Proc. Natl. Acad. Sci. USA
Vol. 75, No. 5, pp. 2055–2058, May 1978
Applied Mathematical Sciences

Mathematical aspects of divergent beam radiography

(x-ray transform/computed tomography)

K. T. Smith*†, D. C. Solmon*, S. L. Wagner‡, and C. Hamaker§

* Mathematics Department and ‡ Environmental Health Sciences Center, Oregon State University, Corvallis, Oregon 97331; and § Mathematics Department, University of Oregon, Eugene, Oregon 97403

Communicated by Peter D. Lax, January 16, 1978

ABSTRACT At the present time, largely because of a breakthrough in radiology called *computed tomography,* the attenuation of x-ray beams is measured in extremely sensitive quantitative ways, and the information from many x-ray sources is assembled and analyzed on a computer. In this situation mathematics can make significant contributions concerning the nature of the information conveyed by x-rays from many sources, the extent to which this information determines the object x-rayed, suitable configurations of sources, methods for using the data to build a detailed reconstruction of the object, etc. This article announces results on these topics for the divergent x-ray beam. The three-dimensionally divergent beam, or cone beam, presents new problems that do not appear in the two-dimensional, or fan beam, case.

Until recently there was little need for mathematics in radiology. Films were examined individually, and by eye, and mathematics had little to offer to the procedure. The picture changed radically in the late 1960s with a breakthrough called *computed tomography,* in which the attenuation of the x-ray beam is measured in an extremely sensitive quantitative way, and the information from many sources is assembled and analyzed on a computer (1). In this new situation mathematics can make significant contributions concerning the nature of the total information conveyed by x-rays from many sources, the extent to which this information determines the object x-rayed, suitable configurations of sources, methods for using the data to build a detailed reconstruction of the object, etc.

In the initial device for computed tomography, the celebrated EMI scanner (1), a parallel x-ray beam was used, and two-dimensional cross sections of the object were reconstructed. The mathematical theory of the parallel beam x-ray transform is developed in refs. 2 and 3. In the current second generation of scanners two-dimensional cross sections are still reconstructed, but a two-dimensionally divergent x-ray beam is used in place of the parallel beam in order to allow faster scan times.

At the present time there has arisen the need for dealing with a three-dimensionally divergent beam because of the extremely fast scan times required in the reconstruction of moving objects. A problem of major interest, for example, is the three-dimensional reconstruction of a beating heart. With human patients the x-ray data for such reconstructions must be collected during the fraction of a second in which heart movement is insignificant. This is impracticable with a succession of two-dimensionally divergent beams.

In the two-dimensional case, practical reconstruction formulas for the divergent beam transform have been obtained by falling back upon known formulas for the parallel beam transform (4, 5), but even in the two-dimensional case very little has been known of the required mathematical theory. In the three-dimensional case, even such formulas have not been known. Moreover, the three-dimensional case differs radically from the former. From a given source point it is feasible to x-ray full two-dimensional cross sections of a three-dimensional object, but it is normally out of the question to x-ray the full three-dimensional object itself. The beam must be coned down to the region of interest. Thus, the three-dimensional problem has a different nature, in that only partial information is available from each x-ray source.

The purpose of this note is to describe the beginnings of the required mathematical theory of the divergent beam x-ray transform. The specific matters addressed are as follows: *uniqueness*—a discussion of configurations of sources and cone beams for which the measured x-ray data determine the object uniquely; *reconstruction procedures*; and *measured regions*—a discussion of the severe limitations on the region x-rayed if each point in this region is to be seen from three or more sources. The field still abounds with open problems of both mathematical and practical importance.

MATHEMATICAL SETTING

The discussion is set in the n-dimensional space R^n so that separate statements are not needed for dimensions 2 and 3. Mathematically, the divergent beam x-ray transform or radiograph of a function f on R^n is the function $\mathcal{D}_a f$ defined by

$$\mathcal{D}_a f(\theta) = \int_0^\infty f(a + t\theta)dt \text{ for } \theta \in S^{n-1}, \qquad [1]$$

in which $S^{n-1} = \{x \in R^n: |x| = 1\}$ is the unit sphere in R^n. Physically, the function f is the density function of the object x-rayed, so that $\mathcal{D}_a f(\theta)$ is the total mass of the object along the half line with origin at the source a and direction θ. In practice this number is determined by measuring the attenuation of the x-ray beam along this half line. The basic problem is the extraction of information about the unknown density function f from information about certain of the radiographs $\mathcal{D}_a f$ from sources a lying in a source set A. Throughout the article it is assumed that the density function f is square integrable and that it vanishes outside a fixed bounded open set Ω with closure $\overline{\Omega}$ and closed convex hull $\hat{\Omega}$.

In the three-dimensional case the beam is usually coned down to a particular region of interest. For each source point a there is chosen a cone C_a with vertex a, and the attenuation is measured only along half lines in C_a. If

$$S_a = (C_a - a) \cap S^{n-1} \qquad [2]$$

is the corresponding set of directions, then the *measured x-ray data* consist of the numbers $\mathcal{D}_a f(\theta)$, for a in the source set A and

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

† Current address (until January 1979): Department of Physiology and Biophysics, Mayo Clinic, Rochester, MN 55901.

θ in the chosen set of directions S_a. A point in $C_a \cap \Omega$ is said to be *seen from the source a*. The set

$$\Omega_m = \bigcup_{a \in A} (C_a \cap \Omega), \qquad [3]$$

consisting of points seen from at least one source, is the *measured region*. Outside the measured region, alterations in the density function produce no changes in the measured x-ray data.

The subject of this note is the extraction of information about the unknown density function from the measured x-ray data.

MATHEMATICAL RESULTS

As stated above, it is assumed throughout that the density function f is square integrable and that it vanishes outside a fixed bounded open set Ω with closure $\overline{\Omega}$ and closed convex hull $\hat{\Omega}$. Theorems additional to the ones stated, and all of the proofs, are given in ref. 6.

Uniqueness. The basic uniqueness problem is to determine the configurations of source sets A and corresponding sets of directions S_a (or cones C_a) which guarantee that, if two density functions have identical measured x-ray data, then the functions themselves must be identical on some region of interest Ω_0. The region Ω_0 must be contained in the measured region Ω_m, because the functions and the data are independent outside Ω_m. When the full object can be x-rayed from each source there is a very strong uniqueness theorem.

THEOREM 1. (*Full information from each source*) *Let* A *be any infinite set of sources bounded away from* $\hat{\Omega}$. *If* $\mathcal{D}_a f(\theta) = \mathcal{D}_a g(\theta)$ *for each source* a *and each direction* θ, *then* f = g *on* $\Omega_m = \Omega$.

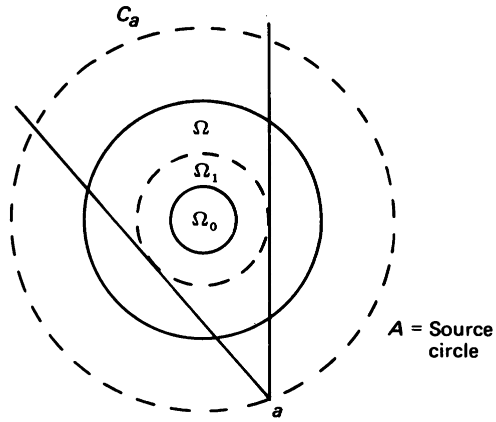
When only part of the object can be x-rayed from each source, as is usually the case in dimension 3, the situation is peculiar. Two examples are given in Figs. 1 and 2. The basic uniqueness problem is not solved in general, but the following theorem provides a special case that is feasible in practice. Figs. 2 and 3 contain two- and three-dimensional examples.

THEOREM 2. (*Partial information from each source*) *Let* C *be an open cone with vertex 0, and let* $\overline{A}$, *the closure of the source set* A, *be a rectifiable arc outside* $\overline{\Omega}$ *such that for each half line* ℓ *in* C *there is a point* $a \in \overline{A}$ *for which* $a + \ell$ *misses* $\overline{\Omega}$. *Let* $S_a = C \cap S^{n-1}$. *If* $\mathcal{D}_a f(\theta) = \mathcal{D}_a g(\theta)$ *for each source* a *and each direction* $\theta \in S_a$, *then* f = g *on* $\Omega_m = (A + C) \cap \Omega = (\overline{A} + C) \cap \Omega$.

The relevant situation for practice is of course that of a *finite* number of sources. For parallel beam x-rays it has been shown that objects with identical shapes and identical radiographs from any given finite number of directions can still differ in a *completely arbitrary* way on any interior compact set (3). At the present time somewhat weaker theorems are known for divergent beam x-rays (6), but these weaker theorems still imply a very high degree of nonuniqueness (and the stronger analog must certainly be true). Practical implications of the interplay between uniqueness and nonuniqueness are discussed briefly below.

Reconstruction. The reconstruction problem is the problem of giving a constructive procedure for recovering the unknown density function f on the region of interest from its measured x-ray data. Because of the nonuniqueness, such reconstructions cannot be achieved exactly with a finite number of sources, even in theory. However, because the uniqueness theorems require only a countably infinite set of sources, they provide the possibility of arbitrarily good approximations.

The most convenient setting for the problem is the Hilbert space $L_0^2(\Omega)$ composed of square integrable functions vanishing



$$f(x) = \int_{|x|}^{\infty} \rho'(s)(s^2 - |x|^2)^{-1/2} ds$$

FIG. 1. Simulated cross section of the chest Ω with the heart Ω_0 the region of interest. The source set A is the entire exterior circle. For each source a, the measured cone C_a is the cone determined by a and the circle Ω_1. The measured region Ω_m is all of Ω. If Ω and Ω_1 are circles centered at 0 with radii r and r_1, and ρ is any function of one variable that vanishes outside the interval (r_1, r), then the function f defined above has 0 measured x-ray data, and there is a high degree of nonuniqueness, even on the region Ω_0.

outside Ω. If the source set A lies outside $\overline{\Omega}$, then the set

$$N_A = \{u \in L_0^2(\Omega) : \mathcal{D}_a u(\theta) = 0 \text{ for } a \in A \text{ and } \theta \in S_a\}$$

is a closed subspace of this Hilbert space. The closed plane $f + N_A$ consists of the functions in $L_0^2(\Omega)$ with the same measured x-ray data as f. Let P_A be the orthogonal projection on this plane.

In any situation where uniqueness holds on the region of interest (e.g., *Theorems 1* and *2*) the plane $f + N_A$ contains only functions identical to f on this region, and for any $g \in L_0^2(\Omega)$, $P_A g$ is therefore identical to f on the region of interest. Consequently, the reconstruction problem can be reinterpreted as the problem of calculating the projection operator P_A. If the source set A is the union of an increasing sequence $\{A_j\}$, then

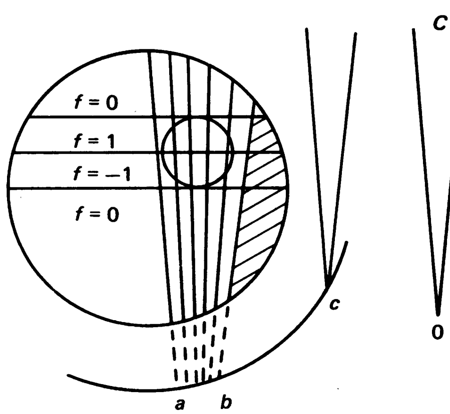


FIG. 2. Simulated cross section of the chest (large circle) with the heart (small circle) the region of interest. For each source a, the measured cone C_a is the translate $a + C$ of the fixed cone C. The source set A is either the arc ab or the arc ac. In the first case the measured region is the region shaded by lines originating on ab, and in the second it is the entire shaded region. In the first case the density function is not uniquely determined on the region of interest, as is shown by the illustrated function f with 0 measured x-ray data. In the second case the density function is uniquely determined, in accordance with *Theorem 2*. It is curious that adjunction of the seemingly irrelevant arc bc entails uniqueness.

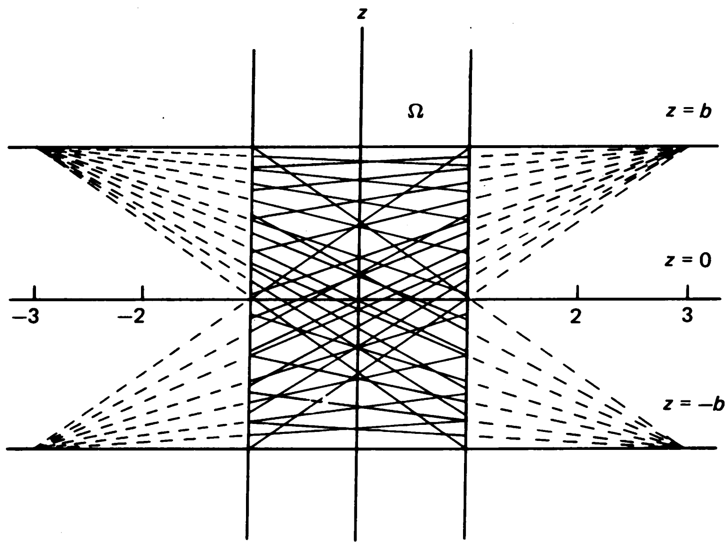


FIG. 3. Vertical plane section through the origin of the two-circle x-ray setup. The three-dimensional picture is obtained by revolving this one around the z axis. Inspection shows clearly what the measured region would be, and which points in it would be seen from just one source, if only one of the two source circles were used.

P_A is the limit of the P_{A_j}, i.e.,

$$\text{For each } g \in L_0^2(\Omega),\ P_{A_j}g \to P_A g \text{ in } L_0^2(\Omega), \qquad [4]$$

as is well known in Hilbert space theory. Thus, the approximation of P_A is reduced to the approximation of the P_{A_j}. When A is a countable set the A_j can be taken to be finite sets, and approximations can be obtained from finite source sets any time the union provides uniqueness on the region of interest.

Now let $A = \{a_1, \cdots, a_M\}$ be a finite set of sources. A constructive procedure for approximating P_A in terms of the individual projections P_{a_j} is given by a theorem of Halperin (7):

$$\text{If } Q = P_{a_M} \cdots P_{a_1}, \text{ then for any } g \in L_0^2(\Omega),\ Q^m g \to P_A g \text{ in } L_0^2(\Omega). \qquad [5]$$

Again this is a theorem in abstract Hilbert space adapted to the present situation.

What remains is the calculation of the projection P_a corresponding to an individual source a. For this an explicit formula can be given. Let $\Omega_{a,\theta} = \{t > 0 : a + t\theta \in \Omega\}$ if $\theta \in S_a$, and $\Omega_{a,\theta} = \phi$ if $\theta \notin S_a$, and set

$$\mu(a, \theta) = \int_{\Omega_{a,\theta}} t^{1-n} dt,$$

and

$$c(a, \theta) = (\mathcal{D}_a f(\theta) - \mathcal{D}_a g(\theta))/\mu(a, \theta) \quad \text{if } \theta \in S_a$$
$$c(a, \theta) = 0 \quad \text{if } \theta \notin S_a.$$

THEOREM 3. *For* $g \in L_0^2(\Omega)$, $P_a g(a + t\theta) = g(a + t\theta) + c(a, \theta)t^{1-n}$ *on* Ω.

The procedure of choosing an initial guess g and improving it successively by iteration of the projection formula in *Theorem 3* is called the Kacmarz procedure. An analogous procedure for solving finite systems of linear equations was introduced by Kacmarz in ref. 8. The parallel beam analog was the reconstruction procedure used in the original EMI scanner (1, 3).

Regions Measured from Multiple Sources. With the computed tomography scanners in current use, each point of the measured region is seen from a great many sources, usually between 150 and 600, depending on the size of the reconstruction matrix. In the two-dimensional case this does not pose a problem. In the three-dimensional case, however, when the number of sources is finite and each point of the measured region is to be seen from three or more sources, the possible configurations of sources and measured regions become very limited. In the sample theorem below it is assumed that the dimension is 3, that the number of sources is finite, and that the boundary of Ω_m satisfies a very mild smoothness condition which is always satisfied in practice.

THEOREM 4. (a) *If each point of* Ω_m *is seen from four noncoplanar sources, then* Ω_m *is a union of components of* Ω; *in particular,* $\Omega_m = \Omega$ *if* Ω *is connected.*

(b) *Suppose that the sources lie in a plane* π, *and let* H *be either of the half spaces determined by* π. *If each point of* Ω_m *is seen from three noncolinear sources, then* $\Omega_m \cap H$ *is a union of components of* $\Omega \cap H$; *in particular,* $\Omega_m \cap H = \Omega \cap H$ *if* $\Omega \cap H$ *is connected.*

Any open set in R^n separates into "connected pieces" called components; two points belong to the same component if they can be joined by a path that lies in the given open set. To illustrate the theorem, take, as in the case of principal interest, Ω to be the human body. Each point of Ω must lie in the component containing the heart, for, by definition points outside this component could not receive blood. In accordance with part a, if each point of the measured region is to be seen from four noncoplanar sources, then the measured region *must* be the entire body. To illustrate part b, suppose that the sources lie in a horizontal plane approximately through the midsection, and let H_+ be the half space above this plane, and H_- the half space below, and assume, as in part b, that each point of the measured region is to be seen from three noncolinear sources. $H_+ \cap \Omega$ has just one component, so if any part of the body above the source plane is to be measured, then *all* of it must be measured. $H_- \cap \Omega$ has three components (coming from the torso and legs, and the two arms), and if any part of one of these components is to be measured, then the entire component must be measured. (In practice Ω is usually a cylinder containing the body, and all sets have just one component.)

In the case of colinear sources, similar results hold with the half space H replaced by wedges bounded by planes through the source line.

DISCUSSION

Uniqueness. Although the nonuniqueness theorems show that, even in theory, exact reconstructions cannot be achieved with a finite number of sources, the uniqueness theorems, involving only a countable number of sources, show that arbitrarily good approximations can be. However, the infinite sequence $\{P_{A_j}\}$ called for in statement 4 (*Reconstruction* section) cannot be computed. Some fixed P_{A_j} must be used. Partly on theoretical and partly on empirical grounds, a reasonable choice for j can be determined. Nevertheless, no matter how large j is, the nonuniqueness theorems show that $P_{A_j}g$ can differ dramatically from f.

In practice the x-ray setup probably should be designed so that in the limiting case of infinitely many sources uniqueness does hold, and then *a priori* information not coming from x-rays should be introduced into the reconstruction method in order to combat the nonuniqueness inherent in the finitely many sources actually used. Examples of the use of *a priori* information can be found in refs. 3, 9, and 10, but much remains to be done.

On the other hand, if very fast reconstructions are needed (e.g., if it is desirable that the physician participate in the process), it may be useful to compromise with uniqueness. In particular, an x-ray setup like the one illustrated in Fig. 1 is attractive because it requires manipulation of a minimum amount of data, and, indeed, similar setups are now being tried in computed tomography algorithms (11). With the setup of Fig. 1 and the reconstruction $f_0 = P_A 0$ (the function with correct measured data and minimum L^2 norm), the error function $u = f - f_0$ satisfies

$$\max_{x \in \Omega_0} |u(x) - u(0)| \le C\|f\|_{L^2}. \qquad [6]$$

Preliminary computations indicate that with reasonable ratios of the radii of Ω, Ω_1, and Ω_0 (e.g., 4:2:1), the constant C is very small. If this is the case, then the error function u is nearly constant on the region of interest and the lack of uniqueness does little damage to the reconstruction on this region.

Reconstruction. It is easy to give explicit formulas for the inverse of the divergent beam x-ray transform (6). The known formulas, however, require the knowledge of $\mathcal{D}_a f(\theta)$ for all sources a on the surface of a sphere surrounding the object and all directions $\theta \in S^{n-1}$—in practice, therefore, the knowledge of $\mathcal{D}_a f(\theta)$ for a finite number of sources and directions more or less uniformly distributed over these spheres. In most two-dimensional problems this is a satisfactory situation, and numerical implementations of the inversion formulas are now being used in many of the computed tomography scanners. [Convolution method (4, 5).]

In three dimensions, however, it is usually necessary to confine the sources to a one-dimensional set, such as an arc or union of arcs, and to cone the beam down. In this situation inversion formulas are not known, even in dimension 2, and stable, practical formulas may well not exist. The Kacmarz method, in a sense, is independent of the distribution of sources and measured cones. For any finite set A of sources and measured cones it always leads to a function $P_A g$ with the correct measured x-ray data, though with bad distribution the convergence is slow. (See ref. 12.) The Kacmarz method also lends itself to the incorporation of *a priori* information. A disadvantage, however, is that iterative methods apparently are intrinsically slower than noniterative methods—which was a principal reason for the switch to the convolution method in the current scanners. A noniterative inversion procedure, though not an explicit formula, is contained implicitly in the proof of the second uniqueness theorem. It is not known whether this procedure can be implemented effectively in practice.

Regions Measured from Multiple Sources. In the computed tomography scanners each point of the measured region is seen from a great many sources. The extent to which this is necessary for results of uniform high quality has not been determined, and probably can be determined only by experience. If each point of the measured region is to be seen from three or more sources, then in three dimensions the possible source configurations are very limited:

(*a*) Colinear sources are possible. In this case the measured region must be effectively the intersection of Ω with a wedge bounded by planes through the source line. Uniqueness holds, but the problem is likely to be rather ill conditioned.

(*b*) Coplanar sources are possible only if it is feasible to measure all of Ω on one side of the source plane, or if the sources are effectively colinear, e.g., on a polygonal arc. There is no evidence available as to whether a plane polygonal arc of sources can produce a well-conditioned problem.

(*c*) Strongly noncoplanar sources, configurations in which each point of the measured region is seen from four noncoplanar sources, are likely to produce the best-conditioned problems, but they are possible only when it is possible to measure the entire region Ω.

A practical configuration—in which uniqueness holds, the measured region is reasonable and each point is seen from many sources, and the problem is probably well conditioned—may be the configuration in which the sources are distributed around two circles. In the typical three-dimensional reconstruction problem, Ω is effectively a vertical cylinder, and the region of interest Ω_0 is a slice cut off by the planes $z = \pm b$. Let C be the cone of half lines in the upper half space $z > 0$ with initial point 0 and making an angle $\alpha = \arctan(b/2)$ with the horizontal x, y plane. If the radius of the cylinder Ω is 1, let the source set A consist of the two circles with center 0 and radius 3 in the planes $z = \pm b$. For sources a in the upper circle, let $C_a = a - C$; and for sources a in the lower circle, let $C_a = a + C$. With this configuration the measured region is the slice Ω_0 of interest, uniqueness holds on this region, and each point is seen either from all sources on the upper circle or from all sources on the lower circle. [The same effect is achieved if the radius of the source circles is any number $r \ge 3$ and the angle α satisfies $b/(r - 1) \le \tan \alpha \le 2b/(r + 1)$.] The two-circle x-ray setup is illustrated in Fig. 3.

Figures were prepared by Mrs. Donna Balow of the Biodynamics Research Unit of the Mayo Clinic.

1. Hounsfield, G. N. (1973) *Br. J. Radiol.* **46,** 1011–1022.
2. Solmon, D. C. (1976) *J. Math. Anal. Appl.* **56,** 61–83.
3. Smith, K. T., Solmon, D. C. & Wagner, S. L. (1977) *Bull. Am. Math. Soc.* 1227–1270.
4. Lakshminarayanan, A. V. (1975) *Reconstruction from Divergent X-Ray Data*, SUNY technical report no. 92 (Computer Science Department, Buffalo, NY).
5. Herman, G. T., Lakshminarayanan, A. V. & Naparstek, A. (1976) in *Reconstruction Tomography in Diagnostic Radiology and Nuclear Medicine*, eds. Ter-Pergosian, M., Phelps, M. E., Brownell, G. L., Cox, J. R., Davis, D. O. & Evens, R. G. (University Park Press, Baltimore, MD), pp. 105–117.
6. Hamaker, C., Smith, K. T., Solmon, D. C. & Wagner, S. L. (1978) *Rocky Mt. J. Math.*, in press.
7. Halperin, I. (1962) *Acta Sci. Math.* **23,** 96–99.
8. Kacmarz, S. (1937) *Bull. Int. Acad. Pol. Sci. Lett. Cl. Sci. Math. Nat. Ser. A*, 355–357.
9. Hounsfield, G. N. (1976) *Am. J. Roentgenol.* **127,** 3–9.
10. Herman G. T. & Lent A. (1976) *Comput. Biol. Med.* **6,** 273–294.
11. Herman, G. T. & Naparstek, A. (1977) *SIAM J. Appl. Math.* **33,** 511–533.
12. Hamaker, C. & Solmon, D. C. (1978) *J. Math. Anal. Appl.* **62,** 1–23.