On the Direct Fourier Method for Computer Tomography

David Gottlieb
Brown University
Division of Applied Mathematics
and
Bertil Gustafsson
Uppsala University
Department of Scientific Computing

March 28, 1998

Abstract

The paper considers the Direct Fourier Methods for reconstructing an image from its given X-ray projections. The main purpose here is to use concepts from numerical analysis to estimate the errors in the methods. We also suggest an alternative way to the interpolations involved in the Direct Fourier Methods and estimate the number of terms involved. Realizing that one of the main reasons for the degradation of the quality of the reconstruction is the Gibbs phenomenon, we suggest a way to overcome this difficulty and give numerical evidence to its efficacy.

1 Introduction

In this paper we consider the Direct Fourier Method (DFM) for the two-dimensional computer tomography problem of reconstructing an image from its given X-ray projections p. We limit ourselves to the case with parallel beams. Modern machines use so called fan-beam projections, where many non-parallel beams are emitted from one point in the form of a fan. However, the data obtained that way are interpolated to obtain intermediate values corresponding to the parallel case, and our analysis applies.

The projections are mathematically given by the Radon transform of the image function f. Therefore, the theoretical solution to the problem is given by the inverse Radon transform operating on p. In practice, computational methods must be used, requiring discretization of the Radon transform. Many such methods have been developed, and effective commercial tomography machines have been in use for many years. Still, there is a need for faster methods with higher image quality. In [4] a presentation is given of the most common computational methods, and the inherent difficulties associated with the problem. For a more comprehensive presentation of the problem and solution methods, see [2] and [5].

The main purpose of this paper is to introduce ideas and concepts from numerical analysis in order to get a better understanding of the computational procedures involved in the inversion of the Radon transform. In particular we will develop a new version of the Direct Fourier Method (DFM), and provide error estimates for it.

The basic idea with the DFM is to perform a Fourier transform of the projections p, and relate it to the Fourier transform of the image function f. The computational implementation involves taking the discrete Fourier transform of p and the inverse discrete Fourier transform of f using the Fast Fourier Transform (FFT). Consequently, for an image with $N \times N$ pixels, the DFM has the potential of being very fast, with the number of arithmetic operations proportional to $N^2 \log N$. However, there are two main difficulties with the method.

The first one is that some form of interpolation must be used in Fourier space, and this is a difficult procedure. We will show in this paper that any standard polynomial interpolation is bound to fail. Instead we develop another method by replacing the interpolation by the Fourier Series Method (FSM), and using the so called *Sampling Theorem*.

The second difficulty has to do with the fact that typical images in applications have sharp contours giving rise to discontinuities in the image function f. It is well known that if such a function is represented by a finite Fourier series, then the Gibbs phenomenon occurs and produces nonphysical oscillations in the image. Many ways of treating this difficulty have been tried, but still, the DFM is not giving sufficiently good quality images compared to other methods. In this paper we shall further consider this important issue, and we suggest two different ways to eliminate the major part of the oscillations. The first one is simply to apply an exponential filter on the Fourier coefficients. The second way is based on the fact that the magnitude of the oscillations is proportional to the strength of the jump in f. For typical applications there are jumps of different magnitudes, the human head is a typical example. The skull has a much higher density than the brain inside, giving rise to large jumps in f compared to the variations inside. Therefore, the part of the projections corresponding to the skull should be subtracted prior to the computation of the image. We present some preliminary tests along these lines, that demonstrate the high potential for improving image quality substantially.

The error analysis of the new FSM shows that the complete computation can be carried out with $\mathcal{O}(N^2 \log N + N^{2.4})$ arithmetic operations. Considering the image quality, we believe that this is an effective method.

In Section 2 we present basic concepts and give some examples of the Radon transform. We review in Section 3 the Direct Fourier method, and in Section 4 we discuss the most common interpolations in the Direct Fourier method and show that polynomial interpolations in the Fourier space cause large errors in the image. In Section 5 we present the FSM method and in Section 6 we derive some error estimates. Section 7 is devoted to the Gibbs phenomenon and its removal, and in Section 8 we present numerical results showing the high quality images reconstructed by the new method.

2 The Radon Transform

Consider an object with density (or image) function $f(x, y)$, and parallel X-ray beams with their normals at an angle θ with the x-axis as shown in Fig 1.

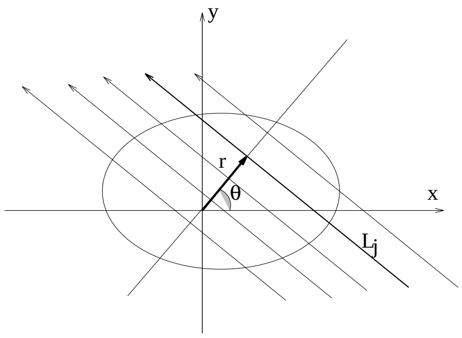

Figure 1: Parallel x-ray beams and the polar coordinate system r, θ

The beam along L_j is uniquely determined by the coordinates r, θ. The corresponding measured value is the integral of $f(x, y)$ along the beam path L_j. By introducing the Dirac delta-function $\delta(x)$, we have

$$p(r, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\delta(x \cos\theta + y \sin\theta - r) dx dy \,, \tag{2.1}$$

where the right hand side is the Radon transform $[\mathcal{R}f](r, \theta)$ of f. We can also express the integral in another way:

$$p(r, \theta) = \int_{-\infty}^{\infty} f(r\cos\theta - s\sin\theta, r\sin\theta + s\cos\theta) ds \,.$$

The reconstructed image function $f(x, y)$ is formally obtained as

$$f(x, y) = [\mathcal{R}^{-1}p](x, y) \,,$$

and there is actually an analytical formula

$$f(x, y) = \frac{1}{2\pi^2} \int_0^{\pi} \int_{-\infty}^{\infty} \frac{\frac{\partial p}{\partial r}}{x\cos\theta + y\sin\theta - r} dr d\theta \,.$$

Let us consider several examples.

Example 1. *A circle with constant unit density* .

$$f(x, y) = \begin{cases} 1 \,, & x^2 + y^2 \le c^2 \\ 0 & \text{else} \,. \end{cases} \tag{2.2}$$

The projections $p(r, \theta)$ are for this case independent of θ, and is simply given by the length of the ray-trace inside the disc:

$$p(r) = \begin{cases} 2\sqrt{c^2 - r^2}\,, & |r| \le c \\ & \text{else}\,. \end{cases} \tag{2.3}$$

This example delineates an essential difficulty in inverting the Radon transform: the image has sharp edges, and therefore the density function $f(x, y)$ has a jump discontinuity. The Radon transform $p(r)$ is continuous at the boundary but the derivative is not. *The Radon transform, in general, is smoother than its image function.*

□

Example 2. *An ellipse with constant density ρ.*
The boundary is defined by

$$(\frac{x - x_1}{a})^2 + (\frac{y - y_1}{b})^2 = 1\,,$$

i.e., it is centered at (x_1, y_1) and has its axes of length $2a$ and $2b$ aligned with the $x-$ and $y-$ axes respectively. A direct evaluation of (2.1) gives

$$p(r, \theta) = g(\rho, a, b, x_1, y_1, r, \theta)$$

where

$$g(\rho, a, b, x_1, y_1, r, \theta) = \frac{2\rho ab}{a^2\cos^2\theta + b^2\sin^2\theta} \times$$
$$\sqrt{(a^2\cos^2\theta + (b^2 - y_1^2)\sin^2\theta - (r - x_1\cos\theta)(r - x_1\cos\theta - 2y_1\sin\theta))^+}\,. \tag{2.4}$$

Here

$$(\xi)^+ = \begin{cases} \xi & \text{if } \xi > 0 \\ 0 & \text{otherwise}\,. \end{cases}$$

Fig 2 is a contour plot of $p(r, \theta)$, and Fig 3 is a one-dimensional cut at $\theta = \pi/2$. The parameter values are

$$\rho = 1,\ a = 1/2,\ b = 1,\ x_1 = 0,\ y_1 = 0\,.$$

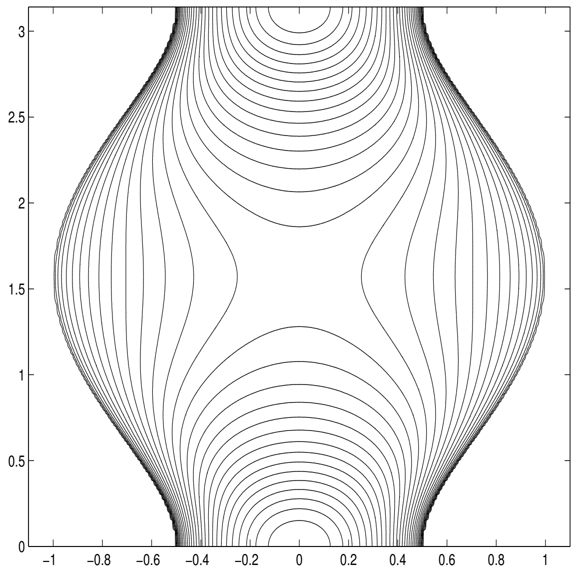

Figure 2: Contour plot of $p(r,\theta)$

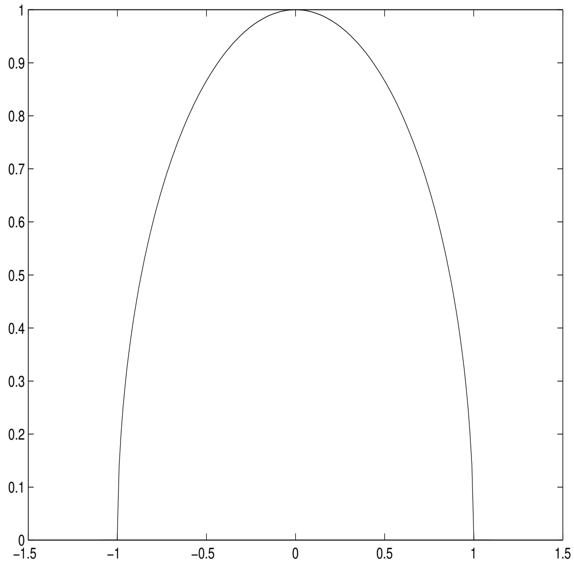

Figure 3: $p(r,\pi/2)$

As in the first example the function p is continuous, but has discontinuous and infinite derivatives at the boundary of the ellipse.

$\square$

Example 3. *One ellipse E_2 with constant density ρ_2 added to another one E_1 with constant density ρ_1.* (The density at the inner part is $\rho_1 + \rho_2$.)

$$E_1 : \text{ Boundary } \left(\frac{x - x_1}{a_1}\right)^2 + \left(\frac{y - y_1}{b_1}\right)^2 = 1 \quad , \text{ density } \rho_1 \, ,$$

$$E_2 : \text{ Boundary } \left(\frac{x - x_2}{a_2}\right)^2 + \left(\frac{y - y_2}{b_2}\right)^2 = 1 \quad , \text{ density } \rho_2 \, .$$

Direct integration gives

$$p(r, \theta) = g(\rho_1, a_1, b_1, x_1, y_1, r, \theta) + g(\rho_2, a_2, b_2, x_2, y_2, r, \theta) \, , \qquad (2.5)$$

where g is defined in (2.4). Fig 4 and 5 show the contour plot of $p(r, \theta)$ and the one-dimensional cut at $\theta = \pi/2$ as in Example 1. The parameter values are

$\rho_1 = 1$	$a_1 = 1/2$	$b_1 = 1$	$x_1 = 0$	$y_1 = 0$
$\rho_2 = -0.9$	$a_2 = 0.45$	$b_2 = 0.9$	$x_2 = 0$	$y_2 = 0.05$

$\square$

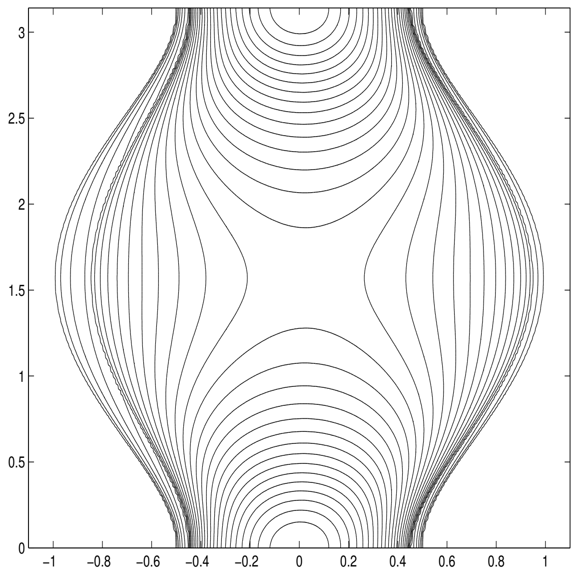

Figure 4: Contour plot of $p(r, \theta)$ for double ellipse

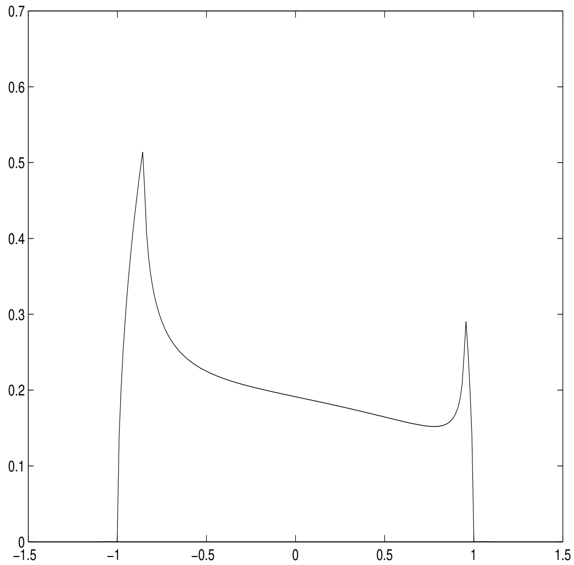

Figure 5: $p(r, \pi/2)$ for double ellipse

In the reconstruction problem, the Radon Transform $p(r, \theta)$ is assumed given, and we need to compute $f = \mathcal{R}^{-1}p$. Since p is continuous and f is discontinuous, the inverse Radon transform apparently introduces a deterioration of smoothness. This type of ill-conditioning is something that we have to deal with when constructing the computational method.

3 The Direct Fourier Method (DFM)

The DFM is based on the *Slice Theorem* that relates the two dimensional Fourier transform of the image density function $f(x, y)$ to the r-Fourier transform of the Radon transform $p(r, \theta)$.

We assume that the image function $f(x, y)$ is compactly supported in the square $[-c, c] \times [-c, c]$, and consequently $p(r, \theta)$ has compact support in $[-d, d] \times [0, \pi]$, where $d \geq c\sqrt{2}$. Thus $p(r, \theta)$ can be represented by

$$p(r, \theta) = \int_{-\infty}^{\infty} \hat{p}(\rho, \theta) e^{2\pi i \rho r} d\rho \,, \tag{3.6}$$

where $\hat{p}(\rho, \theta)$ denotes the one-dimensional Fourier transform of p in the r-

direction:

$$\hat{p}(\rho,\theta) = \int_{-d}^{d} p(r,\theta) e^{-2\pi i \rho r} dr \,. \tag{3.7}$$

In the same manner the density function $f(x,y)$ can be represented by

$$f(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{f}(\xi,\eta) e^{2\pi i(\xi x + \eta y)} d\xi d\eta \,, \tag{3.8}$$

where $\hat{f}(\xi,\eta)$ denotes the two-dimensional Fourier transform of f:

$$\hat{f}(\xi,\eta) = \int_{-c}^{c} \int_{-c}^{c} f(x,y) e^{-2\pi i(\xi x + \eta y)} dx dy \,. \tag{3.9}$$

A simple relation between $\hat{f}(\xi,\eta)$ and $\hat{p}(\rho,\theta)$ is provided by

Theorem 3.1 (**The Slice Theorem**)
Let $\hat{p}(r,\theta)$ be the r-Fourier transform (3.7) *of the Radon transform $p(r,\theta)$, and let $\hat{f}(\xi,\eta)$ be the two-dimensional Fourier transform* (3.9) *of the image function $f(x,y)$. Then*

$$\begin{aligned}
\hat{p}(\rho,\theta) &= \hat{f}(\rho\cos\theta, \rho\sin\theta) \,, \quad -\infty < \rho < \infty \,, \;\; 0 \le \theta < \pi \,, && (3.10)\\
\hat{f}(\xi,\eta) &= \hat{p}(\sqrt{\xi^2+\eta^2}, \tan^{-1}(\frac{\eta}{\xi})) \,, \quad -\infty < \xi, \eta < \infty \,. && (3.11)
\end{aligned}$$

The proof is a direct consequence of the integral relation between f and p.

□

The *Slice Theorem* suggests a potentially attractive method to recover the image function $f(x,y)$ from its Radon transform $p(r,\theta)$:

(a) Get $\hat{p}(\rho,\theta)$ from $p(r,\theta)$ by (3.7).

(b) Solve for $\hat{f}(\xi,\eta)$ by using the *Slice Theorem* (3.11).

(c) Find the image function $f(x,y)$ from its Fourier transform $\hat{f}(\xi,\eta)$ by using (3.8).

The above three steps cannot be carried out numerically as stated, we need to discretize (3.7) and (3.8). To do that, we must use appropriate quadrature formulas.

We start with step (a). In order to approximate the integral in (3.7) we use the trapezoidal formula based on M points :

$$\hat{p}(\rho,\theta) \quad \approx \quad \frac{2d}{M}\sum_{j=-M/2}^{M/2-1} p(j\Delta r,\theta)e^{-2\pi i\rho 2dj/M}\,, \tag{3.12}$$

where $\Delta r = 2d/M$.

In order to evaluate (3.12) efficiently , we use the points $\rho_m = m/2d$, to get

$$\tilde{p}_m(\theta) = \frac{2d}{M}\sum_{j=-M/2}^{M/2-1} p(j\Delta r,\theta)e^{-2\pi ijm/M}\,,\quad -M/2 \le m \le M/2\,, \tag{3.13}$$

which can be evaluated, using the Fast Fourier Transform (FFT), by $\mathcal{O}(MlogM)$ arithmetic operations. *Note that $\tilde{p}_m(\theta)$ approximates $\hat{p}(m/2d,\theta)$.* Step (a) is now replaced by (3.13).

We turn next to step (c). Since we would like to have the density function at the pixels $x_k = 2ck/N$, $y_l = 2cl/N$, we are required to approximate

$$f(x_k,y_l) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{f}(\xi,\eta)e^{4\pi ic(\xi k+\eta l)/N}d\xi d\eta\,,\ -N/2 \le k,l \le N/2\,. \tag{3.14}$$

In order to be able to use the FFT we need to discretize the integral at the points $\xi_\nu = \nu/2c$, $\eta_\mu = \mu/2c$ to get

$$f(x_k,y_l) \approx \sum_{\nu=-N/2}^{N/2-1}\sum_{\mu=-N/2}^{N/2-1}\hat{f}(\frac{\nu}{2c},\frac{\mu}{2c})e^{2\pi i(k\nu+l\mu)/N}\,,\ -N/2 \le k,l \le N/2\,. \tag{3.15}$$

We now face the main difficulty of the Direct Fourier Method:
In order to use the FFT we need to know the values of the Fourier transform of the image function at the points $\xi = \nu/2c$, $\eta = \mu/2c$. By the *Slice Theorem*, these values are given by $\hat{p}(\sqrt{\mu^2+\nu^2}/2c,\theta)$. However, numerically

we can compute only $\tilde{p}_m(\theta)$ (see (3.13)) that approximates $\hat{p}(m/2d, \theta)$! Thus we do not have $\hat{p}$ at the right points.

Some sort of interpolation is needed to overcome this difficulty. The common method is to use standard polynomial interpolation. In the next section we will discuss the drawbacks of this type of interpolation when done in the Fourier space. We will then in Section 5 introduce a better method to reconstruct the Fourier transform of the image function from the Fourier transform of the Radon transform function by using the *Sampling Theorem*.

4 Interpolation in Fourier space

As was described above, the Fourier transform of the line integrals is given on a grid corresponding to a polar coordinate system, but we need to represent it on a Cartesian grid. We shall first show that traditional, standard interpolation based on polynomial fits through neighboring points can cause erroneous results when applied in the Fourier space. Specifically we will show the effect, in the physical space, of an interpolation in the Fourier space.

We investigate first a one-dimensional model problem where the function $p(r)$ has compact support in $[-d, d]$, and is given by the Fourier integral

$$p(r) = \int_{-\infty}^{\infty} \hat{p}(\rho) e^{2\pi i r \rho} d\rho \, .$$

As an example, we consider the case where $\hat{p}(\rho)$ must be found at the center in between the given points. With linear interpolation and the step-size $\Delta\rho$, we substitute $\hat{p}(\rho)$ by

$$\hat{q}(\rho) = \frac{1}{2}(\hat{p}(\rho + \frac{\Delta\rho}{2}) + \hat{p}(\rho - \frac{\Delta\rho}{2})) \ .$$

The new function $q(r)$ in physical space is

$$q(r) = \int_{-\infty}^{\infty} \frac{1}{2}(\hat{p}(\rho + \frac{\Delta\rho}{2}) + \hat{p}(\rho - \frac{\Delta\rho}{2}))e^{2\pi i \rho r} d\rho$$

$$= \int_{-\infty}^{\infty} \frac{1}{2}(\hat{p}(\rho) e^{2\pi i (\rho - \frac{\Delta\rho}{2})r} + \hat{p}(\rho) e^{2\pi i (\rho + \frac{\Delta\rho}{2})r}) d\rho = \cos(\pi r \Delta\rho) p(r) \ .$$

The error is $(1 - \cos(\pi r \Delta\rho))p(r)$. If $r\Delta\rho$ is small, then the error is $\approx \pi^2 r^2 (\Delta\rho)^2 p(r)/2$, which is small.

In the actual computation we use the FFT. We recall (see the discussion after (3.13)), that if $p(r)$ is given on a grid with M points, the transform gives the values

$$\tilde{p}_m \approx \hat{p}(\frac{m}{2d}) \ , \quad m = -M/2, \ldots, M/2 - 1 \ .$$

Therefore

$$\Delta\rho = \frac{1}{2d} \ ,$$

and we have the error

$$(1 - \cos(\frac{\pi r}{2d}))p(r) \ .$$

Basically we have approximated the function 1 by $\cos(\pi r/2d)$, however, this approximation is good only if $\pi r/2d$ is small. *Traditional interpolation in the Fourier space will produce good results in the physical space only for* $r << 2d$.

For the image reconstruction problem, this means that the interpolated Fourier transform corresponds to modified line integrals which are accurate only for those projections passing close to the center of the object. This shows that the traditional interpolation cannot be used for any reasonable image quality. Taking a larger domain (d large) will reduce the error, however, we will also lose resolution.

Higher order interpolations do not provide cure for the problem, as they will result in approximations of the function 1 by series of *cosine* and *sine* functions that will approximate it better for small arguments only. The error will still contain higher powers of $\pi r/2d$, which is not necessarily small. However, the coefficients multiplying these $\Delta\rho$ powers may become smaller, giving some improvement.

It should be emphasized that this observation about poor performance by polynomial interpolation holds independent of the image resolution. If M, the number of grid points, increases, the *number of* $\tilde{p}_m$ increases. However, $\Delta\rho$ doesn't become smaller, it is always equal to $1/2d$.

In our analysis so far, we have limited ourselves to the one-dimensional case. In the two-dimensional case, the projections are functions $p = p(r, \theta)$, where θ is the angle. The Fourier transform is taken over r, giving $\hat{p}(\rho, \theta)$ as in (3.7). This function will be represented on a grid with a large step-size $\Delta\rho = 1/2d$ in the ρ-direction, just as described above. However, if there are

J projections in the θ-direction for each r, the step-size in the θ-direction is $\Delta\theta = \pi/J$. This is a small quantity, and if p is not too rough in the θ-direction, we can expect standard polynomial interpolation to work well.

In conclusion: better interpolation methods are needed in the ρ-direction though not in the θ-direction.

5 Replacing polynomial interpolation by the Fourier Series Method (FSM)

In the FSM we abandon the idea of using interpolation in the Fourier space. The new technique is based on using the *Slice Theorem* to relate the Fourier series of $p(r,\theta)$ to the truncated Fourier series of the image function $f(x,y)$.

Let $\hat{p}_m(\theta)$ be the Fourier coefficients (in the radial direction) of the Radon transform $p(r,\theta)$, namely

$$\hat{p}_m(\theta) = \int_{-d}^{d} p(r,\theta) e^{-\pi i m r/d} dr \quad , \; m = 0, \pm 1, \pm 2, \ldots \; . \tag{5.16}$$

It is well known that $p(r,\theta)$ has the Fourier series expansion:

$$p(r,\theta) = \frac{1}{2d} \sum_{m=-\infty}^{\infty} \hat{p}_m(\theta) e^{\pi i m r/d} \, . \tag{5.17}$$

We use now the series (5.17) to get a representation of the r-Fourier transform $\hat{p}(\rho,\theta)$ of $p(r,\theta)$ in terms of the Fourier coefficients $\hat{p}_m(\theta)$:

Lemma 5.1 (**The Sampling Theorem**) *Let $\hat{p}_m(\theta)$ be the Fourier coefficients of the function $p(r,\theta)$ given in* (5.16)*, and let $\hat{p}(\rho,\theta)$ be the Fourier transform (in the r-direction) of $p(r,\theta)$ defined in* (3.7)*. Then*

$$\hat{p}(\rho,\theta) = \sum_{m=-\infty}^{\infty} \hat{p}_m(\theta) \mathrm{sinc}(\pi(m - 2d\rho)) \, . \tag{5.18}$$

Proof:

The proof follows immediately from the identity

$$\frac{1}{2d} \int_{-d}^{d} e^{\pi i m r/d} e^{-2\pi i \rho r} dr = \mathrm{sinc}(\pi(m - 2d\rho) \, . \tag{5.19}$$

$\square$

In the same way , with $\hat{f}_{kl}$ being the Fourier coefficients of $f(x,y)$, namely

$$f(x,y) = \frac{1}{4c^2} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \hat{f}_{kl} e^{\pi i(kx+ly)/c}\,, \tag{5.20}$$

we get:

$$\begin{aligned} \hat{f}(\xi,\eta) &= \frac{1}{4c^2} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \hat{f}_{kl} \int_{-c}^{c} \int_{-c}^{c} e^{\pi i(kx+ly)/c} e^{-2\pi i(\xi x+\eta y)} dx dy \\ &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \hat{f}_{kl} \mathrm{sinc}(\pi(k-2c\xi))\mathrm{sinc}(\pi(l-2c\eta))\,. \end{aligned}$$

The *Slice Theorem* now implies

$$\sum_{m=-\infty}^{\infty} \hat{p}_m(\theta)\mathrm{sinc}(\pi(m-2d\rho))$$

$$= \sum_{j=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \hat{f}_{jn} \mathrm{sinc}(\pi(j-2c\rho\cos\theta))\mathrm{sinc}(\pi(n-2c\rho\sin\theta))\,.$$

For every integer pair (k,l) we now choose θ and ρ such that

$$\begin{aligned} 2c\rho\cos\theta &= k\,, \\ 2c\rho\sin\theta &= l\,, \end{aligned} \tag{5.21}$$

which leaves only one term $\hat{f}_{kl}$ in the right hand side. Since we are working in the domain $-d \leq \rho \leq d,\ 0 \leq \theta < \pi$, we get from (5.21)

$$\begin{aligned} \theta &= \arctan(\tfrac{l}{k}) + \pi s_0(kl)\ , \\ \rho &= \pi s_1(l)\sqrt{k^2+l^2}\ , \end{aligned}$$

where

$$s_0(j) = \begin{cases} 0 & \text{if } j \geq 0 \\ 1 & \text{else} \end{cases} \quad , \quad s_1(l) = \begin{cases} 1 & \text{if } l \geq 0 \\ -1 & \text{else} \end{cases} \quad .$$

The final formula is

$$\hat{f}_{kl} = \sum_{m=-\infty}^{\infty} \hat{p}_m(\arctan(\frac{l}{k}) + \pi s_0(kl))\ \mathrm{sinc}(\pi m - \pi s_1(l)\frac{d}{c}\sqrt{k^2+l^2})\,. \tag{5.22}$$

Note that this is an exact formula for evaluating the Fourier coefficients, no approximations have been made. For the computational implementation, the idea is to use a small number of terms in the sum, this is our approximation in the ρ-direction. By the arguments above, traditional interpolation can be used in the θ-direction.

Obviously we must keep the number of terms in the sum low, since in the two-dimensional case with N^2 pixels, we need to evaluate the sum at N^2 points. It is important therefore to retain only those terms that contribute significantly to the sum.

The Fourier coefficients $\hat{p}_m(\theta)$ decay with m. As an example, consider the unit density function in a circle (Example 1, Section 2). A standard (though tedious) calculation yields the following expression for the Fourier coefficients of the Radon transform $p(r, \theta) = p(r)$

$$\hat{p}_m = \frac{2d^2 J_1(\pi m)}{m} \ ,$$

where $J_1(x)$ is the Bessel function of order 1. Since $J_1(x)$ decays like $|x|^{-1/2}$, we find that

$$\hat{p}_m \sim |m|^{-3/2} \,. \tag{5.23}$$

Furthermore, the function $\text{sinc}(\rho)$ has its main weight around $\rho = 0$, and decays like ρ^{-1} (see Fig. 6).

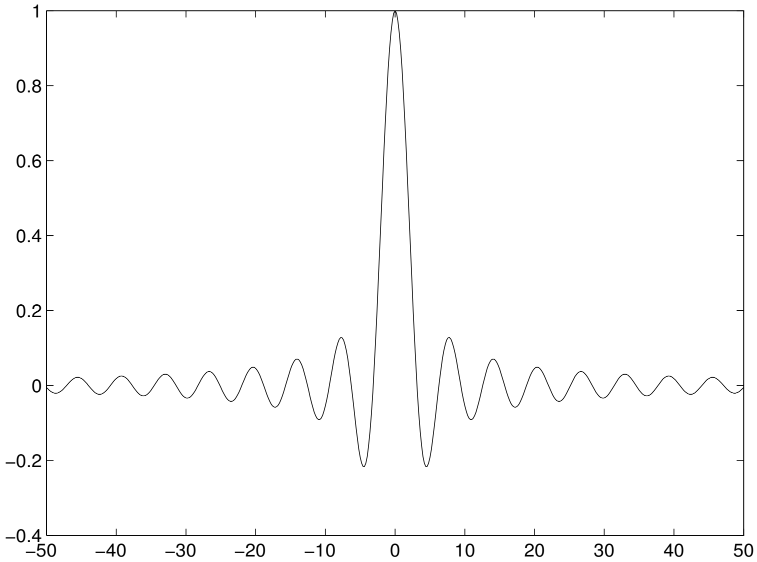

Figure 6: The function $\mathrm{sinc}(\xi)$

This example is representative for typical cases. Therefore, we define the new method to compute the Fourier coefficients of the image function f as a two-part summation. For any given (k,l), define the center m_0 for the second part of the summation as $m_0 = s_1(l)[\sqrt{k^2+l^2}+0.5]$, where $[\rho]$ denotes the integer part of ρ. The new formula is

$$\hat{f}_{kl} \approx \hat{h}_{k,l} = \sum_{m=-m_1}^{m_1} g_m(k,l) + \sum_{\substack{m\,=\,m_0-m_2 \\ |m|>m_1}}^{m_0+m_2} g_m(k,l)\,, \qquad (5.24)$$

where

$$g_m(k,l) = \hat{p}_m(\arctan(\frac{l}{k}) + \pi s_0(kl))\ \mathrm{sinc}(\pi m - \pi s_1(l)\frac{d}{c}\sqrt{k^2+l^2})$$

The choice of m_1 and m_2 will be discussed in the next section.

The image function $f(x,y)$ is finally obtained by

$$f(x_i,y_j) = \frac{1}{4c^2}\sum_{k=-N/2}^{N/2-1}\sum_{l=-N/2}^{N/2-1}\hat{h}_{kl}e^{\pi i(kx_i+ly_j)/c}\ , \quad -N/2 \le i,j \le N/2\ . \quad (5.25)$$

In actuality, we can not compute the true Fourier coefficients $\hat{p}_m$, but rather $\tilde{p}_m$ given in (3.13). The difference between $\hat{p}_m$ and $\tilde{p}_m$ is depicted in Fig 7 and in Fig 8 for the Radon transform of a unit density function in a circle. The first figure corresponds to 8 interpolating points and the second to 16 points.

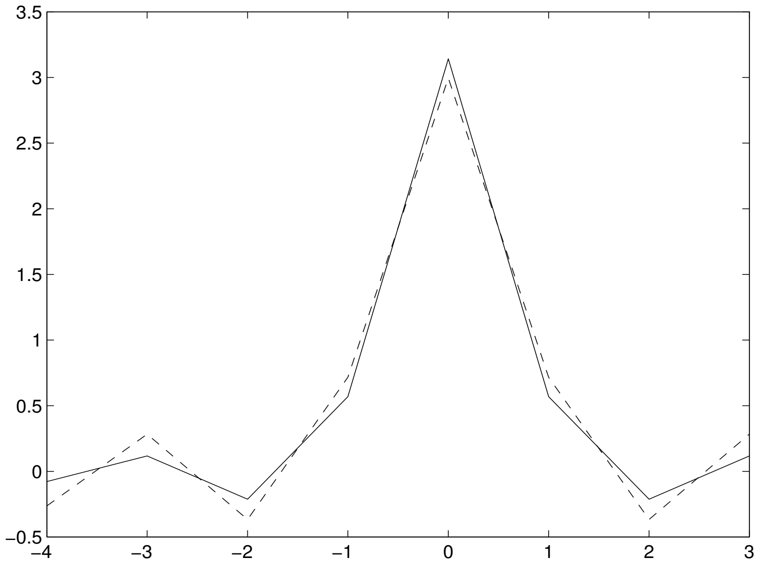

Figure 7: $\hat{p}_m$ ($-$) and $\tilde{p}_m$ ($--$) for $M = 8$

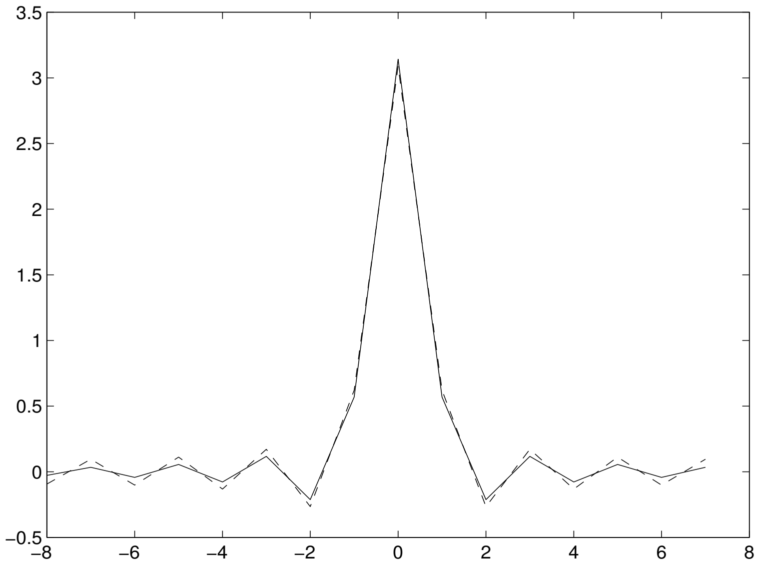

Figure 8: $\hat{p}_m$ $(-)$ and $\tilde{p}_m$ $(--)$ for $M = 16$

As indicated in the last section, polynomial interpolations in the θ direction are justified. We thus approximate $\hat{p}_m(\arctan(\frac{l}{k}) + \pi s_0(kl))$ by bilinear interpolation using its neighboring points.

A different formula can be obtained in the following way: The formula (5.18) was obtained by performing

$$\frac{1}{2d}\int_{-d}^{d} e^{\pi imr/d} e^{-2\pi i\rho r}\, dr = \mathrm{sinc}(\pi(m - 2d\rho))\,.$$

Suppose now that we replace the integration by the trapezoidal rule, i.e we use the points $r_j = 2dj/J$, $j = -J/2, ..., J/2$, then we get

$$\frac{1}{2d}\int_{-d}^{d} e^{\pi imr/d} e^{-2\pi i\rho r} \approx \frac{1}{J}\sum_{j=-J/2}^{J/2}{}'' e^{ij(2\pi m/J - 4\pi d\rho/J)}\,,$$

where the notation $\sum''$ means that the first and the last term in the sum are multiplied by $1/2$.

The trigonometric series can be summed up to get

$$\sin(\pi(m - 2d\rho))\frac{1}{J}\cot(\pi\frac{m - 2d\rho}{J}) \tag{5.26}$$

which replaces the function $sinc(\pi(m-2d\rho))$ in (5.18).

We note here that a similar formula was suggested by Magnusson ([5]), who derived it from a totally different point of view. Also, Magnusson advocated using only the second part of the summation and she took J to be the number of terms $2m_2$ retained in the summation.

Still another approach is the so called gridding method introduced in [6] and analyzed in [7]. The method uses a convolution sum with a windowed Kaiser-Bessel function in the transformed space, and it can be derived by expanding the image function $f(x,y)$ in a series of orthogonal functions related to the Gegenbauer polynomials.

6 Error Estimates

There are three sources of errors in the DFT method

- Truncating the sum as in (5.24).
- Replacing $\hat{p}_m$ by $\tilde{p}_m$.
- Truncating the sum as in (5.25).

The second source of errors is called the *aliasing* error and it is treated extensively in the literature. The aliasing error is of the same order as the truncation error, i.e the error resulting from terminating the sum in (5.17) after the first M terms. (see [8])

The third source of error is more serious. In most of the cases the image density function $f(x,y)$ is discontinuous. In this case the convergence of the Fourier series is slow and the number N in (5.25) has to be large. Moreover, the approximation by finite N introduces oscillations in the image . Methods to enhance the convergence are known in the literature and will be addressed in the next section.

In this section we will concentrate on the choice of m_1 and m_2 in the sum (5.24). On the one side it is important to keep the number of terms m_1+m_2 small, but on the other hand we do not want to reduce the accuracy.

A reasonable assumption is that the Fourier coefficients $\hat{f}_{kl}$ of an image function $f(x,y)$ decay like $\hat{f}_{kl}=\mathcal{O}(1/kl)$. This is, for example, the case for the density function which is 1 inside a sub-square of the original domain, and vanishes at the rest of the domain.

The accuracy requirement for $\hat{h}$ is determined by matching the error committed by replacing $\hat{f}$ with $\hat{h}$ in (5.24) with the error committed by the truncating the sum in (5.25). The first term neglected in (5.25) is of order $\mathcal{O}(1/N)$. We would like therefore the terms in (5.24) to be accurate of the same order.

We assume, as in Example 1, Section 2 (see (5.23)) that the Fourier coefficients $\hat{p}_m(\theta)$ of the Radon transform $p(r,\theta)$ behave as $\hat{p}_m(\theta) = \mathcal{O}(m^{-3/2})$. Note also that $|\mathrm{sinc}x| \leq 1/|x|$.

The terms included in the sum (5.24) are those with index m, such that $|m| \leq m_1$ or $|m - m_0| \leq m_2$. Thus the terms neglected have indices m such that $|m| > m_1$ and $|m - m_0| > m_2$. Taking all that into account we conclude that the first term neglected in (5.24) is of order $\mathcal{O}((m_1^{3/2}m_2)^{-1}$. The first factor is due to the decay of p_m, whereas the second factor is due to the behavior of the *sinc*-function.

For most of the coefficients $\hat{f}_{kl}$, and for realistic values of m_1, m_2, N, the two parts of the sum don't overlap each other, i.e. $|m_0| > m_1 + m_2$. Therefore $m_1 + m_2$ measures the amount of work involved, and we can formulate the following problem:

$$\text{Minimize } (m_1 + m_2)$$

under the constraint

$$m_1^{3/2} m_2 = N \, .$$

The solution of the above problem is given by

$$m_1 = (\frac{3}{2})^{\frac{2}{5}} N^{\frac{2}{5}} \, , \tag{6.27}$$

$$m_2 = (\frac{2}{3})^{\frac{3}{5}} N^{\frac{2}{5}} \, . \tag{6.28}$$

We can draw two conclusions from (6.27), (6.28):

- In order to keep the accuracy the number of terms have to be proportional to $N^{0.4}$. Thus the cost in step (b) using the FSM method is $N^{2.4}$, whereas the cost of steps (a) and (c) is $N^2 logN$. However, the additional factor of $N^{0.4}/logN$ is not very significant. For example if $N = 1024$ the factor is 1.6.

- The number of terms that should be included in the first sum in (5.24) is larger than that in the second. The proper relation between m_1 and m_2 is
$$m_1 = 1.5m_2 \tag{6.29}$$

7 Treatment of the Gibbs phenomenon

We now come to the error associated with step (c) of the DFM method, i.e. the computation (5.25). Even if the computed Fourier coefficients $\hat{h}_{k,l}$ are identical with the true coefficients $\hat{f}_{kl}$, we still face the problem of the deterioration of accuracy due to the Gibbs phenomenon, see [3].

It is well known that approximating a discontinuous function by its *finite* Fourier expansion (retaining only N terms) results in poor accuracy. The following lemma is well known:

Lemma 7.1 (**The Gibbs Phenomenon**) *Assume that we have a piecewise continuous function $F(x)$, $-\pi \le x \le \pi$, with one jump of magnitude D at a point $x = a$. ($D = |f(a-) - f(a+)|$.) Let $\hat{F}_k$ be the Fourier coefficients of $F(x)$. Then we have for every positive δ*

$$|F(x) - \sum_{k=-N}^{N} \hat{F}_k e^{ikx}| \le \mathcal{O}(\frac{D}{N}), \qquad |x-a| > \delta \tag{7.30}$$

$$\max_{-\pi \le x \le \pi} |F(x) - \sum_{k=-N}^{N} \hat{F}_k e^{ikx}| = \frac{DP_1}{2\pi}, \tag{7.31}$$

where
$$P_1 = \int_{-\pi}^{\pi} \frac{\sin x}{x} dx .$$

□

A typical image function $f(x, y)$ has a considerable jump corresponding to the outer boundary of the image. This jump and the Gibbs phenomenon it causes effects the quality of the reconstruction.

In Fig. 9 we show the two-dimensional finite Fourier expansion of the function (see Example 1, Section 2)

$$f(x,y) = \begin{cases} 1\,, & x^2+y^2 \le 1 \\ 0 & \text{else}\,, \end{cases}$$

with $N = 64$. Fig. 10 shows a side view of the same expansion. The oscillations are clearly visible.

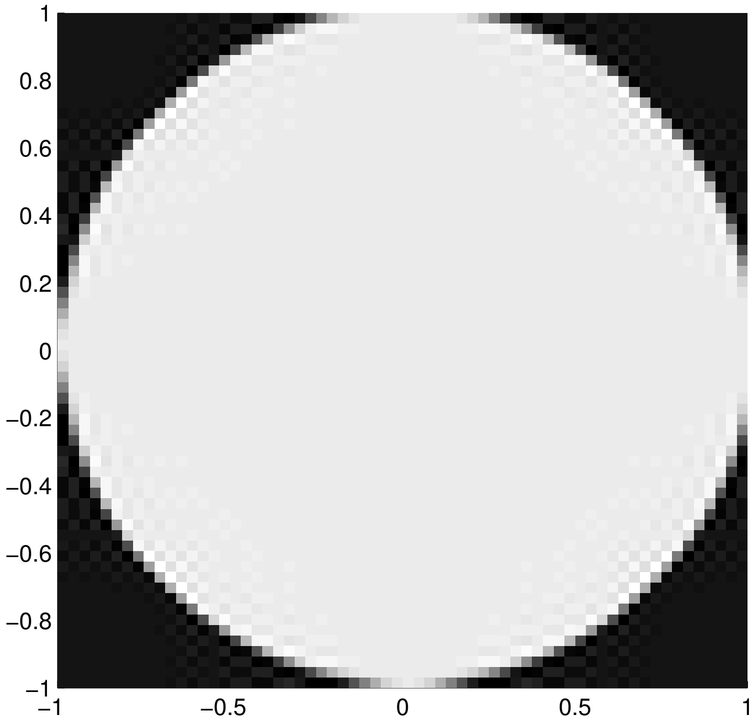

Figure 9: $f(x,y)$, $N = 64$. No filtering.

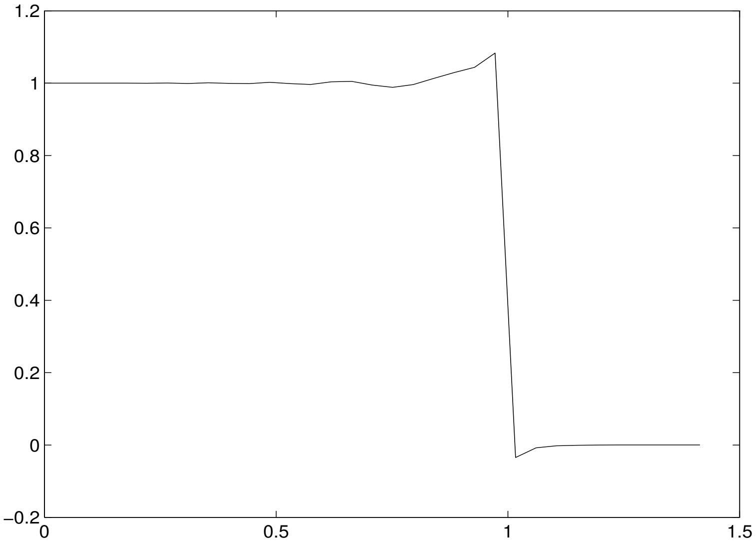

Figure 10: Side view of $f(x, y)$ along a radius in the direction 45 degrees, $N = 64$. No filtering.

A very efficient way to enhance the accuracy *away* from the discontinuities is to use smoothing procedures. We have used a filter in Fourier space, which has the form

$$\hat{h}_{kl} \rightarrow \hat{g}_{kl} = e^{-\alpha((2k/N)^{\beta}+(2l/N)^{\beta})}\hat{h}_{kl} \ .$$

In most of the cases we have used $\alpha = 5$, and $\beta = 2$. Note that this method requires minimal amount of extra computations. The filter functions can be computed and stored once and for all. To demonstrate the efficacy of this filter we repeat the same calculation as before using the filter with the parameters mentioned above. The result is shown in Fig. 11. Note that the checkerboard pattern is no longer visible, however, the boundaries are now smeared.

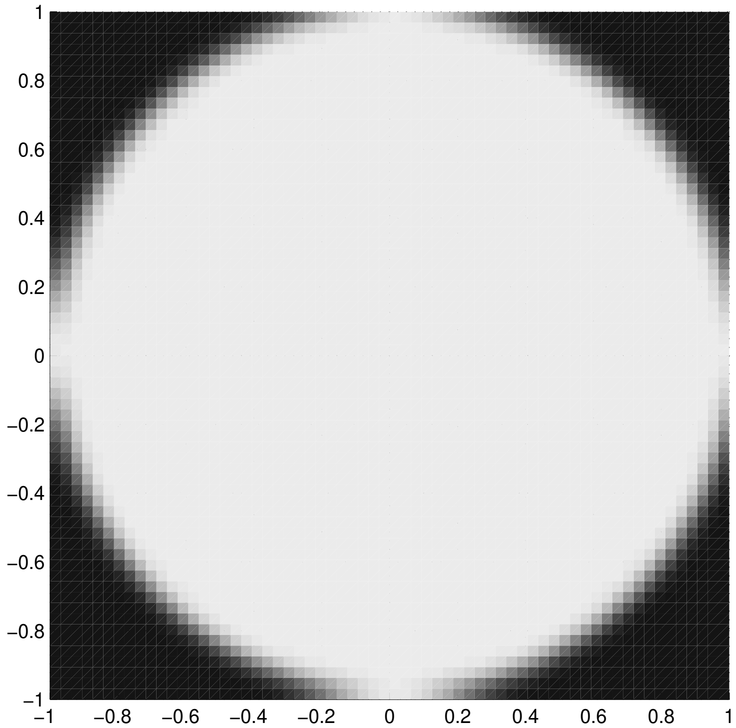

Figure 11: $f(x, y)$, $N = 64$, filter coefficients $\alpha = 5$, $\beta = 2$.

In many applications there are discontinuities of different magnitudes. For example, the skull of the human head has a much higher density than the brain inside, and it has well defined edges. Furthermore, the outer boundary of any object is an always present well defined edge. The corresponding strong discontinuities in the image function f causes the type of oscillations as described above. The magnitude of the oscillations is proportional to the strength of the discontinuity (see (7.30), (7.31)). Therefore, if these discontinuities are subtracted prior to the computation, and added on after the computation, then we should expect much better image quality. This idea has been used in a different context, see [1].

In general we cannot expect to know exactly where the discontinuities are. However, the outer edge is easy to find from the projection function p, since it is the contour where p becomes non-zero. For inner discontinuities in f, we have to look for discontinuities in the derivatives of p. If these contours correspond to a known shape with constant density, as for instance in Example 1, 2 and 3 in Section 2, an a priori evaluated function p_0 can be subtracted from p, and the reconstruction algorithm is applied to the modified function $p - p_0$. At the end we add the image function f_0 that corresponds to p_0, back to the computed result.

In Section 8 we shall present numerical results from the application of this procedure.

8 Numerical experiments for the Shepp- Logan phantom

In this section we shall present numerical experiments applied to the Shepp-Logan phantom shown in Fig. 12 and defined in Appendix A.

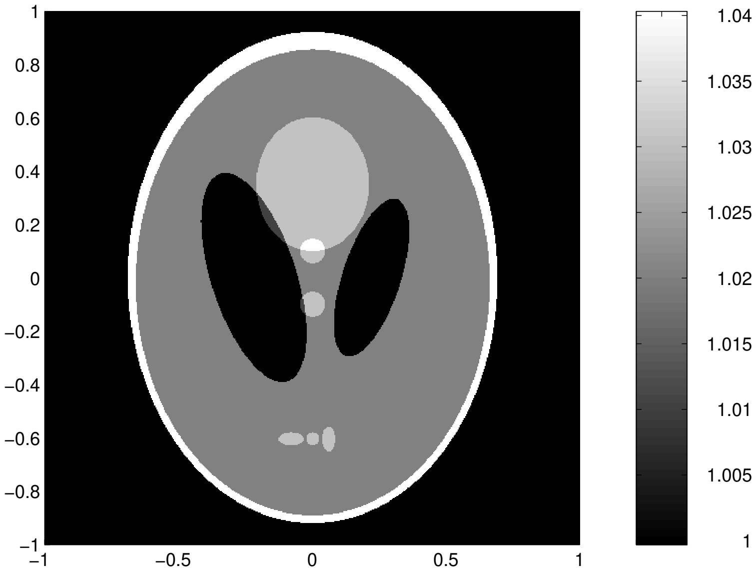

Figure 12: The Shepp-Logan phantom.

The domain $\{-1 \leq r \leq 1\,,\ 0 \leq \theta < \pi\}$ is discretized using a 512×256 grid, and the corresponding line integrals $p(r, \theta)$ are computed by using the trapezoidal rule with 2048 points along each line.

In Fig 13 we first show the result from a computation where the formula (5.18) for the Fourier coefficients is substituted by traditional bilinear interpolation.

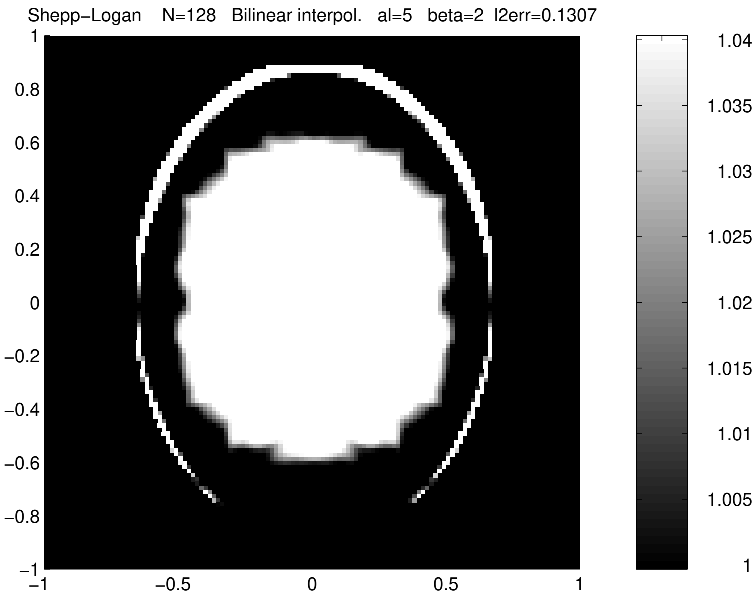

Figure 13: Shepp-Logan, $N = 128$, *Bilinear interpolation*, $\alpha = 5$, $\beta = 2$.

Obviously the inner part of the image is completely wrong, confirming the theoretical analysis in Section 4.

We next illustrate the efficacy of the two-part summation (5.24). Fig. 14 shows the reconstructed image for $m_1 = 32\,,\ m_2 = -1$, i.e. the second part of the sum is not included at all. The result is quite poor. With $m_1 = -1\,,\ m_2 = 32$, i.e. the first part of the sum is not included at all, we also get a poor result as shown in Fig. 15.

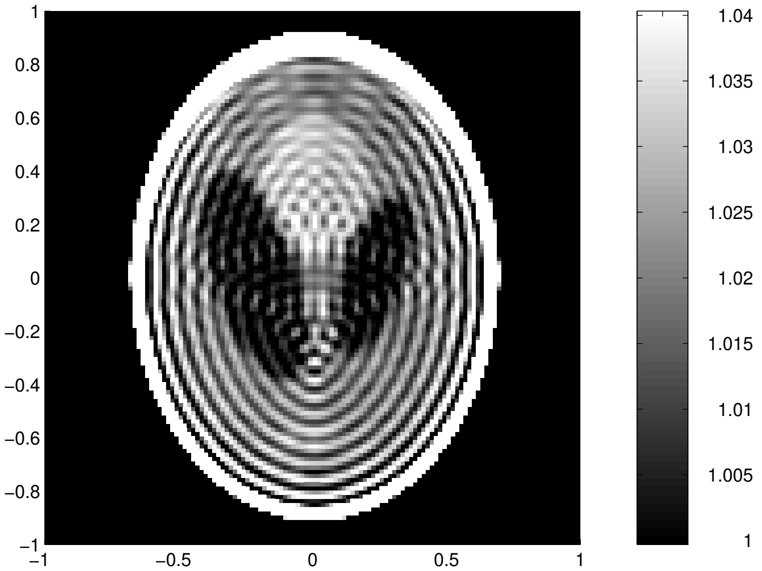

Figure 14: Shepp-Logan, $N = 128$, $m_1 = 32$, $\alpha = 5$, $\beta = 2$. No second part in the interpolating sum.

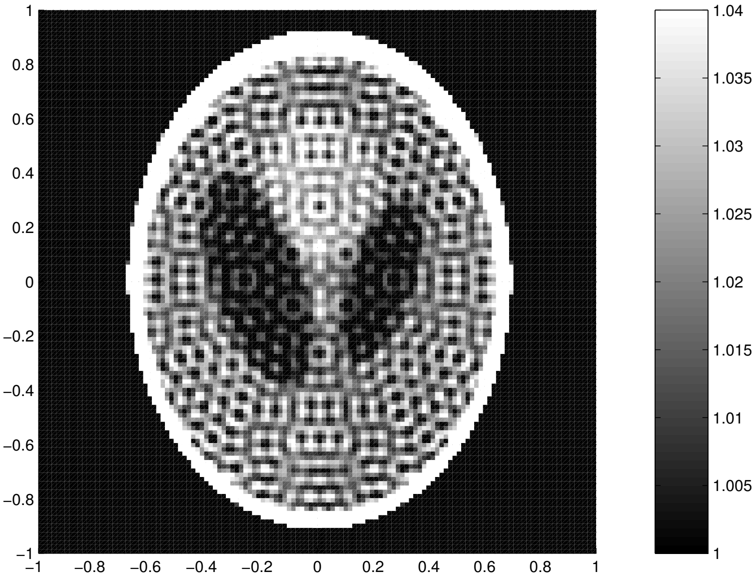

Figure 15: Shepp-Logan, $N = 128$, $m_2 = 32$, $\alpha = 5$, $\beta = 2$. No first part in the interpolating sum.

Essentially the same cost is obtained with $m_1 = 19$, $m_2 = 13$, where the number of terms has been chosen according to (6.29). A significant improvement is now obtained, as shown in Fig. 16. In Fig. 17 the result of finer resolution is shown.

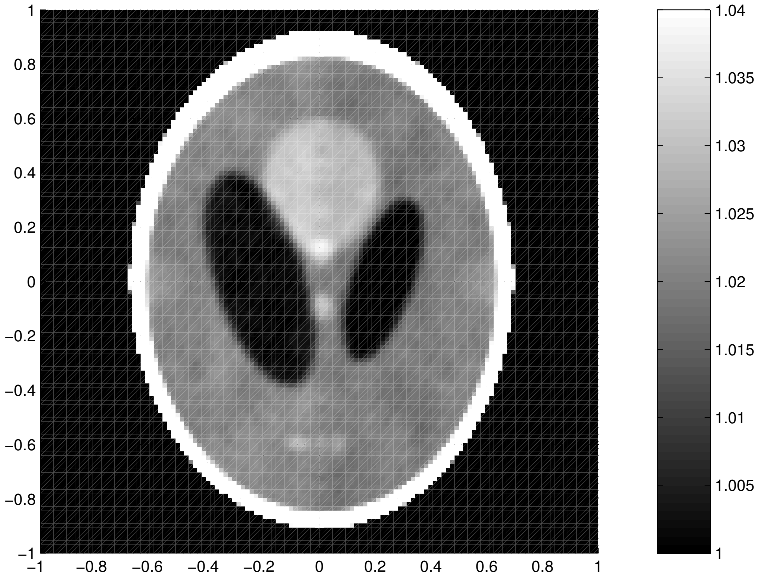

Figure 16: Shepp-Logan, $N = 128,\ m_1 = 19,\ m_2 = 13,\ \alpha = 5,\ \beta = 2.$

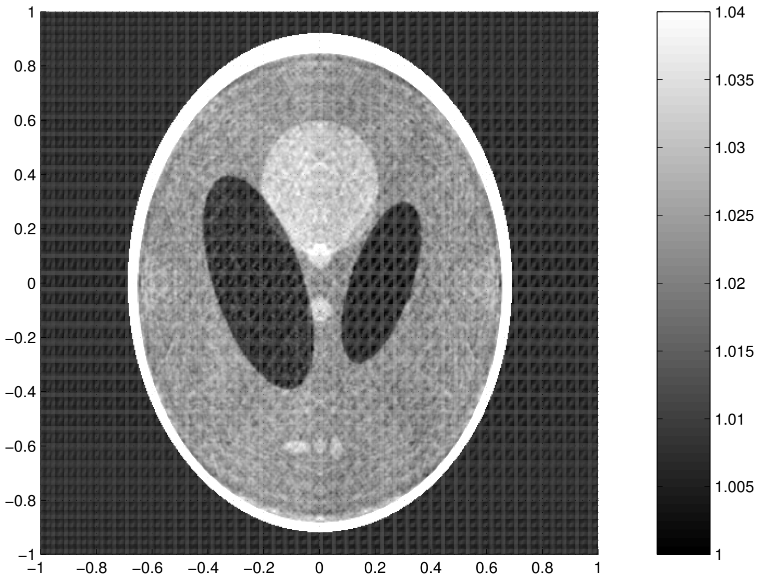

Figure 17: Shepp-Logan, $N = 512,\ m_1 = 19,\ m_2 = 13,\ \alpha = 5,\ \beta = 2,\ l_2$-error=0.0384.

Finally, we present the result from a reconstruction where the method of subtracting the large discontinuities has been applied. The outer part of the phantom, simulating the skull, has a much higher density than the rest. The edges of the two ellipses defining the skull are known, and since the density is constant, we can subtract the function p defined by (2.5) in Example 3 of Section 2. We then applied the FSM including the two-part summation formula (5.24) to the modified problem. The result, after adding the outer skull again, is shown in Fig 18. The error in the l_2-norm is now 0.0005 compared to 0.04 for the original method.

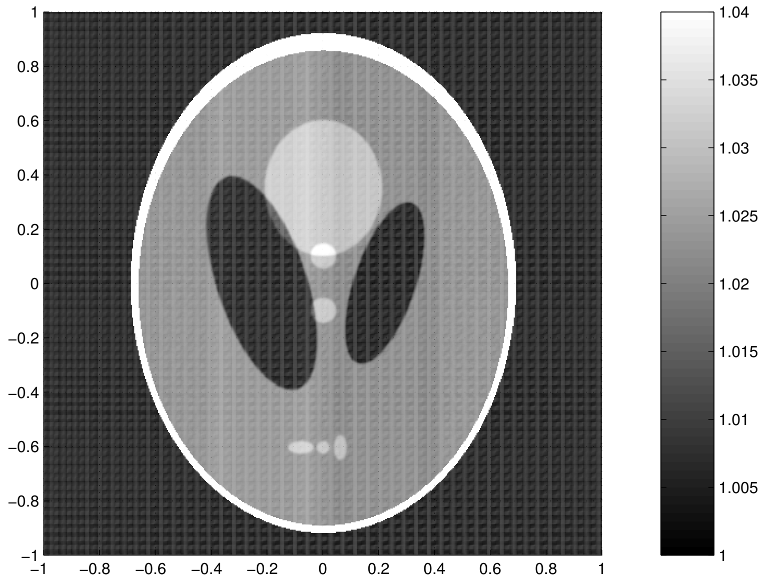

Figure 18: Shepp-Logan, $N = 512$, $m_1 = 19$, $m_2 = 13$, $\alpha = 5$, $\beta = 2$, l_2-error=0.0005.

Acknowledgment

Our thanks are due to Uno Eriksson, Uppsala University Hospital, who brought the tomography problem to our attention. Our thanks are also due to Per-Erik Danielsson, Linköping University for valuable comments on the manuscript.

A The Shepp-Logan phantom

The Shepp-Logan phantom consists of ten ellipses E_j partly overlapping each other, each one with constant density f_j. If f_j is defined as zero outside E_j, then at any given point (x, y), the total density is

$$f(x,y) = \sum_{j=1}^{10} f_j(x,y)\,.$$

Note that f_j may be negative. The following algorithm defines the density f.

$f = 0$
$\text{if}(x/0.69)^2 + (y/0.92)^2 \leq 1 \text{ then } z = 2$
$\text{if}(x/0.6624)^2 + ((y + 0.0184)/0.874)^2 \leq 1 \text{ then } z = z - 0.98$
$xi = (x - 0.22) * \cos(0.4 * pi) + y * \sin(0.4 * pi)$
$eta = -(x - 0.22) * \sin(0.4 * pi) + y * \cos(0.4 * pi)$
$\text{if}(xi/0.31)^2 + (eta/0.11)^2 \leq 1 \text{ then } z = z - 0.02$
$xi = (x + 0.22) * \cos(0.6 * pi) + y * \sin(0.6 * pi)$
$eta = -(x + 0.22) * \sin(0.6 * pi) + y * \cos(0.6 * pi)$
$\text{if}(xi/0.41)^2 + (eta/0.16)^2 \leq 1 \text{ then } z = z - 0.02$
$\text{if}(x/0.21)^2 + ((y - 0.35)/0.25)^2 \leq 1 \text{ then } z = z + 0.01$
$\text{if}(x/0.046)^2 + ((y - 0.1)/0.046)^2 \leq 1 \text{ then } z = z + 0.01$
$\text{if}(x/0.046)^2 + ((y + 0.1)/0.046)^2 \leq 1 \text{ then } z = z + 0.01$
$\text{if}((x + 0.08)/0.046)^2 + ((y + 0.605)/0.023)^2 \leq 1 \text{ then } z = z + 0.01$
$\text{if}(x/0.023)^2 + ((y + 0.605)/0.023)^2 \leq 1 \text{ then } z = z + 0.01$
$\text{if}((x - 0.06)/0.023)^2 + ((y + 0.605)/0.046)^2 \leq 1 \text{ then } z = z + 0.01$

References

[1] S. Abarbanel and D. Gottlieb: *Information content in spectral calculations.* "Progress and Super-computing in CFD", Birkhauser, pp 345-356, (1985).

[2] S. R. Deans: *The Radon Transform and Some of Its Applications.* John Wiley and Sons, (1983).

[3] D. Gottlieb and C.W. Shu: *On the Gibbs Phenomenon and its Resolution.* SIAM Review, Vol. 39, No. 4, pp 644-668 (1997).

[4] B. Gustafsson: *Mathematics for Computer Tomography.* Physica Scripta, Vol. T61, pp 38-43, (1996).

[5] M. Magnusson: *Linogram and Other Direct Fourier Methods for Tomographic Reconstruction.* Ph.D thesis, Linköping University (1993).

[6] J. D. O'Sullivan: *A Fast Sinc Function Gridding Algorithm for Fourier Inversion in Computer Tomography.* IEEE Transactions on Medical Imaging, Vol. MI-4, No. 4, pp 200-207, (1985).

[7] H. Schomberg and J. Timmer: *The Gridding method for Image Reconstruction by Fourier Transformation.* IEEE Transactions on Medical Imaging, Vol. 14, No. 3, pp 596-607, (1995).

[8] E. Tadmor: *The exponential accuracy of Fourier and Chebyshev differencing methods.* SIAM J. Numer. Anal , Vol. 23, No. 1 pp 1-10, (1986).