entered the graduate program in Mathematics at Texas A&M University in September 2001, and received his Ph.D. in Mathematics in August 2006.

Gaik Ambartsoumian may be reached at the Department of Mathematics, University of Texas at Arlington, P.O. Box 19408, Arlington, TX 76019-0408. His e-mail address is gaik@tamu.edu.