X-ray Tomography

Frank Natterer

Institut für Numerische und Angewandte Mathematik, University of Münster,
Einsteinstraße 62, 48149 Münster, Germany
natterer@math.uni-muenster.de

Summary. We give a survey on the mathematics of computerized tomography. We start with a short introduction to integral geometry, concentrating on inversion formulas, stability, and ranges. We then go over to inversion algorithms. We give a detailed analysis of the filtered backprojection algorithm in the light of the sampling theorem. We also describe the convergence properties of iterative algorithms. We shortly mention Fourier based algorithms and the recent progresses made in their accurate implementation. We conclude with the basics of algorithms for cone beam scanning which is the standard scanning mode in present days clinical practice.

1 Introduction

X-ray tomography, or computerized tomography (CT) is a technique for imaging cross sections or slices (slice = $\tau o \mu o s$ (greek)) of the human body. It was introduced in clinical practice in the 70s of the last century and has revolutionized clinical radiology. See Webb (1990), Natterer and Ritman (2002) for the history of CT.

The mathematical problem behind CT is the reconstruction of a function f in $\mathbb{R}^2$ from the set of its line integrals. This is a special case of integral geometry, i.e. the reconstruction of a function from integrals over lower dimensional manifolds. Thus, integral geometry is the backbone of the mathematical theory of CT.

CT was soon followed by other imaging modalities, such as emission tomography (single particle emission tomography (SPECT) and positron emission tomography (PET)), magnetic resonance imaging (MRI). CT also found applications in various branches of science and technology, e.g. in seismics, radar, electron microscopy, and flow. Standard references are Herman (1980), Kak and Slaney (1987), Natterer and Wübbeling (2001).

In these lectures we give an introduction into the mathematical theory and the reconstruction algorithms of CT. We also discuss matters of resolution and stability. Necessary prerequisites are calculus in $\mathbb{R}^n$, Fourier transforms, and the elements of sampling theory.

Since the advent of CT many imaging techniques have come into being, some of them being only remotely related to the straight line paradigm of CT, such as impedance tomography, optical tomography, and ultrasound transmission tomography. For the study of these advanced technique a thorough understanding of CT is at least useful. Many of the fundamental tools and issues of CT, such as backprojection, sampling, and high frequency analysis, have their counterparts in the more advanced techniques and are most easily in the framework of CT.

The outline of the paper is as follows. We start with a short account of the relevant parts of integral geometry, with the Radon transform as the central tool. Then we discuss in some detail reconstruction algorithms, in particular the necessary discretizations for methods based on inversion formulas and convergence properties of iterative algorithm. Finally we concentrate on the 3D case which is the subject of current research.

2 Integral Geometry

In this section we introduce the relevant integral transforms, derive inversion formulas, and study the ranges. For a thorough treatment see Gelfand, Graev, and Vilenkin (1965), Helgason (1999), Natterer (1986).

2.1 The Radon Transform

Let f be a function in $\mathbb{R}^n$. To avoid purely technical difficulties we assume f to be smooth and of compact support, if not said otherwise.

For $\theta \in S^{n-1}$, $s \in \mathbb{R}^1$ we define

$$(Rf)(\theta, s) = \int_{x\cdot\theta=s} f(x)dx \; . \tag{1}$$

R is the Radon transform. dx is the restriction of the Lebesgue measure in $\mathbb{R}^n$ to $x \cdot \theta = s$. Other notations are

$$(Rf)(\theta, s) = \int_{\theta^\perp} f(s\theta + y)dy \tag{2}$$

with $\theta^\perp$ the subspace of $\mathbb{R}^n$ orthogonal to θ, and

$$(Rf)(\theta, s) = \int \delta(x \cdot \theta - s)f(x)dx \tag{3}$$

with δ the Dirac δ-function.

The Radon transform is closely related to the Fourier transform

$$\hat{f}(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)dx \; .$$

Namely, we have the so-called projection-slice theorem.

Theorem 2.1.

$$(Rf)^\wedge(\theta,\sigma) = (2\pi)^{(n-1)/2}\hat{f}(\sigma\theta)\ .$$

Note that the Fourier transform on the left hand side is the 1D Fourier transform of Rf with respect to its second variable, while the Fourier transform on the right hand side is the nD Fourier transform of f.

For the proof of Theorem 2.1 we make use of the definition (2) of Radon transform, yielding

$$\begin{aligned}(Rf)^\wedge(\theta,\sigma) &= (2\pi)^{-1/2}\int\limits_{\mathbb{R}^1} e^{-is\sigma}(Rf)(\theta,s)ds \\ &= (2\pi)^{-1/2}\int\limits_{\mathbb{R}^1} e^{-is\sigma}\int\limits_{\theta^\perp} f(s\theta+y)dyds\ .\end{aligned}$$

Putting $x = s\theta + y$, i.e. $s = x\cdot\theta$, we have

$$\begin{aligned}(Rf)^\wedge(\theta,s) &= (2\pi)^{-1/2}\int\limits_{\mathbb{R}^n} e^{-i\sigma x\cdot\theta}f(x)dx \\ &= (2\pi)^{(n-1)/2}\hat{f}(\sigma\theta)\ .\end{aligned}$$

The proofs of many of the following results follow a similar pattern. So we omit proofs unless more sophisticated tools are needed.

It is possible to extend R to a bounded operator $R : L_2(\mathbb{R}^n) \to L_2(S^{n-1}\times\mathbb{R}^1)$. As such R has an adjoint $R^* : L_2(S^{n-1}\times\mathbb{R}^1) \to L_2(\mathbb{R}^n)$ which is easily seen to be

$$(R^*g)(x) = \int\limits_{S^{n-1}} g(\theta, x\cdot\theta)d\theta\ . \tag{4}$$

R^* is called the backprojection operator in the imaging literature.

We also will make use of the Hilbert transform

$$(Hf)(s) = \frac{1}{\pi}\int\frac{f(t)}{s-t}dt \tag{5}$$

which, in Fourier domain, is given by

$$(Hf)^\wedge(\sigma) = -i\ sgn(\sigma)\hat{f}(\sigma) \tag{6}$$

with $sgn(\sigma)$ the sign of the real number σ. Now we are ready to state and to prove Radon's inversion formula:

Theorem 2.2. *Let $g = Rf$. Then*

$$f = \frac{1}{2}(2\pi)^{1-n}R^*H^{n-1}g^{(n-1)}$$

where $g^{(n-1)}$ stands for the derivative of order $n-1$ of g with respect to the second argument.

For the proof we write the Fourier inversion formula in polar coordinates and make use of Theorem 2.1 and the evenness of g, i.e. $g(\theta, s) = g(-\theta, -s)$.

Special cases of Theorem 2.2 are

$$f(x) = \frac{1}{4\pi^2} \int_{S^1} \int_{\mathbb{R}^1} \frac{g'(\theta, s)}{x \cdot \theta - s} ds d\theta \tag{7}$$

for $n = 2$ and

$$f(x) = -\frac{1}{8\pi^2} \int_{S^2} g''(\theta, x \cdot \theta) d\theta \tag{8}$$

for $n = 3$. Note that there is a distinctive difference between (7) and (8): The latter one is local in the sense that for the reconstruction of f at x_0 only the integrals $g(\theta, s)$ over those planes $x \cdot \theta = s$ are needed that pass through x_0 or nearby. In contrast, (7) is not local in this sense, due to the integral over R^1.

In order to study the stability of the inversion process we introduce the Sobolev spaces H^α on $\mathbb{R}^n$, $S^{n-1} \times \mathbb{R}^1$ with norms

$$\|f\|^2_{H^\alpha(\mathbb{R}^n)} = \int_{\mathbb{R}^n} (1 + |\xi|^2)^\alpha |\hat{f}(\xi)|^2 d\xi \,,$$

$$\|g\|^2_{H^\alpha(S^{n-1} \times \mathbb{R}^1)} = \int_{S^{n-1}} \int_{\mathbb{R}^1} (1 + \sigma^2)^\alpha |\hat{g}(\theta, \sigma)|^2 d\sigma d\theta \,.$$

As in Theorem 2.1, the Fourier transform $\hat{g}$ is the 1D Fourier transform of g with respect to the second argument.

Theorem 2.3. *There exist constants c, $C > 0$ depending only on α, n, such that*

$$c\|f\|_{H^\alpha(\mathbb{R}^n)} \leq \|Rf\|_{H^{\alpha+(n-1)/2}(S^{n-1} \times \mathbb{R}^1)} \leq C\|f\|_{H^\alpha(\mathbb{R}^n)}$$

for $f \in H^\alpha(\mathbb{R}^n)$ with support in $|x| < 1$.

The conclusion is that Rf is smoother than f by the order $(n-1)/2$. This indicates that the reconstruction process, i.e. the inversion of R, is slightly ill-posed.

Finally we study the range of R.

Theorem 2.4. *For $m \geq 0$ an integer let*

$$p_m(\theta) = \int_{\mathbb{R}^1} s^m (Rf)(\theta, s) ds \,.$$

Then, p_m is a homogeneous polynomial of degree m in θ.

The proof is by verification. The question wether or not the condition of Theorem 2.4 is sufficient for a function g of θ, s being in the range of R is the subject of the famous Helgason–Ludwig theorem.

Theorem 2.4 has applications to cases in which the data function g is not fully specified.

2.2 The Ray Transform

For f a function in $\mathbb{R}^n$, $\theta \in S^{n-1}$, $x \perp \theta^\perp$ we define

$$(Pf)(\theta, x) = \int_{\mathbb{R}^1} f(x + t\theta)dt \tag{9}$$

P is called the ray (or X-ray) transform. The projection-slice theorem for P reads

Theorem 2.5.

$$(Pf)^\wedge(\theta, \xi) = (2\pi)^{1/2}\hat{f}(\xi), \xi \in \theta^\perp .$$

The Fourier transform on the left hand side is the $(n-1)D$ Fourier transform in $\theta^\perp$, while the Fourier transform on the right hand side is the nD Fourier transform in $\mathbb{R}^n$.

The adjoint ray transform, which we call backprojection again, is given by

$$(P^*g)(x) = \int_{S^{n-1}} g(\theta, E_\theta x)d\theta \tag{10}$$

with E_θ the orthogonal projection onto $\theta^\perp$, i.e. $E_\theta x = x - (x \cdot \theta)\theta$. We have the following analogue of Radon's inversion formula:

Theorem 2.6. *Let $g = Pf$. Then*

$$f = \frac{1}{2\pi|S^{n-1}|} P^* I^{-1} g$$

with the Riesz potential

$$(I^{-1}g)^\wedge(\xi) = |\xi|\hat{g}(\xi)$$

in $\theta^\perp$.

Theorem 2.6 is not as useful as Theorem 2.2. The reason is that for $n \geq 3$ (for $n = 2$ Theorems 2.2 and 2.6 are equivalent) the dimension $2(n-1)$ of the data function g is greater than the dimension n of f. Hence the problem of inverting P is vastly overdetermined. A useful formula would compute f from the values $g(\theta, \cdot)$ where θ is restricted to some set $S_0 \subseteq S^{n-1}$. Under which conditions on S_0 is f uniquely determined?

Theorem 2.7. *Let $n = 3$, and assume that each great circle on S^2 meets S_0. Then, f is uniquely determined by $g(\theta, \cdot)$, $\theta \in S_0$.*

The condition on S_0 in Theorem 2.7 is called Orlov's completeness condition. In the light of Theorem 2.5, the proof of Theorem 2.7 is almost trivial: Let $\xi \in \mathbb{R}^3 \setminus \{0\}$. According to Orlov's condition there exists $\theta \in S_0$ on the great circle $\xi^\perp \cap S^2$. For this θ, $\xi \in \theta^\perp$, hence

$$\hat{f}(\xi) = (2\pi)^{-1/2}\hat{g}(\theta, \xi)$$

by Theorem 2.5. Thus $\hat{f}$ is uniquely (and stably) determined by g on S_0.

An obvious example for a set $S_0 \subseteq S^2$ satisfying Orlov's completeness condition is a great circle. In that case inversion of P is simply 2D Radon inversion on planes parallel to S_0. Orlov gave an inversion formula for S_0 a spherical zone.

2.3 The Cone-Beam Transform

The relevant integral transform for fully 3D X-ray tomography is the cone beam transform in $\mathbb{R}^3$

$$(Cf)(a,\theta) = \int_0^\infty f(a+t\theta)dt \tag{11}$$

where $\theta \in S^2$. The main difference to the ray transform P is that the source point a is restricted to a source curve A surrounding the object to be imaged.

Of course the following question arises: Which condition on A guarantees that f is uniquely determined by $(Cf)(a,\cdot)$?

A first answer is: Whenever A has an accumulation point outside supp (f) (provided f is sufficiently regular). Unfortunately, inversion under this assumptions is based on analytic continuation, a process that is notoriously unstable. So this result is useless for practical reconstruction.

In order to find a stable reconstruction method (for suitable source curves A) we make use of a relationship between C and R discovered by Grangeat (1991):

Theorem 2.8.

$$\frac{\partial}{\partial s}(Rf)(\theta,s)\,|_{s=a\cdot\theta} = \int_{\theta^\perp \cap S^2} \frac{\partial}{\partial\theta}(Cf)(a,\omega)d\omega\ .$$

The notation on the right hand side needs explication: $(Cf)(a,\cdot)$ is a function on S^2. θ is a tangent vector to S^2 for each $\omega \in S^2 \cap \theta^\perp$. Thus the derivative $\frac{\partial}{\partial\theta}$ makes sense on $S^2 \cap \theta^\perp$.

Since Theorem 2.8 is the basis for present day's work on 3D reconstruction we sketch two proofs. The first one is completely elementary. Obviously it suffices to put $\theta = e_3$, the third unit vector, in which case it reads

$$\frac{\partial}{\partial s}(Rf)(\theta,a_3) = \int_{\omega\in S^1} \left[\frac{\partial}{\partial z}(Cf)\left(a, \begin{pmatrix}\omega\\ z\end{pmatrix}\right)\right]_{z=0} d\omega$$

where a_3 is the third component of the source point a. It is easy to verify that right- and left-hand side both coincide with

$$\int_{\mathbb{R}^2} \frac{\partial f}{\partial x_3}\left(a + \begin{pmatrix}x'\\ 0\end{pmatrix}\right) dx'\ .$$

The second proof makes use of the formula

$$\int_{S^{n-1}} (Cf)(a,\omega)h(\theta\cdot\omega)d\omega = \int_{\mathbb{R}^1} (Rf)(\theta,s)h(s-a\cdot\theta)ds$$

that holds – for suitable functions f in $\mathbb{R}^2$ – for h a function in $\mathbb{R}^1$ homogeneous of degree $1-n$; see Hamaker et al. (1980).

Putting $h=\delta'$ yields Theorem 2.8.

Theorem 2.9. *Let the source curve A satisfy the following condition: Each plane meeting supp (f) intersects A transversally. Then, f is uniquely (and stably) determined by $(Cf)(a,\theta)$, $a\in A$, $\theta\in S^2$.*

The condition on A in Theorem 2.9 is called the Kirillov–Tuy condition.

The proof of Theorem 2.9 starts out from Radon's inversion formula (8) for $n=3$:

$$f(x) = -\frac{1}{8\pi^2}\int_{S^2}\left[\frac{\partial^2}{\partial s^2}Rf(\theta,s)\right]_{s=x\cdot\theta} d\theta \tag{12}$$

Now assume A satisfies the Kirillov–Tuy condition, and let $x\cdot\theta=s$ be a plane hitting supp (f). Then there exists $a\in A$ such that $a\cdot\theta=s$, and

$$\begin{aligned}\left[\frac{\partial}{\partial s}Rf(\theta,s)\right]_{s=x\cdot\theta} &= \left[\frac{\partial}{\partial s}Rf(\theta,s)\right]_{s=a\cdot\theta}\\ &= \int_{\theta^\perp\cap S^2}\frac{\partial}{\partial\theta}(Cf)(a,\omega)d\omega\end{aligned}$$

is known by Theorem 2.8. Since A intersects the plane $x\cdot\theta=s$ transversally this applies also to the second derivative in (12). Hence the integral in (12) can be computed from the known values of Cf for all planes $x\cdot\theta=s$ hitting supp (f), and for the others it is zero.

Theorem 2.8 in conjunction with Radon's inversion formula (8) is the basis for reconstruction formulas for C with source curves A satisfying the Kirillov–Tuy condition. The simplest source curve one can think of, a circle around the object, does not satisfy the Kirillov–Tuy condition. For a circular source curve an approximate inversion formula, the FDK formula, exists; see Feldkamp et al. (1984). In medical applications the source curve is a helix.

2.4 The Attenuated Radon Transform

Let $n=2$. The attenuated Radon transform of f is defined to be

$$(R_\mu f)(\theta,s) = \int_{x\cdot\theta=s} e^{-(C_\mu)(x,\theta_\perp)}f(x)dx$$

where μ is another function in $\mathbb{R}^2$ and $\theta_\perp$ is the unit vector perpendicular to θ such that det $(\theta, \theta_\perp) = 1$. This transform comes up in SPECT, where f the (sought-for) activity distribution and μ the attenuation map.

R_μ admits an inversion formula very similar to Radon's inversion formula for R:

Theorem 2.10. *Let $g = R_\mu f$. Then,*

$$f = \frac{1}{4\pi} Re\ div\, R^*_{-\mu}(\theta e^{-h} H e^h g)$$

where

$$h = \frac{1}{2}(I + iH) R\mu \,,$$
$$(R^*_\mu g)(x) = \int_{S^1} e^{-(C\mu)(x,\theta_\perp)} g(\theta, x\cdot\theta) d\theta \,.$$

This is Novikov's inversion formula; see Novikov (2000). Note that the back-projection $R^*_{-\mu}$ is applied to a vector valued function. For $\mu = 0$ Novikov's formula reduces to Radon's inversion formula (7).

There is also an extension of Theorem 2.4 to the attenuated case:

Theorem 2.11. *Let $k > m \geq 0$ integer and h the function from Theorem 2.10. Then,*

$$\int_{\mathbb{R}^1}\int_{S^1} s^m e^{ik\varphi + h(\theta,s)} (R_\mu f)(\theta, s) d\theta ds = 0$$

where $\theta = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$.

The proofs of Theorems 2.10 and 2.11 are quite involved. They are both based on the following remark: Let

$$u(x,\theta) = h(\theta, x\cdot\theta) - (C\mu)(x, \theta_\perp) \,.$$

Then, u admits a Fourier series representation as

$$u(x,\theta) = \sum_{\ell>0 \text{ odd}} u_\ell(x) e^{i\ell\varphi}$$

with certain functions $u_\ell(x)$ and $\theta = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$.

Of course this amounts to saying that $u(x,\cdot)$ admits an analytic continuation from S^1 into the unit disk.

2.5 Vectorial Transforms

Now let f be a vector field. The scalar transforms P, R can be applied componentwise, so Pf, Rf make sense. They give rise to the vectorial ray transform

$$(\mathcal{P}f)(\theta,s) = \theta \cdot (Pf)(\theta,s)\ , \quad x \in \theta^\perp$$

and the Radon normal transform

$$(\mathcal{R}^\perp f)(\theta,s) = \theta \cdot (Rf)(\theta,s)\ .$$

Obviously, vectorial transforms can't be invertible. In fact their null spaces are huge.

Theorem 2.12. *$\mathcal{P}f = 0$ if and only if $f = \nabla\psi$ for some ψ.*

The nontrivial part of the proof makes use of Theorem 2.5: If $\mathcal{P}f = 0$, then, for $\xi \in \theta^\perp$

$$(\mathcal{P}f)^\wedge(\theta,\xi) = \theta \cdot (Pf)^\wedge(\theta,\xi) = (2\pi)^{1/2}\theta \cdot \hat{f}(\xi) = 0\ .$$

Hence $\theta^\perp$ is invariant under $\hat{f}$. This is only possible if $\hat{f}(\xi) = i\hat{\psi}(\xi)\xi$ with some $\hat{\psi}(\xi)$. It remains to show that $\hat{\psi}$ is sufficiently regular to conclude that $f = \nabla\psi$.

A similar proof leads to

Theorem 2.13. *Let $n = 3$. $\mathcal{R}^\perp f = 0$ if and only if $f =$ curl ϕ for some ϕ.*

Inversion formulas can be derived for those parts of the solution that are uniquely determined. We make use of the Helmholtz decomposition to write a vector f in the form

$$f = f^s + f^i$$

where f^s is the solenoidal part (i.e. div $f^s = 0$ or $f^s =$ curl ϕ) and f^i is the irrotational part (i.e. curl $f^i = 0$, $f^i = \nabla\psi$). Then, Theorem 2.12 states that the solenoidal part of f is uniquely determined by $\mathcal{P}f$, while the irrotational part of f is uniquely determined by $\mathcal{R}^\perp f$.

We have the following inversion formula:

Theorem 2.14. *Let $g = \mathcal{P}f$ and let f be solenoidal (i.e. $f^i = 0$). Then,*

$$f = \frac{n-1}{2\pi|S^{n-2}|}\,\mathcal{P}^* I^{-1} g\ ,$$

$$(\mathcal{P}^* g)(x) = \int_{S^{n-1}} g(\theta, E_\theta x)\theta d\theta$$

and I^{-1} is the Riesz potential already used in Theorem 2.6.

For more results see Sharafutdinov (1994).

3 Reconstruction Algorithms

There are essentially three classes of reconstruction algorithms. The first class is based on exact or approximate inversion formulas, such as Radon's inversion formula. The prime example is the filtered backprojection algorithm. It's not only the work horse in clinical radiology, but also the model for many algorithms in other fields. The second class consists of iterative methods, with Kaczmarz's method (called algebraic reconstruction technique (ART) in tomography) as prime example. The third class is usually called Fourier methods, even though Fourier techniques are used also in the first class. Fourier methods are direct implementations of the projection slice theorem (Theorems 2.1, 2.5) without any reference to integral geometry.

3.1 The Filtered Backprojection Algorithm

This can be viewed as an implementation of Radon's inversion formula (Theorem 2.2). However it is more convenient to start out from the formula

Theorem 3.1. *Let $V = R^*v$. Then,*

$$V * f = R^*(v * g)$$

where the convolution on the left hand side is in $\mathbb{R}^n$, i.e.

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

and the convolution on the right hand side is in $\mathbb{R}^1$, i.e.

$$(v * g)(\theta, s) = \int_{\mathbb{R}^1} v(\theta, s-t)g(\theta, t)dt \ .$$

The proof of Theorem 3.1 require not much more than the definition of R^*.

Theorem 3.1 is turned into an (approximate) inversion formula for R by choosing $V \sim \delta$. Obviously V is the result of the reconstruction procedure for $f = \delta$. Hence V is the point spread function, i.e. the function the algorithm would reconstruct if the true f were the δ-function.

Theorem 3.2. *Let $V = R^*v$ with v independent of θ. Then,*

$$\hat{V}(\xi) = 2(2\pi)^{(n-1)/2}|\xi|^{1-n}\hat{v}(|\xi|) \ .$$

For V to be close to δ we need $\hat{V}$ to be close to $(2\pi)^{-\frac{n}{2}}$, hence

$$\hat{v}(\sigma) \sim \frac{1}{2}(2\pi)^{-n+1/2}\sigma^{n-1} \ .$$

Let ϕ be a low pass filter, i.e. $\phi(\sigma) = 0$ for $|\sigma| > 1$ and, for convenience, ϕ even. Let Ω be the (spatial) bandwidth. By Shannon's sampling theorem Ω corresponds to the (spatial) resolution $2\pi/\Omega$:

Theorem 3.3. *Let f be Ω-bandlimited, i.e. $\hat{f}(\xi) = 0$ for $|\xi| > \Omega$. Then, f is uniquely determined by $f(h\ell)$, $\ell \in \mathbb{Z}^n$, $h \le \pi/\Omega$. Moreover, if $h \le 2\pi/\Omega$, then*

$$\int_{\mathbb{R}^n} f(x)dx = h^n \sum_\ell f(h\ell) \ .$$

An example for a Ω-bandlimited function in $\mathbb{R}^1$ is $f(x) = \mathrm{sinc}(\Omega x)$ where $\mathrm{sinc}(x) = \sin(x)/x$ is the so-called sinc-function. From looking at the graph of f we conclude that bandwidth Ω corresponds to resolution $2\pi/\Omega$.

We remark that Theorem 3.3 has the following corollary:
Assume that f, g are Ω-bandlimited. Then,

$$\int_{\mathbb{R}^n} f(x)g(x)dx = h^n \sum_\ell f(h\ell)g(h\ell)$$

provided that $h \le \pi/\Omega$. This is true since fg has bandwidth 2Ω.

We put

$$\hat{v}(\sigma) = 2(2\pi)^{(n-1)/2}|\sigma|^{n-1}\phi(\sigma/\Omega) \ .$$

v is called the reconstruction filter. As an example we put $n = 2$ and take as ϕ the ideal low pass, i.e.

$$\phi(\sigma) = \begin{cases} 1 \ , & |\sigma| \le 1 \ , \\ 0 \ , & |\sigma| > 1 \ . \end{cases}$$

Then,

$$v(s) = \frac{\Omega^2}{4\pi^2}u(\Omega s) \ , \quad u(s) = \mathrm{sinc}(s) - \frac{1}{2}\left(\mathrm{sinc}\left(\frac{s}{2}\right)\right)^2 \ .$$

Now we describe the filtered backprojection algorithm for the reconstruction of the function f from $g = Rf$. We assume that f has the following properties:

$$f \text{ is supported in } |x| \le \rho \tag{13}$$

$$f \text{ is essentially } \Omega\text{-bandlimited, i.e.} \tag{14}$$

$$\hat{f}(\xi) \text{ is negligible in some sense for } |\xi| > \Omega \ .$$

Since f is essentially Ω-bandlimited we conclude from Theorem 2.1 that the same is true for g. Hence, by a loose application of the sampling theorem,

$$(v * g)(\theta, s_\ell) = \Delta s \sum_k v(\theta, s_\ell - s_k)g(\theta, s_k) \tag{15}$$

where $s_\ell = \ell\Delta s$ and $\Delta s \le \pi/\Omega$. Equation (15) is the filtering step in the filtered backprojection algorithm. In the backprojection step we compute $R^*(v * g)$ in a discrete setting. We restrict ourselves to the case $n = 2$, i.e.

$$(R^*(v * g))(x) = \int_{S^1} (v * g)(\theta, x \cdot)d\theta \ .$$

In order to discretize this integral properly we make use of Theorem 3.3 again. For this we have to determine the bandwidth of the function $\varphi \mapsto (v * g)(\theta, x \cdot \theta)$ where $\theta = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$.

Theorem 3.4. *For k an integer we have*

$$\int_0^{2\pi} e^{-ik\varphi}(v * g)(\theta, x \cdot \theta)d\varphi = \frac{i^k}{2\pi} \int_{|y|<\rho} f(y)e^{-ik\psi} \int_{-\Omega}^{\Omega} \hat{v}_k(\sigma)J_k(-\sigma|x-y|)d\sigma dy$$

where $\psi = \arg(y)$ and J_k is the Bessel function of order k of the first kind.

All that is needed for the proof is the integral representation

$$J_k(s) = \frac{i^{-k}}{2\pi} \int_0^{2\pi} e^{-ik\varphi + s\cos\varphi} d\varphi$$

of the Bessel function.

By Debye's asymptotic relation (see Abramowitz and Stegun (1970)), $J_k(s)$ is negligibly small for $|s| < |k|$ and $|k|$ large. Hence the left hand side in Theorem 3.4 is negligible for $2\Omega\rho < |k|$. In other words, the function $\varphi \mapsto (v * g)(\theta, x \cdot \theta)$ has essential bandwidth $2\Omega\rho$. According to Theorem 3.3,

$$(R^*(v * g))(x) = \Delta\varphi \sum_j (v * g)(\theta_j, x \cdot \theta_j) \tag{16}$$

where $\theta_j = \begin{pmatrix} \cos\varphi_j \\ \sin\varphi_j \end{pmatrix}$, $\varphi_j = j\Delta\varphi$, provided that $\Delta\varphi \leq \pi/\Omega\rho$.

The formulas (16), (17) describe the filtered backprojection algorithm – except for an interpolation step in the second argument. It is generally acknowledged that linear interpolation suffices.

3.2 Iterative Algorithms

The equations one has to solve in CT are usually of the form

$$R_j f = g_j \ , \qquad j = 1, \dots, p \tag{17}$$

where R_j are certain linear operators from a Hilbert space into an Hilbert space H_j. For instance, for the 2D problem we have

$$(R_j f)(s) = (Rf)(\theta_j, s) \; . \tag{18}$$

Kaczmarz's method solves linear systems of the kind (18) by successively projecting orthogonally onto the affine subspaces $R_j f = g_j$ of H:

$$f_j = P_{j \bmod p} f_{j-1} \; , \qquad j = 1, 2, \ldots \; .$$

Here, P_j is the orthogonal projection onto $R_j f = g_j$, i.e.

$$P_j f = f - R_j^* (R_j R_j^*)^{-1} (R_j f - g_j) \; .$$

For the application to tomography we replace the operator $R_j R_j^*$ by a simpler one, C_j, and introduce a relaxation parameter ω. Putting

$$\begin{aligned} P_j f \; &= f - R_j^* C_j^{-1} (R_j f - g_j) \; , \\ P_j^\omega f &= (1-\omega) f + \omega P_j f \end{aligned}$$

we have the following convergence result.

Theorem 3.5. *Let (17) be consistent. Assume that $C_j \leq R_j R_j^*$ (i.e. C_j hermitian and $C_j - R_j R_j^*$ positive semidefinite) and $0 < \omega < 2$. Then, $f_j = P^\omega_{j \bmod p} f_{j-1}$ converges for $f_0 = 0$ to the solution of (18) with minimal norm.*

The condition $0 < \omega < 2$ is reminiscent of the SOR (successive overrelaxation) method of numerical analysis. In fact Kaczmarz's method can be viewed as an SOR method applied to the linear system $R^* R f = R^* g$ where $R = R_1 \times \cdots \times R_p$.

Theorem 3.5 can be applied directly to tomographic problems. The convergence behaviour depends critically on the choice of ω and of the ordering of (17). We analyse this for the standard case (18). For convenience we assume that f is supported in $|x| < 1$ and $g_j = 0$, $j = 1, \ldots, p$.

The following result is due to Hamaker and Solmon (1978):

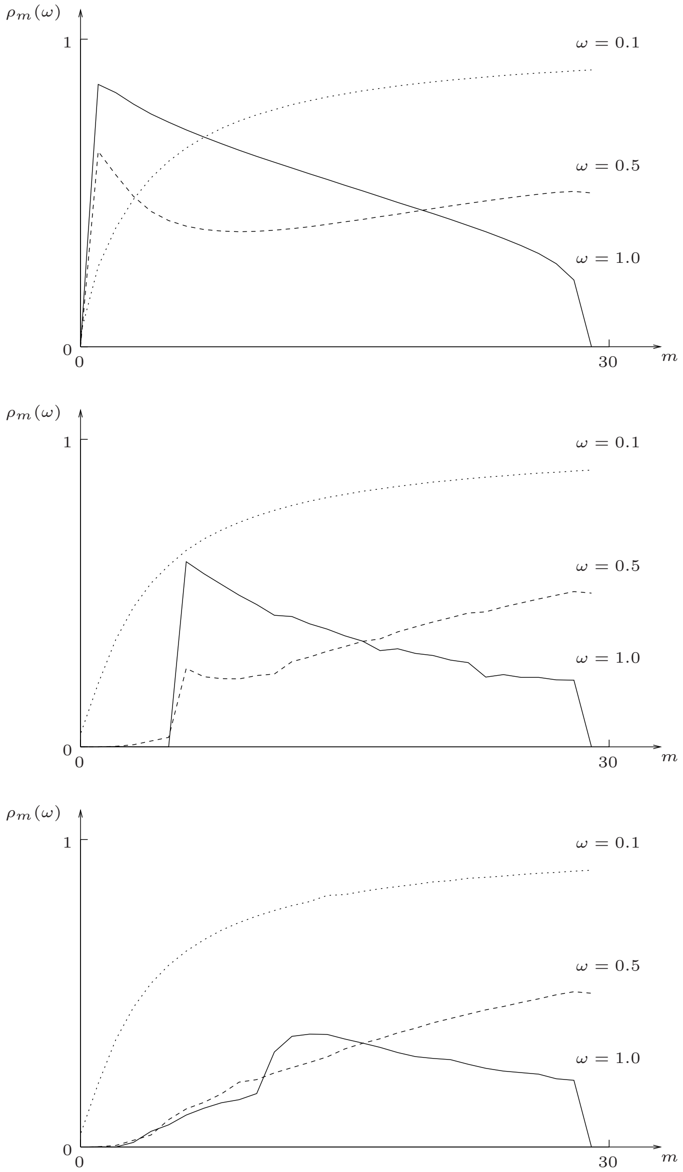
Theorem 3.6. *Let $\mathcal{C}_m$ be the linear subspace of $L_2(|x| < 1)$ spanned by $T_m(x \cdot \theta_1), \ldots, T_m(x \cdot \theta_p)$, T_m the first kind Chebysheff polynomials. Then,*

$$P_j \mathcal{C}_m \subseteq \mathcal{C}_m \; .$$

This means that $\mathcal{C}_m$ is an invariant subspace under the iteration of Theorem 3.6. Thus we can study the convergence behaviour on each of these finite dimensional subspaces separately. Let $\rho_m(\omega)$ be the spectral radius of

$$P^\omega = P_1^\omega \cdots P_p^\omega \; .$$

$\rho_m(\omega)$ can be computed numerically:



First we consider the solid lines. They correspond to $\omega = 1$. For the sequential ordering $\varphi_j = j\pi/p$, $j = 1, \ldots, p$ the values of $\rho_m(\omega)$ are displayed in the top figure. We see that $\rho_m(\omega)$ is fairly large for low order subspaces $\mathcal{C}_m$, indicating slow convergence for low frequency parts of f, i.e. for the overall features of f. The other figures shows the corresponding values of $\rho_m(\omega)$ for a non-consecutive ordering of the φ_j suggested by Herman and Meyer (1993) and for a random ordering respectively. Both orderings yield much smaller values of $\rho_m(\omega)$, in particular on the low order subspaces. The conclusion is that for practical work one should never use the consecutive ordering, and random ordering is usually as good as any other more sophisticated ordering.

In particular in emission tomography the use of multiplicative iterative methods is quite common. The most popular of these multiplicative algorithms is the expectation-maximization (EM) algorithm: Let $Af = g$ be a linear system, with all the elements of the (not necessarily square) matrix A being non-negative. Assume that A is normalized such that $A1 = 1$ where 1 stands for vectors of suitable dimension containing only $1'$s. The EM algorithm for the solution of $Af = g$ reads

$$f^{j+1} = f^j A^* \frac{g}{Af^j} \, , \quad j = 1, 2, \dots \, ,$$

where division and multiplication are understood componentwise. One can show that f^j converges to the minimizer of the log likelihood function

$$\ell(f) = \sum_i \left(g_i \log(Af)_i - (Af)_i \right) \, .$$

There exists also a version of the EM algorithm that decomposes the system $Af = g$ into smaller ones, $A_j f = g_j$, $j = 1, \dots, p$, as in the Kaczmarz method (ordered subset EM, OSEM; see Hudson and Larkin (1992)). As in the Kaczmarz method one can obtain good convergence properties by a good choice of the ordering, and the convergence analysis is very much the same.

For more on iterative methods in CT see Censor (1981).

3.3 Fourier Methods

Fourier methods make direct use of relations such as Theorem 2.1. Let's consider the 2D case in which we have

$$\hat{f}(\sigma\theta) = (2\pi)^{-1/2} \hat{g}(\theta, \sigma) \, , \quad g = Rf \, . \tag{19}$$

In order to implement this we first have to do a 1D discrete Fourier transform on g for each direction θ_j:

$$\hat{g}(\theta_j, \sigma_\ell) = (2\pi)^{-1/2} \Delta s \sum_k e^{-i\sigma_\ell s_k} g(\theta_j, s_k)$$

where $s_k = k\Delta s$ and $\sigma_\ell = \ell\Delta\sigma$. This provides $\hat{f}$ at the points $\sigma_\ell \theta_j$ which form a grid in polar coordinates. The problem is now to do a 2D inverse Fourier transform on this polar coordinate grid. Various methods have been developed for doing this, in particular the gridding method and various fast Fourier transform on non-equispaced grids; see Fourmont (1999), Beylkin (1995), Steidl (1998).

4 Reconstruction from Cone Beam Data

One of the most urgent needs of present day's X-ray CT is the development of accurate and efficient algorithms for the reconstruction of f from $(Cf)(a,\theta)$ where a is on a source curve A (typically a helix) and $\theta \in S^2$. In practice, θ is restricted to a subset of S_0, but we ignore this additional difficulty.

We assume that the source curve A satisfies Tuy's condition, i.e. A intersects each plane hitting supp (f) transversally. Let A be the curve $x = a(\lambda)$, $\lambda \in \Lambda$. Tuy's condition implies that for each s with $(Rf)(\theta, s) \neq 0$ there is λ such that $s = a(\lambda) \cdot \theta$. In fact there may be more than one such λ, but we assume that their number is finite. Then we may choose a function $M(\theta, s)$ such that

$$\sum_{\substack{\lambda \\ s=a(\lambda)\cdot\theta}} M(\theta,\lambda) = 1$$

for each s. Let $g(\lambda,\theta) = (Cf)(a(\lambda),\theta)$ be the data function. Put

$$G(\lambda,\theta) = \int_{\theta^\perp \cap S^2} \frac{\partial}{\partial\theta} g(\lambda,\omega)d\omega \ .$$

Then we have

Theorem 4.1. *Let w be a function in $\mathbb{R}^1$ and $V = R^* w'$. Then,*

$$(V * f)(x) = \int_{S^2}\int_{\Lambda} w((x-a(\lambda))\cdot\theta)G(\lambda,\theta)|a'(\lambda)\cdot\theta|M(\theta,\lambda)d\lambda d\theta \ .$$

For the proof we start out from the 3D case of Theorem 3.1:

$$\begin{aligned}(V * f)(x) &= \int_{S^2}\int_{\mathbb{R}^1} w'(x\cdot\theta - s)(Rf)(\theta,s)dsd\theta \\ &= \int_{S^2}\int_{\mathbb{R}^1} w(x\cdot\theta - s)(Rf)'(\theta,s)dsd\theta \ .\end{aligned}$$

In the s integral we make the substitution $s = a(\lambda)\cdot\theta$, obtaining

$$(V * f)(x) = \int_{S^2}\int_{\Lambda} w((x-a)\cdot\theta)(Rf)'(\theta,a(\lambda)\cdot\theta)|a'(\lambda)\cdot\theta|M(\theta,\lambda)d\lambda d\theta \ ,$$

the factor $M(\theta, s)$ being due to the fact that $\lambda \to a(\lambda)\cdot\theta$ is not one-to-one. By Theorem 2.8,

$$(Rf)'(\theta, a(\lambda)\cdot\theta) = G(\lambda,\theta) \ ,$$

and this finishes the proof.

Theorem 4.1 is the starting point for a filtered backprojection algorithm. For

$$w' = -\frac{1}{8\pi^2}\delta' ,$$

δ the 1D δ-function, we have $V = \delta$, the 3D δ-function, hence

$$f(x) = \int\limits_{\Lambda} |x-a(\lambda)|^{-2} G^w\left(\lambda, \frac{x-a(\lambda)}{|x-a(\lambda)|}\right) d\lambda ,$$

$$G^w(\lambda,\omega) = -\frac{1}{8\pi^2}\int\limits_{S^2} \delta'(\omega\cdot\theta)G(\lambda,\theta)|a'(\lambda)\cdot\theta|M(\theta,\lambda)d\theta,\ \omega\in S^2 .$$

This is the formula of Clack and Defrise (1994) and Kudo and Saito (1994). For more recent work see Katsevich (2002).

There exist various approximate solutions to the cone-beam reconstruction problem, most prominently the FDK formula for a circular source curve (which doesn't suffice Tuy's condition). It can be viewed as a clever adaption of the 2D filtered backprojection algorithm to 3D; see Feldkamp, Davis and Kress (1984). An approximate algorithm based on Orlov's condition of Theorem 2.7 is the π method of Danielsson et al. (1999).

References

1. M. Abramowitz and I.A. Stegun (1970): *Handbook of Mathematical Functions.* Dover.
2. G. Beylkin (1995): 'On the fast Fourier transform of functions with singularities', *Appl. Comp. Harm. Anal.* **2**, 363-381.
3. Y. Censor (1981): 'Row–action methods for huge and sparce systems and their applications', *SIAM Review* **23**, 444-466.
4. R. Clack and M. Defrise (1994): 'A Cone–Beam Reconstruction Algorithm Using Shift–Variant Filtering and Cone–Beam Backprojection', *IEEE Trans. Med. Imag.* **13**, 186-195.
5. P.E. Danielsson, P. Edholm, J. Eriksson, M. Magnusson Seger, and H. Turbell (1999): The original π-method for helical cone-beam CT, Proc. Int. Meeting on Fully 3D-reconstruction, Egmond aan Zee, June 3-6.
6. L.A. Feldkamp, L.C. Davis, and J.W. Kress (1984): 'Practical cone–beam algorithm', *J. Opt. Soc. Amer. A* **6**, 612-619.
7. K. Fourmont (1999): 'Schnelle Fourier–Transformation bei nicht–äquidistanten Gittern und tomographische Anwendungen', Dissertation *Fachbereich Mathematik und Informatik der Universität Münster, Münster, Germany.*
8. I.M. Gel'fand, M.I. Graev, N.Y. Vilenkin (1965): *Generalized Functions, Vol. 5: Integral Geometry and Representation Theory.* Academic Press.
9. P. Grangeat (1991): 'Mathematical framework of cone–beam reconstruction via the first derivative of the Radon transform' in: *G.T. Herman, A.K. Louis and F. Natterer (eds.): Lecture Notes in Mathematics* **1497**, 66-97.
10. C. Hamaker, K.T. Smith, D.C. Solmon and Wagner, S.L. (1980): 'The divergent beam X-ray transform', *Rocky Mountain J. Math.* **10**, 253-283.

11. C. Hamaker and D.C. Solmon (1978): 'The angles between the null spaces of X-rays', *J. Math. Anal. Appl.* **62**, 1-23.
12. S. Helgason (1999): *The Radon Transform.* Second Edition. Birkhäuser, Boston.
13. G.T. Herman (1980): *Image Reconstruction from Projection. The Fundamentals of Computerized Tomography.* Academic Press.
14. G.T. Herman and L. Meyer (1993): 'Algebraic reconstruction techniques can be made computationally efficient', *IEEE Trans. Med. Imag.* **12**, 600-609.
15. H.M. Hudson and R.S. Larkin (1994): 'Accelerated EM reconstruction using ordered subsets of projection data', *IEEE Trans. Med. Imag.* **13**, 601-609.
16. A.C. Kak and M. Slaney (1987): *Principles of Computerized Tomography Imaging.* IEEE Press, New York.
17. A. Katsevich (2002): 'Theoretically exact filtered backprojection-type inversion algorithm for spiral CT', *SIAM J. Appl. Math.* **62**, 2012-2026.
18. H. Kudo and T. Saito (1994): 'Derivation and implementation of a cone–beam reconstruction algorithm for nonplanar orbits', *IEEE Trans. Med. Imag.* **13**, 196-211.
19. F. Natterer (1986): *The Mathematics of Computerized Tomography.* John Wiley & Sons and B.G. Teubner. Reprint SIAM 2001.
20. F. Natterer and F. Wübbeling (2001): *Mathematical Methods in Image Reconstruction.* SIAM, Philadelphia.
21. F. Natterer and E.L. Ritman (2002): 'Past and Future Directions in X-Ray Computed Tomography (CT)', to appear in *Int. J. Imaging Systems & Technology.*
22. R.G. Novikov (2000): 'An inversion formula for the attenuated X-ray transform', *Preprint, Departement de Mathématique, Université de Nantes.*
23. V.A. Sharafutdinov (1994): *Integral geometry of Tensor Fields.* VSP, Utrecht.
24. G. Steidl (1998): 'A note on fast Fourier transforms for nonequispaced grids', *Advances in Computational Mathematics* **9**, 337-352.
25. S. Webb (1990): *From the Watching of Shadows.* Adam Hilger.