INSTITUTE OF PHYSICS PUBLISHING INVERSE PROBLEMS

Inverse Problems **21** (2005) 473–485 doi:10.1088/0266-5611/21/2/004

On the injectivity of the circular Radon transform

Gaik Ambartsoumian and Peter Kuchment

Mathematics Department, Texas A & M University, College Station, TX 77843-3368, USA

E-mail: kuchment@math.tamu.edu and haik@tamu.edu

Received 13 July 2004, in final form 14 January 2005
Published 1 February 2005
Online at stacks.iop.org/IP/21/473

Abstract
The circular Radon transform integrates a function over the set of all spheres with a given set of centres. The problem of injectivity of this transform (as well as inversion formulae, range descriptions, etc) arises in many fields from approximation theory to integral geometry, to inverse problems for PDEs and recently to newly developing types of tomography. A major breakthrough in the 2D case was made several years ago in a work by Agranovsky and 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, especially those based on simple PDE techniques, which would be less restrictive in more general situations. This paper provides some new results that one can obtain by methods that essentially involve only the finite speed of propagation and domain dependence for the wave equation.

1. Introduction

The circular Radon transform integrates a function over the set of all spheres with a given set of centres. 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 ([1–7, 9–13, 16, 17, 20–25, 27–33, 37–40]). Although significant progress has been achieved, some related analytic problems have proven to be rather hard and remain unresolved till now. A new wave of interest in such transforms has been sparked by the recent development of the thermoacoustic tomography (TAT or TCT) as one of the promising methods of medical imaging (e.g., [21, 37–40]). The TAT procedure can be described as follows: a short microwave or radiofrequency electromagnetic pulse is sent through the biological object. At each internal location x certain energy $H(x)$ is absorbed. The cancerous cells happen to absorb several times more MW (or RF) energy than the normal ones, which means that significant increases of the values of $H(x)$ are expected at tumorous locations. 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. It has been shown that here one effectively measures the integrals of $H(x)$ over all spheres centred at the transducers' locations. In other words, one needs to invert the above-mentioned generalized Radon transform of H ('generalized,' since integration is done over spheres). It is clear from the dimension considerations that it should be sufficient to run the transducers along a curve in the case of a 2D problem or a surface in 3D. The most popular geometries of these surfaces (curves) that have been implemented are spheres, planes and cylinders [37–39].

The central problems that arise in these studies are: uniqueness of reconstruction, reconstruction formulae and algorithms, stability of the reconstruction, description of the range of the transform and incomplete data problems.

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) [26, 27]. However, they are much more complex and not that well understood for the circular Radon transform that arises in TAT.

This paper contains some new approaches and results concerning the uniqueness problem. The reader should be aware that for the currently practically used geometries of TAT the uniqueness issue has been resolved. For example, for the spherical location of the centres (transducers) uniqueness follows for instance from corollary 5, first proven in [22] (see also [2, 3] and references therein). For the planar location, it has been known for a long time [9, 20] that only odd functions with respect to this plane cannot be reconstructed from the spherical integral data. However, the complete understanding of the uniqueness problem for general locations of the transducers remains elusive (especially in dimensions higher than two). The aim of this paper is to make progress in filling this gap by obtaining new uniqueness results, as well as by reproving some known results by simpler means, which makes them easier to extend to higher dimensions and other geometries.

The next section contains the mathematical formulation of the problem and its brief history. The following section contains the main results of this paper. It is followed by sections containing further remarks and acknowledgments.

2. Formulation of the problem and known results

The discussion of the previous section motivates the study of the following 'circular' Radon transform. Let $f(x)$ be a continuous function on $\mathbb{R}^n$, $n \geqslant 2$.

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

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

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

In this definition we do not restrict the set of centres 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 suggests to restrict the set of centres 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:

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

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

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

This problem has been around in different guises for quite a while [3, 11, 23, 24]. The paper [3] contains a survey of some other problems that lead to the injectivity question for R_S.

The first important observations concerning non-injectivity sets were made by Lin and Pincus [23, 24] and by Zobin [41]. Their results imply in particular 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 4. *Any set $S \subset \mathbb{R}^n$ of uniqueness for the harmonic polynomials is a set of injectivity for the transform R.*

In particular, this implies

Corollary 5. *If $U \subset \mathbb{R}^n$ is a bounded domain, then $S = \partial U$ is an injectivity set of R.*

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

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.

Definition 6. *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*[1]*:*

$$L_k = \{t\, \mathrm{e}^{\mathrm{i}\pi k/n} \mid -\infty < t < \infty\}.$$

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. 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 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 Lin and Pincus that these are the only non-injectivity sets for compactly supported functions on the plane. Proving this conjecture, Agranovsky and Quinto established the following result:

Theorem 7 [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 8 [3]**.** *The following condition is necessary and sufficient for S to be a set of injectivity for the circular Radon transform on $C_c(\mathbb{R}^n)$:*

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

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.*

The reader notes that for $n = 2$ this boils down to theorem 7.

The beautiful proof of theorem 7 by Agranovsky and 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 [19]), and known geometric structure of level sets of harmonic polynomials in 2D (e.g., [14]). 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, 22] that stems from known relations between spherical integrals and the wave equation (e.g., [9, 20]). Namely, consider the initial value problem for the wave equation in $\mathbb{R}^n$:

$$u_{tt} - \triangle u = 0, \qquad x \in \mathbb{R}^n, \quad t \in \mathbb{R} \tag{1}$$

$$u|_{t=0} = 0, \qquad u_t|_{t=0} = f. \tag{2}$$

Then

$$u(x,t) = \frac{1}{(n-2)!}\frac{\partial^{n-2}}{\partial t^{n-2}}\int_0^t r(t^2 - r^2)^{(n-3)/2}(Rf)(x,r)\,\mathrm{d}r, \qquad t \geqslant 0.$$

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 9 [3, 22]. *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 (1)–(2) 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 [22] of corollary 5 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 10 [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 $q \leqslant 2n/(n-1)$. 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 7 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 [13] by Finch, Rakesh and 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:

Theorem 11 [13]**.** *Let D be a bounded, open, subset of $\mathbb{R}^n$, $n \geqslant 2$, 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 $f = 0$.*

Although this theorem follows from microlocal results in [3][2], its significance lies in the proof provided in [13] (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 11 in [13]. They will be relevant for our purpose as well.

Proposition 12 [13]**.** *Let $B_\epsilon(p) = \{x \in \mathbb{R}^n \mid |x-p| < \epsilon\}$. If u is a distribution and satisfies (1) 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

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

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 $\mathbb{R}^n \backslash D$ joining p to q. Then $d(p,q)$ is a metric on $\mathbb{R}^n \backslash D$. For any point p in $\mathbb{R}^n \backslash D$ and any positive number r, define $E_r(p)$ to be the ball of radius r and centre at p in $\mathbb{R}^n \backslash D$ with respect to this metric, i.e.

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

Proposition 13 [13]**.** *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*

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

$$u = h \qquad on \quad S \times R.$$

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

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

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

3. Further injectivity results by PDE means

We will show now how simple tools similar to propositions 12 and 13, namely finite speed of propagation and domain of dependence for the wave equation allow one to obtain more results

[2] Results of [3] make the situation described in theorem 11 impossible, since the support of f lies on one side of a tangent plane to Γ. See also theorem 19 and [25].

concerning geometry of non-injectivity sets, as well as to re-prove some known results with much simpler means. The final goals were to recover the full result of [3] in 2D and to prove its analogues 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 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 5 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. Note that theorem 11 does exactly that when S contains an open part of the boundary of a smooth strictly convex domain where f is supported. Theorem 7, on the other hand excludes all compactly supported f in $\mathbb{R}^2$, unless $S = \omega\Sigma_N$. Similarly, theorem 10 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, $j = 1, \ldots, m$. One can define the interior metric in H^j as follows:

$$d^j(p, q) = \inf\{\text{length of } \gamma\}, \tag{3}$$

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

Theorem 14. *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*

$$\begin{aligned} \operatorname{dist}(x, \operatorname{supp} f \cap H^j) &= \operatorname{dist}^j(x, \operatorname{supp} f \cap H^j) \\ &\leqslant \operatorname{dist}(x, \operatorname{supp} f \cap H^k), k \neq j, \end{aligned} \tag{4}$$

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

$$\operatorname{dist}(x, \operatorname{supp} f \cap H^j) = \operatorname{dist}^j(x, \operatorname{supp} f \cap H^j) = \operatorname{dist}(x, \operatorname{supp} f). \tag{5}$$

Thus, the expressions in (5) in fact do not depend on $j = 1, \ldots, m$.

Remark 15. Note 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 behaviour 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. Note first of all, that the function $d^j(p,x)$ has gradient $|\nabla_x d^j(p,x)| \leqslant 1$ a.e.[3]

Let us prove now the equality

$$\operatorname{dist}(x, \operatorname{supp} f \cap H^j) = \operatorname{dist}^j(x, \operatorname{supp} f \cap H^j). \tag{6}$$

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

$$\operatorname{dist}(x, \operatorname{supp} f \cap H^j) = d_1 < d_2 = \operatorname{dist}^j(x, \operatorname{supp} f \cap H^j). \tag{7}$$

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

$$\operatorname{dist}(p, \operatorname{supp} f \cap H^j) \leqslant d_3 < d_4 \leqslant \operatorname{dist}^j(p, \operatorname{supp} f \cap H^j). \tag{8}$$

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 $\tau \leqslant (d_3 + d_4)/2$. Consider the solution $u(x,t)$ of the wave equation problem (1)–(2) 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, chapter 1 in [8]) we conclude that $u = 0$ in V. For the reader's convenience, let us provide brief details of the corresponding calculations from [8]: since $\Box u = 0$, $u = u_t = 0$ on Σ_1 and $u|_S = 0$ for all times, we get by integration by parts

$$\begin{aligned} 0 &= \int_V u_t \Box u \,\mathrm{d}x\,\mathrm{d}t = \int_{t=\phi(x)} \frac{1}{2}\left(|\nabla u|^2 + u_t^2 + 2u_t \nabla\phi \nabla u\right) \mathrm{d}x \\ &= \frac{1}{2}\int_{\phi(x)\geqslant 0} (|\nabla(u(x,\phi(x))|^2 + (1 - |\nabla\phi|^2) u_t(x,\phi(x))^2)\,\mathrm{d}x. \end{aligned} \tag{9}$$

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

$$\int_{\phi(x)\geqslant 0} (|\nabla(u(x,\phi(x))|^2)\,\mathrm{d}x = 0$$

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 $|t| \leqslant (d_3 + d_4)/2$. Note that $(d_3 + d_4)/2 > d_3$. Now applying proposition 12 to the wave equation in the whole space, we conclude that

$$\operatorname{dist}(p, \operatorname{supp} f) > d_3, \tag{10}$$

and hence

$$\operatorname{dist}(p, \operatorname{supp} f \cap H^j) > d_3, \tag{11}$$

which is a contradiction. This proves (6). It is now sufficient to prove

$$\operatorname{dist}(x, \operatorname{supp} f \cap H^j) \leqslant \operatorname{dist}(x, \operatorname{supp} f \cap H^k) \tag{12}$$

for $k \neq j$. This in fact is an immediate consequence of (10). Alternatively, we can repeat the same consideration as above in a simplified version. Namely, suppose that

$$\operatorname{dist}(x, \operatorname{supp} f \cap H^j) > d_2 > d_1 > \operatorname{dist}(x, \operatorname{supp} f \cap H^k) \tag{13}$$

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 $t = d_2 - |x - p|$

[3] In order to justify legality of the calculation presented below, one can either use geometric measure theory tools, as in [13], or just note that due to algebraicity of S, the function $d^j(p,x)$ is piece-wise analytic.

(p fixed in the small ball) and the boundary $S \times \mathbb{R}$. Consider the solution $u(x, t)$ of the wave equation problem (1)–(2) 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, chapter 1 in [8]) we conclude that $u = 0$ in V. In particular, $u(p, t) = 0$ for all $p \in B$ and $|t| \leqslant d_2$. Now applying proposition 12 to the wave equation in the whole space, we conclude that

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

and hence

$$\text{dist}(p, \text{supp } f \cap H^k) > d_2, \tag{14}$$

which is a contradiction. □

We will now show several corollaries that can be extracted from theorem 14.

Corollary 16. *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 14 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 $\bigcap_j H^j$, 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 notes that a similar proof establishes the following

Corollary 17. *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 $d_S(p, q) \geqslant |p - q|$. Using this metric, we can define the corresponding distances $dist_S$ from points to sets.

Theorem 18. *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 $p \in \mathbb{R}^n - S$*

$$\text{dist}_S(p, \text{supp } f) = \text{dist}(p, \text{supp } f). \tag{15}$$

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

Proof. Assume that (15) does not hold, i.e.

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

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

$$\mathrm{dist}^j(p, \mathrm{supp}\, f \cap H^j) \geqslant \mathrm{dist}_S(p, \mathrm{supp}\, f) \tag{16}$$

and hence

$$\mathrm{dist}^j(p, \mathrm{supp}\, f \cap H^j) > \mathrm{dist}(p, \mathrm{supp}\, f). \tag{17}$$

This, however, contradicts theorem 14. □

Let us formulate another example of a geometric constraint on pairs S, f such that $R_S f = 0$.[4]

Theorem 19. *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 $f = 0$.*[5]

Proof of the theorem. Let us denote by $K_p(\mathrm{supp}\, f)$ 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 $K_{p_0}(\mathrm{supp}\, f)$, 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 $K_p(\mathrm{supp}\, f)$ will intersect S. This means in particular, that for this point p we have $\mathrm{dist}_S(p, \mathrm{supp}\, f) > \mathrm{dist}(p, \mathrm{supp}\, f)$. According to theorem 18, this implies that $f = 0$. □

Corollary 20. *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 21. *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.*

Proof. Corollary 5 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 20, 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

[4] A similar statement in the case of analytic surfaces S was announced in [25] for distributions f. The proof is claimed to be based upon microlocal analysis.

[5] This implies, in particular, theorem 11.

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 7 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*,[6] 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 $\mathbb{R}^n \backslash K$ such that $\text{dist}(x, K) = r$ 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).

Theorem 22. *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 7.*

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 14 implies that any irreducible component of S must meet any neighbourhood of the support of f. If we take the neighbourhood of radius $r < r_0$, then each point on S in this neighbourhood will have a unique closest point on supp f. Applying now corollary 17, 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 those odd with respect to L (e.g., [3, 9, 20]). Hence, f is odd with respect to all lines L_j. In particular, every of these lines passes through the centre of mass of the support of f. Hence, lines L_j form a 'cross'[7]. 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 note that if this is not the case, then there is no non-zero function that is odd simultaneously with respect to all the lines. □

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

Proposition 23. *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 (a scroll)*[8].

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.

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

[7] 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.

[8] A *ruled surface*, or a *scroll* is the union of a family of lines (e.g., [36]).

Remark 24. 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 8 for this particular case.

4. Additional remarks

(1) Agranovsky and 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, 4–6] for details.

(2) One of our goals was to obtain the complete theorem 7, 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 its analogues 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 21 and 23 and theorem 22) 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 13 and theorem 18, 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) In this paper 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) Most of our results can be generalized to some Riemannian manifolds, in particular to the hyperbolic plane (where the analogue of theorem 7 has not been proven yet). We plan to address these issues elsewhere. Quinto has recently announced a version of theorem 19 in the case of distributions for the spherical transforms on real-analytic Riemannian manifolds with infinite injectivity radius and an analytic set S of centres [35].

(6) A closer inspection of the results of section 3 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 [25]. We hope to address this issue elsewhere.

(7) As Boman notified us during the April 2004 AMS meeting in Lawrenceville, he jointly with Sjostrand, being unaware of our work, had recently independently obtained some results analogous to some of those presented here (e.g., to theorem 18).

(8) We have not touched the problem of finding explicit inversion formulae for the circular transforms. Such formulae are known for the spherical, planar and cylindrical sets of centres [7, 10, 12, 13, 28, 31, 37–39]. They come in two kinds: those involving expansions into special functions, and those of backprojection type. Exact backprojection-type formulae are known for the planar geometry [10, 31] and recently for the spherical geometry in odd dimensions [13] if the function to be reconstructed is supported inside the sphere of transducers.

    Another problem deserving attention is finding the ranges of transforms R_S. Such knowledge could be used, for instance, to replenish missing data. Some necessary range conditions have been recently obtained in [32] for spherical location of transducers.

An important problem of reconstruction with incomplete data was treated in [25, 40] based on an earlier work by Quinto in [34].

(9) An important integral geometric technique of the so-called κ-operator has been developed in Gelfand's school (e.g., [15, 16]). It has been applied recently to the problems of the circular Radon transform (see [17], the last chapter of [16], and references therein), albeit applicability of this method to the problems of the kind we consider in this paper is not completely clear yet.

Acknowledgments

The authors express their gratitude to M Agranovsky, J Boman, E Chappa, M de Hoop, L Ehrenpreis, D Finch, S Patch, E T Quinto, L Wang, M Xu, Y Xu and N Zobin for information about their work and discussions. The authors are also grateful to the reviewers for useful comments.

This research was partly based upon work supported by the NSF under grants DMS 0296150, 9971674, 0002195 and 0072248. The authors thank the NSF for this support. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

The results of this paper were presented at the special sessions on tomography at the AMS Meetings in Binghamton, NY, USA in October 2003 and in Lawrenceville, NJ, USA in April 2004 and at the Inverse Problems workshop at IPAM in November 2003.

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

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

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

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

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

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

[13] 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

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

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

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

[17] Gindikin S 1995 Integral geometry on real quadrics *Lie groups and Lie algebras: E. B. Dynkin's Seminar (Am. Math. Soc. Transl. Ser. 2, vol 169)* (Providence, RI: American Mathematical Society) pp 23–31

[18] Helgason S 1980 *The Radon Transform* (Basel: Birkhäuser)
[19] Hörmander L 1983 *The Analysis of Linear Partial Differential Operators,* vol 1 (New York: Springer)
[20] John F 1971 *Plane Waves and Spherical Means, Applied to Partial Differential Equations* (New York: Dover)
[21] Kruger R A, Liu P, Fang Y R and Appledorn C R 1995 Photoacoustic ultrasound (PAUS) reconstruction tomography *Med. Phys.* **22** 1605–9
[22] Kuchment P 1993 Unpublished
[23] Ya V, Lin and Pinkus A 1993 Fundamentality of ridge functions *J. Approx. Theory* **75** 295–311
[24] Ya V, Lin 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
[25] 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
[26] Natterer F 1986 *The mathematics of Computerized Tomography* (New York: Wiley)
[27] Natterer F and Wübbeling F 2001 Mathematical methods in image reconstruction *Monographs on Mathematical Modeling and Computation,* vol 5 (Philadelphia, PA: SIAM)
[28] 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
[29] Norton S J 1980 Reconstruction of a two-dimensional reflecting medium over a circular domain: exact solution *J. Acoust. Soc. Am.* **67** 1266–73
[30] 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** 200–2
[31] Palamodov V P 2000 Reconstruction from limited data of arc means *J. Fourier Anal. Appl.* **6** 25–42
[32] Patch S K 2004 Thermoacoustic tomography—consistency conditions and the aprtial scan problem *Phys. Med. Biol.* **49** 1–11
[33] Quinto E T 1982 Null spaces and ranges for the classical and spherical Radon transforms *J. Math. Anal. Appl.* **90** 408–20
[34] 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
[35] Quinto E T 2004 Personal communication
[36] Spivak M 1999 *A Comprehensive Introduction to Differential Geometry* vol 3 (Houston: Publish or Perish)
[37] Xu M and Wang L-H V 2002 Time-domain reconstruction for thermoacoustic tomography in a spherical geometry *IEEE Trans. Med. Imaging* **21** 814–22
[38] Xu Y, Feng D and Wang L-H V 2002 Exact frequency-domain reconstruction for thermoacoustic tomography: I. Planar geometry *IEEE Trans. Med. Imaging* **21** 823–28
[39] Xu Y, Xu M and Wang L-H V 2002 Exact frequency-domain reconstruction for thermoacoustic tomography: II. Cylindrical geometry *IEEE Trans. Med. Imaging* **21** 829–33
[40] Xu Y, Wang L, Ambartsoumian G and Kuchment P 2004 Reconstructions in limited view thermoacoustic tomography *Med. Phys.* **31** 724–33
[41] Zobin N 1993 Private communication