COMBINING IMAGE RECONSTRUCTION AND IMAGE ANALYSIS WITH AN APPLICATION TO 2D – TOMOGRAPHY

ALFRED K. LOUIS*

Abstract. The tasks of image reconstruction from measured data and the analysis of the so produced images are more or less strictly separated. One group computes by applying reconstruction algorithms the images, the other starts out of that by operating on these images to enhance the analysis. First attempts, in a non - systematic way, are known, as Lambda tomography or Tikhonov - Phillips methods with ℓ_1 – norms or with level – set methods.

The aim of this paper is to provide a general tool to combine these two steps; i.e., already in the reconstruction step the future image analysis step is taken into account leading to a new reconstruction kernel. Here we concentrate on linear methods.

As practical example we consider the image reconstruction problem in computerized tomography followed by an edge detection. We calculate a new reconstruction kernel and present results from simulations.

Key words. Image analysis, image reconstruction, approximate inverse

AMS subject classifications. 65R32, 45Q05

1. Introduction. In order to extract information from a given image, analysis tools are used. In a first step one applies operators on that image, then the searched – for information is found by operating on these enhanced versions of the original picture. Images typically are two-dimensional arrays of numbers. Of course, three-dimensional arrays for volume data or even time - dependent data, which may amount to four-dimensional data, are conceivable. Prominent analysis tools are edge detection methods where first partial derivatives of smoothed versions of the image are computed, followed then by recognition methods. A typical example is the Canny edge detector, see [4]. Other operations can be found e.g. in [6, 11]. Here we restrict, as above mentioned, to linear operators. In denoising one can think of solving the heat equation with homogeneous boundary condition and initial condition as the original image, at the final time T the image is considered to be denoised.

We have to mention that of course also non - linear methods play an essential role. But this does – at least at the moment – not fit in our framework.

Attempts to combine reconstruction and analysis are known, but not systematically pursued. As examples we mention the Λ – CT, where local inversion formulas produce images where the singular support is preserved, which means, that those images have jumps wherever the original image has them. See e.g. [10, 13, 21]. The use of Tikhonov – Phillips regularization with ℓ_1 – norms results in smooth images, see e.g. [5]. Level – set methods in combination with tomography data lead to the determination of the boundary of the object, at least if the objects are relatively smooth with jumps along smooth curves, see e.g. [29]. Another possibility is the direct calculation of wavelet coefficients of the searched - for solution, see [22] and for an application to tomography see [3]. These coefficients may be used in classification algorithms or in local reconstructions, see [26]. Here as mollifier, see next section, we use the scaling function and the wavelets.

In order to make the statements more precise we consider the following example. In computerized tomography the images are produced by applying reconstruction al-

*Department of Applied Mathematics, Saarland University D-66041 Saarbrücken, Germany (`louis@num.uni-sb.de`).

gorithms to the measured data. In that way one calculates images which are smoothed version of the original density distributions. The result can be presented as

$$f_\gamma = f * e_\gamma =: E_\gamma f$$

where f is the original object and e_γ is a mollifier depending on the reconstruction method. In the image analysis part, for example in the above mentioned edge detection methods, one computes then derivatives of smoothed versions of this image. Typically one calculates in a first step

$$f_{\gamma\beta k} = \frac{\partial}{\partial x_k}\big(G_\beta * f_\gamma\big) = \frac{\partial}{\partial x_k}\big(W_\beta f_\gamma\big)$$
$$= \frac{\partial}{\partial x_k}\big(W_\beta E_\gamma f\big)$$

where G_β represents a mollifier, for example a Gaussian kernel, and where the two parameters β and γ are chosen independently. The aim of this paper is to provide a method which allows for directly computing in one step the smoothed version of the derivative. To this end in the next chapter we generalize the concept of approximate inverse as introduced in [20]. We precompute independently of the data g a reconstruction kernel ψ_γ by solving an auxiliary problem $A^*\psi_\gamma = e_\gamma$. Then the solution is calculated as $g * \psi_\gamma$. A further advantage is that invariances of the operator combined with suitable mollifiers lead to very efficient reconstruction methods.

In Chapter 2 we present some basic facts about linear ill–posed problems, we then introduce the approximate inverse for combining the two steps on regularization and analyzing. In Chapter 4 we study the regularization properties of the new method. Chapter 5 is devoted to the efficient calculation of the result using invariances of the included operators. Finally in the last section we present results for the case of tomography in combination with edge detection, we derive a new filter and present results from simulations, showing that the results obtained in this way are better than with the classical approach of separately performing reconstruction and differentiation. In addition the computing is much quicker.

2. Linear Ill - Posed Problems. We consider a continuous mapping A between the Hilbert spaces X and Y. The problem (A, X, Y) is called well - posed if $Af = g$ has a unique solution that depends continuously on the data. If one of those conditions is not fulfilled the problem is called ill - posed. It is important to include the spaces in this definition, then by changing the spaces we may get well-posed problems. The reason for choosing the given spaces, is that on one hand the data, including the noise, are mostly not smooth enough to choose a smaller space Y. On the other hand, by selecting a larger space X we may change the concept of solution, including for examples distributions.

Many integral equations of the first kind lead to compact operators between X and Y, which means that if the operator does not have finite rank the range of Y, denoted by $\mathcal{R}(A)$, is not closed and hence the inverse is not continuous. In order to define solutions for these in the classical sense not necessarily solvable problems we introduce the pseudo - inverse with domain of definition $\mathcal{D}(A^\dagger) = \mathcal{R}(A) \oplus \mathcal{N}(A^*) \subset \overline{\mathcal{R}(A)} \oplus \mathcal{N}(A^*) = Y$ mapping $g \in \mathcal{D}(A^\dagger)$ to the uniquely determined $f \in \mathcal{N}(A)^\perp \subset X$ which solves

$$Af = P_{\overline{\mathcal{R}(A)}}g \qquad (2.1)$$

where $P_{\overline{\mathcal{R}(A)}}$ is the orthogonal projection onto the closure of the range of A. Hence the null - space of the so defined $A^\dagger$ is $\mathcal{N}(A^\dagger) = \mathcal{N}(A^*)$.

To measure the degree of ill - posedness, which is important to select the appropriate regularization, in principle two concepts are used. One is based on the decay of the singular values of the compact operator A, see [17]. Another possibility is to consider the smoothing properties of A as introduced by Natterer, see [24], if the spaces are based on L_2 – spaces. We say that A smoothes α steps in a Sobolev scale if

$$c_1\|f\|_{H^{-\alpha}} \leq \|Af\|_{L_2} \leq c_2\|f\|_{H^{-\alpha}} \tag{2.2}$$

or

$$c_1\|f\|_{L_2} \leq \|Af\|_{H^\alpha} \leq c_2\|f\|_{L_2} \tag{2.3}$$

where the Sobolev norms for functions in $\mathrm{I\!R}^N$ are defined as

$$\|f\|^2_{H^\alpha} = \int_{\mathrm{I\!R}^N} \left(1 + |\xi|^2\right)^\alpha |\hat{f}(\xi)|^2 d\xi \tag{2.4}$$

with the Fourier transform

$$\hat{f}(\xi) = (2\pi)^{-N/2} \int_{\mathrm{I\!R}^N} f(x) \exp\left(-\imath \xi^\top x\right) dx \tag{2.5}$$

In the case of Fourier integral operators this definition coincides with the fact that the singular values σ_n of A as mapping between L_2 spaces decay like $O(n^{-\alpha})$. We say, the problem $\left(A, L_2, L_2\right)$ is ill – posed of order α.

In this paper we use the smoothing properties in Sobolev scales to measure the degree of ill - posedness for spaces $X = L_2(U)$ and $Y = L_2(V)$ for suitable domains U and V.

The theory of regularization is concerned with the definition of solutions for arbitrary data in Y with the additional aspect of balancing the influence of the unavoidable data error against the best possible resolution in the reconstruction. This is achieved by constructing operators $T_\gamma : Y \to X$ with the property that

$$\lim_{\varepsilon \to 0} T_{\gamma(\varepsilon, g^\varepsilon)} g^\varepsilon = A^\dagger g$$

when $g \in \mathcal{D}(A^\dagger)$ and the erroneous data g^ε go to g for $\varepsilon \to 0$.

It is shown in [19] that many of the well - known regularization methods, including the truncated singular value decomposition or the Tikhonov - Phillips method or iterative methods like Landweber or CG are, as well as the approximate inverse, of the form

$$T_\gamma = M_\gamma \tilde{A}^\dagger$$

with a smoothing operator M_γ and a suitable continuation of $A^\dagger$ to all of Y. Another possibility is to first smooth the data and then to invert.

In order to extend the operator $A^\dagger$ to all of $Y = \overline{\mathcal{R}(A)} \oplus \mathcal{N}(A^*)$ we put on $\mathcal{N}(A^*)$

$$\tilde{A}^\dagger g = A^\dagger g = 0 \ , \ g \in \mathcal{N}(A^*)$$

The right – hand side of condition (2.3) says that A maps $L_2(U)$ continuously to $H^\alpha(V) \subset L_2(V)$ for $\alpha > 0$, hence $\mathcal{R}(A) \subset H^\alpha(V) \subset L_2(V)$. The left - hand side of condition (2.2) says that A is continuously invertible from $\mathcal{N}(A)^\perp \subset L_2(V)$ to $H^{-\alpha}(U)$ where the norm of the inverse is bounded by c_1^{-1}. Hence we define $\tilde{A}^\dagger$ on

$\overline{\mathcal{R}(A)}$ as this inverse observing that on $\mathcal{R}(A)$ this coincides with the definition of $A^\dagger$ due to (2.1) and get

THEOREM 2.1. *The continuation of the pseudo–inverse $A^\dagger$ to all of $Y = L_2(V)$ is a mapping with $\tilde{A}^\dagger : L_2(V) \to H^{-\alpha}(U)$ with $\|\tilde{A}^\dagger\| \le c_1^{-1}$.*

Proof. We decompose $g \in L_2(V)$ into $g = g_1 + g_0$ with $g_1 \in \overline{\mathcal{R}(A)}$ and $g_0 \in \mathcal{N}(A^*)$, then due to the above construction and (2.2) we get

$$\|\tilde{A}^\dagger g\|_{H^{-\alpha}} \le c_1^{-1}\|g_1\|_{L_2} \le c_1^{-1}\|g\|_{L_2}$$

which completes the proof. $\square$

If we use differential operators in the image analysis step then we consider, with the same arguments as above, these operators L as mapping

$$L : \mathcal{D}(L) \subset X \to X$$

in order not to change the space where the solution or the approximate solution is presented. Obviously, differential operators are unbounded mappings when considered as

$$L : \mathcal{D}(L) \subset L_2(U) \to L_2(U).$$

In order to measure the degree of making the functions less smooth we assume for a $t > 0$

$$\|Lf\|_{H^{s-t}} \le c_{L,s}\|f\|_{H^s} \tag{2.6}$$

i.e., L is a differential operator or a pseudo - differential operator of order t. The problem of determining Lf from $Af = g$ is then ill - posed of order $\alpha + t$, which means that the ill – posedness is enhanced. When first dealing with ill - posed problems one might think, that by choosing the right spaces the problem becomes well - posed, but this is impossible, as the above discussion shows.

3. Approximate Inverse for Combining Reconstruction and Analysis. The motivation for the approximate inverse where the problem $Af = g$ is stably solved by

$$T_\gamma g = \langle g, \psi_\gamma \rangle$$

with a reconstruction kernel fulfilling $A^*\psi_\gamma = e_\gamma$ with a prescribed mollifier e_γ is at least twofold. Firstly the calculation of those functionals of the solution may be stably achieved, in contrast to the calculation of the solution itself. This was already observed with the Backus – Gilbert method, [1]. The method is too time – consuming to evaluate the solution at all points. See also Eckhardt, [8] for the calculation of the derivative of the solution. In connection with tomography it was observed by Grünbaum, [7], that the filtered backprojection leads to a mollified version of the searched–for solution. For an early application of the calculation of functionals of the solution to partial differential equations see [15] and the recent paper by Ovall [27]. The second reason for the introduction of the approximate inverse was to derive fast inversion formulae for the case when the same problem has to be solved repeatedly with different right–hand sides, as it is the case for measuring devices as X–ray scanners. Essentially for the method to be fast is the selection of the mollifier e_γ according to the invariances of the operator A, as was observed in [17], where also applications

to some non – linear problems are treated. The essential difference to the mollifica-tion method of Murio, see e.g. [23], is that $f * e_\gamma$ does not replace f in the equation $Af = g$, where in that way the kernel of the integral operator is smoothed, which even amplifies the ill – posedness of the problem.

In this section we generalize the method of the approximate inverse as analyzed in [18]. Let $A : X \to Y$ be a linear operator between the Hilbert spaces X and Y and $L : X \to Z$ be a linear, possibly unbounded, operator between the Hilbert spaces X and Z. Typical examples that we have in mind are differential operators where $Z = X$ and $L : X \to X$ is unbounded, or $Z = \ell_2$ when we compute the wavelet coefficients of the solution or $Z = \mathrm{I\!R}^N$ when we compute N of those coefficients. In the last two cases the operator L is bounded.

In the reconstruction part we have to solve

$$Af = g \qquad (3.1)$$

and then we apply the operation L on the so computed solution f for the image analysis.

Now we adapt the concept of approximate inverse, first introduced in [20]. In [8, 18] already derivatives of the solution of $Af = g$ are directly calculated. We now compute instead of Lf an approximation

$$(Lf)_\gamma = \langle Lf, e_\gamma \rangle$$

with a prescribed mollifier $e_\gamma(x, \cdot) \in X$. The value x depends on the application. In the situation $X = L_2(U)$ and $L : X \to X$ then $x \in U$ is the reconstruction point where Lf is evaluated. In the wavelet application x is the index of the scaling function or the wavelet coefficient. We formulate in the following theorem the principle of the reconstruction method, the technical details, as conditions on the mollifier, are treated in the next section.

THEOREM 3.1. *Let $e_\gamma(x, \cdot) \in X$ be a suitably chosen mollifier and $\psi_\gamma(x, \cdot) \in Y$ be the solution of the auxiliary problem*

$$A^* \psi_\gamma(x, \cdot) = L^* e_\gamma(x, \cdot) \qquad (3.2)$$

Then the smoothed version of the image analysis operation is directly computed from the given data g as

$$(Lf)_\gamma(x) = \langle g, \psi_\gamma(x, \cdot) \rangle \qquad (3.3)$$

Proof. We write the smoothed version of the image analysis part as

$$(Lf)_\gamma(x) = \langle Lf, e_\gamma(x, \cdot) \rangle$$

Now we use the adjoint operator of L and the auxiliary problem to continue

$$\begin{aligned}
(Lf)_\gamma(x) &= \langle f, L^* e_\gamma(x, \cdot) \rangle \\
&= \langle f, A^* \psi_\gamma(x, \cdot) \rangle \\
&= \langle g, \psi_\gamma(x, \cdot) \rangle
\end{aligned}$$

where in the last step we have used the original equation $Af = g$. □ We remark that if the auxiliary problem is not solvable; i.e., if $L^* e_\gamma$ is not in $\mathcal{R}(A^*) \subset \mathcal{N}(A)^\perp$ then we

solve the normal equation $AA^*\psi_\gamma = AL^*e_\gamma$ which still leads to a regularized pseudo - solution of the problem of finding Lf.

DEFINITION 3.2. *The operator $S_\gamma : Y \to Z$ defined as*

$$S_\gamma g = \langle g, \psi_\gamma(x, \cdot) \rangle \tag{3.4}$$

is called the approximate inverse *of A to compute an approximation of Lf and ψ_γ is called the* reconstruction kernel.

4. Regularization Method. In this section we study the smoothness conditions necessary to guarantee a suitable solution to the whole problem generalizing the results of [19, 12] in the framework of ill – posed problems [9, 17]. Let A be a linear operator between the Hilbert spaces X and Y and L be a linear operator between the Hilbert spaces X and Z.

DEFINITION 4.1. *A regularization of $A_L^\dagger := LA^\dagger$ for finding the enhanced solution $Lf \in Z$ of $Af = g$ and the application of the image analysis operator L is a family of operators*

$$\{T_\gamma\}_{\gamma > 0}, \ T_\gamma : Y \to Z$$

with a mapping: $\gamma : \mathbb{R}^+ \times Y \to \mathbb{R}^+$, such that for all $g \in \mathcal{D}(A_L^\dagger)$ and for all $g^\epsilon \in Y$ with $\|g - g^\epsilon\| \le \epsilon$ the equality

$$\lim_{\epsilon \to 0, g^\epsilon \to g} T_{\gamma(\epsilon, g^\epsilon)} g^\epsilon = LA^\dagger g$$

holds.

If the operator L is bounded then clearly any regularization T_γ for $A^\dagger$ leads with LT_γ to a regularization of $LA^\dagger$. Hence we discuss in the following unbounded operators and use the notation introduced in Section 2. The problem $(A, L_2(U), L_2(V))$ is ill – posed of order α and L is a pseudodifferential operator of order t. In this setting the pseudo – inverse $\tilde{A}^\dagger$ maps $L_2(V)$ to the space $H^{-\alpha}(U)$ and then L to $H^{-(\alpha+t)}(U)$. This space is too large, hence we need some smoothing operator to come back from this large space $H^{-(\alpha+t)}(U)$ to $L_2(U)$. In the rest of this section the spaces are considered as spaces over U, hence we omit U.

THEOREM 4.2. *Let $M_\gamma : H^{-(\alpha+t)} \to L_2$ be a family of linear continuous operators such that*

$$\begin{aligned} &i) \|M_\gamma f\|_{L_2} \le c(\gamma) \|f\|_{H^{-(\alpha+t)}} \quad \textit{for all } f \in \mathcal{N}(A)^\perp, \\ &ii) \quad \lim_{\gamma \to 0} \|M_\gamma Lf - Lf\| = 0 \qquad \textit{for all } f \in \mathcal{N}(A)^\perp. \end{aligned}$$

Then $T_\gamma = M_\gamma L\tilde{A}^\dagger$ is a regularization of $A_L^\dagger$ for finding Lf if we chose γ in such a way that $c(\gamma)\varepsilon \to 0$ for $\varepsilon \to 0$.

Proof. Let $g \in \mathcal{D}(A^\dagger)$ and $g^\epsilon \in L_2$ such that $\|g^\epsilon - g\|_{L_2} \le \epsilon$, then we get with $A^\dagger g = \tilde{A}^\dagger g \quad$ for all $\ g \in \mathcal{D}(A^\dagger)$:

$$\begin{aligned} \|T_\gamma g^\epsilon - A_L^\dagger g\| &\le \|T_\gamma(g^\epsilon - g)\| + \|T_\gamma g - A_L^\dagger g\| \\ &= \|M_\gamma L\tilde{A}^\dagger(g^\epsilon - g)\| + \|T_\gamma g - \tilde{A}_L^\dagger g\| \\ &\le c(\gamma)\|L\tilde{A}^\dagger(g^\epsilon - g)\|_{H^{-(\alpha+t)}} + \|M_\gamma L\tilde{A}^\dagger g - L\tilde{A}^\dagger g\| \\ &\le c(\gamma)c_{L,-\alpha}c_1^{-1}\epsilon + \|M_\gamma L\tilde{A}^\dagger g - L\tilde{A}^\dagger g\| \\ &\overset{\epsilon \to 0}{\longrightarrow} 0 \end{aligned}$$

for $\epsilon \to 0$ and γ such that $c(\gamma)\epsilon \to 0$. □

We now look for conditions for e_γ from Section 3 in order to guarantee that the method presented there is a regularization. The function $e_\gamma(x,y)$ is defined for $x, y \in U$. If we consider a mollifier e_γ of convolution type in $\mathrm{I\!R}^N$, then we can derive the following result. We denote the Fourier transform of e_γ by $\hat{e}_\gamma$.

THEOREM 4.3. *Let $e_\gamma(x,y)$ be of convolution type, i.e. $e_\gamma(x,y) = e_\gamma(x-y)$, and*

$$i) (2\pi)^{N/2} \sup_\xi \{(1 + |\xi|^2)^{(\alpha+t)/2} |\hat{e}_\gamma(\xi)|\} \le c(\gamma),$$

$$ii) \quad \sup_{\xi \in \mathrm{I\!R}^N} \left(|(2\pi)^{N/2} \hat{e}_\gamma(\xi) - 1| \right) \xrightarrow{\gamma \to 0} 0$$

Then $T_\gamma g = \langle e_\gamma(x,.), L\tilde{A}^\dagger g \rangle$ is a regularization of $A_L^\dagger$ for finding Lf.

Proof. We check the conditions of Theorem 4.2.

$$\|M_\gamma f\|^2 = \|\mathcal{F}(M_\gamma f)\|^2$$
$$= (2\pi)^N \int_{\mathrm{I\!R}^N} |\hat{e}_\gamma(\xi)\hat{f}(\xi)|^2 d\xi$$
$$= (2\pi)^N \int_{\mathrm{I\!R}^N} (1 + |\xi|^2)^{-(\alpha+t)} (1 + |\xi|^2)^{(\alpha+t)} |\hat{e}_\gamma(\xi)|^2 |\hat{f}(\xi)|^2 d\xi$$
$$\le \underbrace{(2\pi)^N \sup_\xi \{(1 + |\xi|^2)^{(\alpha+t)/2} |\hat{e}_\gamma(\xi)|\}^2}_{c(\gamma)^2} \|f\|^2_{H^{-(\alpha+t)}}.$$

which proves part i) in Theorem 4.2.

$$\|M_\gamma Lf - Lf\|^2 = \|\mathcal{F}(M_\gamma Lf - Lf)\|^2 = \int_{\mathrm{I\!R}^N} |(2\pi)^{N/2} \hat{e}_\gamma(\xi) - 1|^2 |\hat{Lf}(\xi)|^2 d\xi$$
$$\le \sup_{\xi \in \mathrm{I\!R}^N} \left(|(2\pi)^{N/2} \hat{e}_\gamma(\xi) - 1| \right) \|Lf\|^2_{L_2}$$
$$\longrightarrow 0.$$

which proves part ii) in Theorem 4.2. □

EXAMPLE:
Let $N = 2$ and e_γ be such that

$$\hat{e}_\gamma(\xi) = (2\pi)^{-1} \mathrm{sinc} \frac{\gamma|\xi|\pi}{2} \chi_{[-1/\gamma, 1/\gamma]}(|\xi|)$$

with $\mathrm{sinc}\, x = \sin x / x$. Checking the conditions of Theorem 4.3 we obtain:

$$i) (2\pi)^{1/2} \sup_\xi \{(1 + |\xi|^2)^{(\alpha+t)/2} |\hat{e}_\gamma(|\xi|)|\} = \left(1 + \left| \frac{1}{\gamma} \right|^2 \right)^{(\alpha+t)/2} = c(\gamma),$$

$$ii) \quad \sup_\xi |2\pi \hat{e}_\gamma(\xi) - 1| \to 0 \text{ for } \gamma \to 0$$

where we have used that $\mathrm{sinc}\, (x)| \le 1$ and $\lim_{\gamma \to 0} \mathrm{sinc}\, (\gamma x) = \mathrm{sinc}\, (0) = 1$. We also note that $\sup_\xi |\hat{e}(\xi)| \le (2\pi)^{-1}$. Hence $E_\gamma \tilde{A}_L^\dagger$ is a regularization of $A_L^\dagger$.

5. Invariances. The computational efficiency of the approximate inverse heavily depends on the use of invariances. We mention again the reconstruction problem in tomography. If we chose for each reconstruction point x a special mollifier, namely $e_\gamma(x, \cdot)$, then the reconstruction kernel also depends on x, the number of values to store is then the number of reconstruction points times the number of data. If we use invariances, for example translation and rotational invariances of the Radon transform and we use these invariances to produce the mollifier we can reduce this number of values to compute and store to just the number of views per direction. The mathematical basis for this can be found in [18]. Here we derive the corresponding result for the combination of reconstruction and image analysis.

THEOREM 5.1. *Let $A : X \to Y$ and $L : X \to Z$ be the two operators as above. Let*

$$T_1 : Z \to Z$$
$$T_2 : X \to X$$
$$T_3 : Y \to Y$$

be linear operators with

$$L^* T_1 = T_2 L^* \qquad (5.1)$$
$$T_2 A^* = A^* T_3 \qquad (5.2)$$

and let Ψ_γ be the solution of the auxiliary problem for a general mollifier $E_\gamma \in \mathcal{D}(L^*)$

$$A^* \Psi_\gamma = L^* E_\gamma \qquad (5.3)$$

Then the solution for the special mollifier

$$e_\gamma = T_1 E_\gamma \qquad (5.4)$$

is

$$\psi_\gamma = T_3 \Psi_\gamma \qquad (5.5)$$

Proof. We start with the right - hand side of the auxiliary problem and use the above relations to get

$$L^* e_\gamma = L^* T_1 E_\gamma = T_2 L^* E_\gamma = T_2 A^* \Psi_\gamma = A^* T_3 \Psi_\gamma ,$$

hence $T_3 \Psi_\gamma$ solves the auxiliary problem. □

As a consequence we observe that the solution for a special mollifier fulfilling the condition $e_\gamma = T_1 E_\gamma$ can be found as

$$\langle f, e_\gamma \rangle = \langle g, T_3 \Psi_\gamma \rangle .$$

If for example the operators A and L are of convolution type and if we chose the mollifier e_γ also of convolution type, then the mappings T_k are all of translation type, which means that also the final reconstruction formula is of convolution type.

6. Tomography and Edge Detector. The mathematical model of computerized tomography in two dimensions, for the parallel geometry, is the Radon transform, see e.g. [25]. It is defined as

$$\mathbf{R}f(\theta, s) = \int_{\mathbb{R}^2} f(x)\delta(s - x^\top \theta)dx$$

where $\theta \in S^1$ is a unit vector and $s \in \mathbb{R}$.
We consider $\mathbf{R}$ as mapping

$$\mathbf{R} : L_2(\Omega) \to L_2(S^1 \times \mathbb{R})$$

where Ω is a bounded domain in $\mathbb{R}^2$. In the notation of Section 2 we have $A = \mathbf{R}$, $U = \Omega$ and $V = S^1 \times \mathbb{R}$. The Radon transform has a trivial null space. The relations (2.2) and (2.3) hold with $\alpha = 1/2$, see [25].
In the following we summarize a few results. The central slice theorem, or projection theorem is nothing but the formal application of the adjoint operator for fixed direction θ on $\exp(\imath s\sigma)$

$$\widehat{\mathbf{R}f}(\theta, \sigma) = (2\pi)^{1/2} \hat{f}(\sigma\theta) \tag{6.1}$$

The Radon transform of a derivative is

$$\mathbf{R}\frac{\partial}{\partial x_k} f(\theta, s) = \theta_k \frac{\partial}{\partial s} \mathbf{R}f(\theta, s) \tag{6.2}$$

see e.g. [25], and generalizations for higher derivatives. The inversion formula for the two – dimensional Radon transform is

$$\mathbf{R}^{-1} = \frac{1}{4\pi} \mathbf{R}^* \mathbf{I}^{-1} \tag{6.3}$$

where $\mathbf{R}^*$ is the adjoint operator from L_2 to L_2 known as backprojection

$$\mathbf{R}^* g(x) = \int_{S^1} g(\theta, x^\top \theta)d\theta$$

and the Riesz potential $\mathbf{I}^{-1}$ is defined with the Fourier transform

$$\widehat{\mathbf{I}^{-1}g}(\theta, \sigma) = |\sigma|\hat{g}(\theta, \sigma)$$

where the Fourier transform acts on the second variable.
The following invariances are well established for the Radon transform. Consider for $x \in \mathbb{R}^2$ the shift operators $T_2^x f(y) = f(y - x)$ and $T_3^{x^\top \theta} g(\theta, s) = g(\theta, s - x^\top \theta)$ then

$$\mathbf{R}T_2^x = T_3^{x^\top \theta} \mathbf{R} \tag{6.4}$$

Another couple of intertwining operators is found by rotation. Let Θ be a unitary 2×2 matrix and $D_2^\Theta f(y) = f(\Theta y)$. then

$$\mathbf{R}D_2^\Theta = D_3^\Theta \mathbf{R} \tag{6.5}$$

where $D_3^\Theta g(\theta, s) = g(\Theta\theta, s)$. With the $(T\mathbf{R})^* = \mathbf{R}^* T^*$ we get the relations used in Theorem 5.1. These two invariances lead for a mollifier of convolution type and

independent of the directions; i.e., $e_\gamma(x, y) = E_\gamma(\|x - y\|)$, to a reconstruction kernel for determining f of convolution type, independent of the direction, namely $\psi_\gamma(x; \theta, s) = \Psi_\gamma(s - x^\top \theta)$.

For the edge detectors we use differential operators $\mathbf{L}_k = \frac{\partial}{\partial x_k}$. These operators are considered as

$$\mathbf{L}_k : \mathcal{D}(\mathbf{L}_k) \subset L_2(\Omega) \to L_2(\Omega)$$

hence the scalar products used in the following are L_2 scalar products. They fulfill condition (2.6) with

$$t = 1 .$$

We recapitulate the results necessary to apply the results from the last sections. The Radon transform fulfills conditions (2.2) and (2.3) with $\alpha = 1/2$. The Range of $\mathbf{R}$ is described by consistency conditions, known as Helgason-Ludwig-Gelfand conditions, for references see e.g. [25]. The operators $\mathbf{L}_k$ fulfil (2.6) with $t = 1$, which means that the range of $\mathbf{R}_{\mathbf{L}_k}^\dagger$ consists of those functions of $\mathbf{R}^\dagger$ such that $\mathbf{R}^\dagger g \in H^1(\Omega)$, hence $\mathbf{L}_k \mathbf{R}^\dagger g \in L_2(\Omega)$. The whole problem of determining $\mathbf{L}_k f$ form $\mathbf{R} f = g$ thus is ill–posed of order

$$\alpha + t = 3/2 \tag{6.6}$$

We observe that also $\mathbf{L}_k$ intertwines with the shift operators considered above. Hence, Theorem 5.1 tells immediately that if we choose a mollifier of convolution type, then also the reconstruction is of filtered backprojection type as it is the case for the standard reconstruction algorithm for determining f. In order to find the dependency of the reconstruction kernel with respect to the direction we make use of the relation of the Radon transform and the differential operators given in (6.2).

THEOREM 6.1. *Let the mollifier $e_\gamma(x, \cdot)$ be given as*

$$e_\gamma(x, y) = E_\gamma(\|x - y\|) \tag{6.7}$$

Then the reconstruction kernel for finding $\mathbf{L}_k f$, where $\mathbf{L}_k = \frac{\partial}{\partial x_k}$, is $\psi_\gamma(x, \cdot)$ with

$$\psi_{\gamma k}(x; \theta, s) = \theta_k \Psi_\gamma(s - x^\top \theta) \tag{6.8}$$

where θ_k is the k-th component of θ and where $\Psi_\gamma(s)$ is determined as

$$\Psi_\gamma = -\frac{1}{4\pi} \frac{\partial}{\partial s} I^{-1} \mathbf{R} E_\gamma \tag{6.9}$$

If, in addition, $\hat{E}_\gamma$ fulfills the conditions of Theorem 4.3, then $S_\gamma g := \langle g, \psi_{\gamma k} \rangle$ is a regularization of $\mathbf{R}_{\mathbf{L}_k}^\dagger$.

Proof. We start with the auxiliary problem and use the inversion formula for $\mathbf{R}$

$$\begin{aligned}
\mathbf{R}^* \psi_\gamma &= \mathbf{L}_k^* e_\gamma \\
&= \mathbf{R}^{-1} \mathbf{R} \mathbf{L}_k^* e_\gamma \\
&= \frac{1}{4\pi} \mathbf{R}^* I^{-1} \mathbf{R} \mathbf{L}_k^* e_\gamma
\end{aligned}$$

hence we get

$$\psi_\gamma = \frac{1}{4\pi} I^{-1} \mathbf{R} \mathbf{L}_k^* e_\gamma$$

The relation (5.2) between Radon transform and differential operators together with $L_k^* = -L_k$ results in

$$\psi_\gamma = -\frac{1}{4\pi}\theta_k \mathbf{I}^{-1}\frac{\partial}{\partial s}\mathbf{R}e_\gamma$$

Using Fourier transforms we see that I^{-1} and $\frac{\partial}{\partial s}$ commute, hence

$$\psi_\gamma = -\frac{1}{4\pi}\theta_k \frac{\partial}{\partial s}\mathbf{I}^{-1}\mathbf{R}e_\gamma$$

Now e_γ is independent of a direction, so is then $\mathbf{R}e_\gamma$ and also the derivatives. Combining this with the conclusions of Theorem 5.1 the above statement is proven.
□

In the following we present a special mollifier for verifying the theoretical results of the preceding chapters. The cut–off frequency is denoted as b and it is related to the γ used before by

$$b = 1/\gamma \tag{6.10}$$

In a fist step we choose the mollifier for the reconstruction part. Because of its advantageous properties we select the mollifier stemming from the Shepp – Logan kernel,

$$\widehat{e_b^1}(\xi) = (2\pi)^{-1}\mathrm{sinc}\frac{\|\xi\|\pi}{2b}\chi_{[-b,b]}(\|\xi\|) \tag{6.11}$$

For the differentiation part we choose with a possibly different parameter β

$$\widehat{e_\beta^2}(\xi) = (2\pi)^{-1}\mathrm{sinc}\frac{\|\xi\|\pi}{\beta} \tag{6.12}$$

leading to a combined kernel of the form

$$E_{b\beta} = e_b^1 * e_\beta^2 \tag{6.13}$$

with

$$\widehat{E_{b\beta}}(\xi) = (2\pi)^{-1}\mathrm{sinc}\frac{\|\xi\|\pi}{2b}\mathrm{sinc}\frac{\|\xi\|\pi}{\beta}\chi_{[-b,b]}(\|\xi\|) \tag{6.14}$$

THEOREM 6.2. *The mollifier $E_{b\beta}$ given in (6.13) is of convolution type and radially symmetric. It fulfills with $\gamma = 1/b$, see (6.yyyyy), and $\beta = \tau b$*

$$i) 2\pi\sup_{\xi}\{(1+|\xi|^2)^{3/4}|\hat{E}_{b\beta}(\xi)|\} \leq c(\gamma) = \left(1+\gamma^{-2}\right)^{3/4},$$

$$ii) \quad \sup_{\xi\in\mathbb{R}^2}\left(|(2\pi)^{N/2}\hat{E}_{b\beta}(\xi) - 1|\right) \overset{\gamma\to 0}{\longrightarrow} 0$$

Hence, $T_{b\beta}g = \langle g, \psi_{b\beta}$ with $\psi_{b\beta}$ determined according to Theorem 6.1 is a regularization for determining $\mathbf{L}_k$, using all the invariances presented in this section.

Proof. Compared to the example at the end of Section 4 the Fourier transform of the here used mollifier $E_{b\beta}$ has an additional factor sinc $(\pi\|\xi\|/\beta)$ which is also bounded by 1, hence estimate i) follows. Due to the fact, that the parameter β is tied

to b and hence also tends to ∞ for $\gamma \to 0$, also this factor tends to 1, which proves the condition ii). $\square$ In order to compute the reconstruction kernel for determining $L_k f$ we start by computing the Radon transform of $E_{b\beta}$ where we use that the Radon transform of a convolution is the convolution of the Radon transforms

$$\mathbf{R} E_{b\beta} = \mathbf{R} e_b^1 * \mathbf{R} e_\beta^2$$

Next we use the convolution theorem for Fourier transforms and the projection theorem for the Radon transform to get

$$(\mathbf{R} E_{b\beta})\,\hat{}\,(\sigma) = (2\pi)^{1/2}(\mathbf{R} e_b^1)\,\hat{}\,(\sigma)(\mathbf{R} e_\beta^2)\,\hat{}\,(\sigma)$$
$$= (2\pi)^{3/2} \widehat{e_b^1}(\sigma\theta) \widehat{e_\beta^2}(\sigma\theta)$$

where we used that $E_{b\beta}(x)$ only depends on the length of x, hence its Radon transform is independent of θ and in the last step we can use any θ. In the following we write, for the sake of simplicity, $e_c(\sigma)$ instead of $e_c(\sigma\theta)$. Now we apply differentiation and Riesz potential

$$-\frac{1}{4\pi}\left(\frac{\partial}{\partial s}\mathbf{I}^{-1}\mathbf{R} E_{b\beta}\right)\hat{}\,(\sigma) = -\frac{1}{4\pi}(2\pi)^{3/2}\imath\sigma|\sigma|\widehat{e_\beta^2}(\sigma)\widehat{e_b^1}(\sigma)$$
$$= (2\pi)^{1/2}\left(-\imath\sigma(2\pi)^{1/2}\hat{e}_\beta^2(\sigma)\right)\left(\frac{1}{2}(2\pi)^{-1/2}|\sigma|\hat{e}_b^1(\sigma)\right)$$
$$= (2\pi)^{1/2}\left(\hat{\psi}_\beta^2(\sigma)\right)\left(\hat{\psi}_b^1(\sigma)\right)$$
$$= \left(\psi_\beta^2 * \psi_b^1\right)\hat{}\,(\sigma)$$

The kernel ψ_b^1 is with the above choice of e_b^1 the Shepp – Logan kernel, see [25], page 111,

$$\widehat{\psi_b^1}(\sigma) = \frac{1}{8\pi^2}(2\pi)^{1/2}|\sigma|\operatorname{sinc}\frac{\|\xi\|\pi}{2b}\chi_{[-b,b]}(\|\xi\|)$$

with

$$\psi_b^1(s) = \frac{b^2}{2\pi^3}\frac{\frac{\pi}{2} - bs\sin(bs)}{\frac{\pi}{2}^2 - (bs)^2} \tag{6.15}$$

For the kernel ψ_β^2 we observe

$$\hat{\psi}_\beta^2(\sigma) = -\imath\sigma(2\pi)^{-1/2}\operatorname{sinc}\left(\frac{\sigma\pi}{\beta}\right)$$
$$= -\imath\sigma(2\pi)^{-1/2}\frac{\sin(\sigma\pi/\beta)}{\sigma\pi/\beta}$$
$$= -\imath\frac{\beta}{\pi}(2\pi)^{-1/2}\frac{1}{2\imath}\left(\exp(\imath\sigma\pi/\beta) - \exp(-\imath\sigma\pi/\beta)\right)$$
$$= -\frac{\beta}{2\pi}\left(\hat{\delta}_{-\pi/\beta} - \hat{\delta}_{\pi/\beta}\right)(\sigma)$$

where we used in the last step that

$$\hat{\delta}_s(\sigma) = (2\pi)^{-1/2}\exp(-\imath s\sigma)$$

For the final result of $\psi_{b\beta}$ we then get

$$\psi_{b\beta}(s) = \psi_\beta^2 * \psi_b^1(s)$$

$$= \frac{\beta}{2\pi} \left(\psi_b^1(s + \frac{\pi}{\beta}) - \psi_b^1(s - \frac{\pi}{\beta}) \right)$$

Now we choose $b = \beta$ which means $\tau = 1$ in the last theorem, and, as usual, see again [25], we put $b = \pi/h$ and $s_\ell = \ell h$ and get the new filter

$$\psi_{\pi/h}(s_\ell) = \frac{1}{\pi^2 h^3} \frac{8\ell}{\left(3 + 4\ell^2\right)^2 - 64\ell^2} \ , \ \ell \in Z \tag{6.16}$$

To present the algorithm we assume that the data $\mathbf{R}f(\theta, s)$ are given for $s_k = kh$, $k = -q, ..., q$, $h = 1/q$ and $\theta_j = (\cos\varphi_j, \sin\varphi_j)^\top$ with $\varphi_j = \pi(j-1)/p$, $j = 1, ..., p$. We choose

$$\gamma = \frac{\pi}{h} \tag{6.17}$$

leading to the filter ψ_γ from above.

- Step 1: For $j = 1, ..., p$ evaluate the discrete convolutions

$$v_{m,j} = h \sum_{\ell=-q}^{q} \psi_\gamma(s_j - s_\ell)\mathbf{R}f(\theta_m, s_\ell) \ , \ \ j = -q, ..., q \ . \tag{6.18}$$

- Step 2: To get the partial derivatives with respect to x_k, $k = 1, 2$, multiply

$$v_{m,j}^k = \theta_{m,k} v_{m,j} \tag{6.19}$$

- Step 3: For each reconstruction point x compute the discrete backprojection

$$\left(\frac{\partial}{\partial x_k} f \right)_\gamma (x) = \frac{2\pi}{p} \sum_{m=1}^{p} \left((1 - \eta)v_{m\ell}^k + \eta v_{m,\ell+1}^k \right) \tag{6.20}$$

where, for each m and x, ℓ and η are determined by

$$s = \theta^\top x \ , \ \ell \le s/h < \ell + 1 \ , \ \eta = s/h - \ell$$

When comparing to the standard reconstruction, see e.g. [25], we observe, that the filter changes and in addition we have step 2, the multiplication with $\cos\varphi_m$ or $\sin\varphi_m$ respectively. The filter is the same for both derivatives.

Here we optimized the choice of the different filters and the selection of the parameters. In order to test the algorithm we choose the well – known Shepp – Logan phantom, where we use the densities originally given by Shepp – Logan; i.e., the skull has the value 2 and the brain has the value 1 (in contrast to many authors, where these values are lowered by 1 leading to a brain consisting of air, as in the outside of the skull). The objects inside the brain differ by 1% up to 3% to the surrounding tissue. The number of data are $p = 720$ and $q \approx p/\pi$, namely $q = 326$ leading to 653 rays per view. The reconstruction is computed on a 1025×1025 grid.

Figures 1 and 2 show the result of the above mentioned algorithm with exact data. We observe that even the height of the jumps is correctly computed within the numerical approximation of the derivatives.

Then we added to the data 5% noise. Figure 3 shows the density reconstruction and figure 4 the application of the smoothed derivative in x_1 direction. Figures 5 and 6 show the result of the above algorithm, the contours of the object are clearly visible, which is even the case for the objects differing only by 1% relative to the surrounding tissue. This is not the case for the classical approach in Figure 4, and also not the case for the application of Λ - tomography, where the second derivative of the data is computed and backprojected, see Figure 7.

The artefacts outside the object can easily be removed by implementing the support theorem for the Radon transform stating that the object vanishes on lines parallel to θ not meeting the support of the data, see [2].

Fig. 1: Application of the above algorithm for derivative in x_1 - direction with exact data.
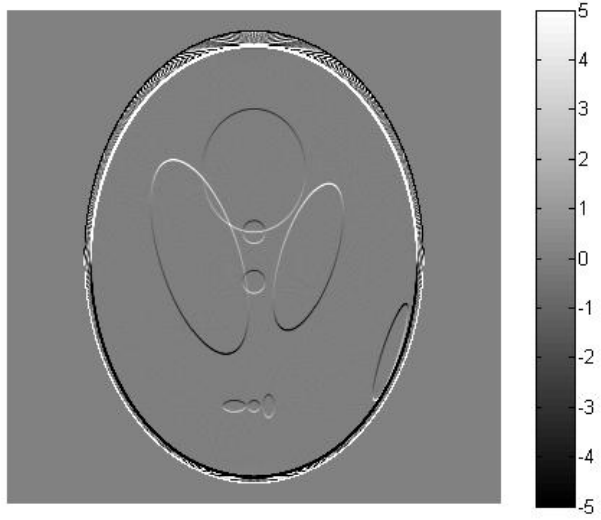
Fig. 2: Application of the above algorithm for derivative in x_2 - direction with exact data.
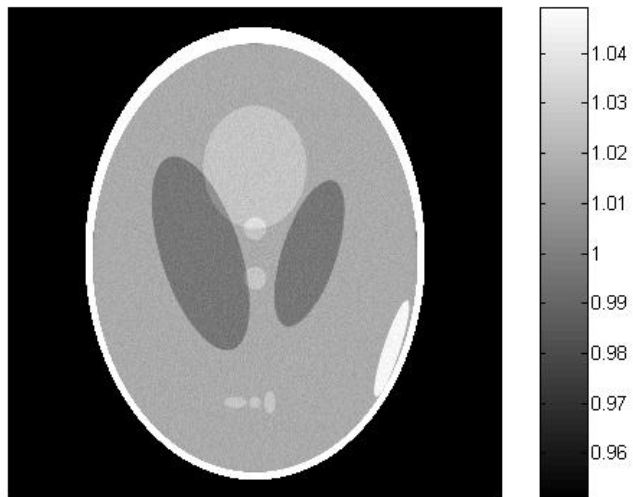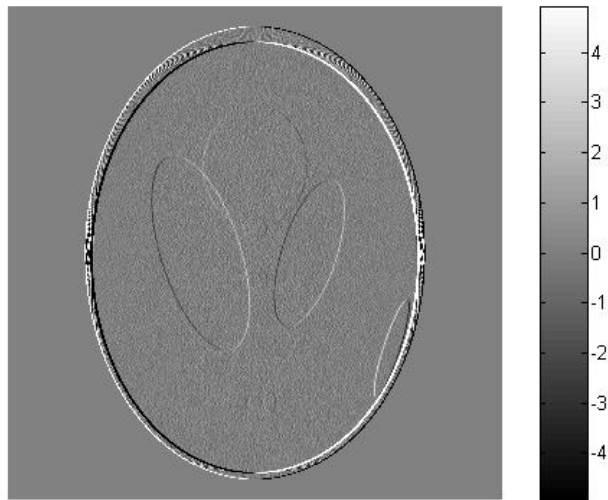

Fig. 3: Classical reconstruction of noisy data.

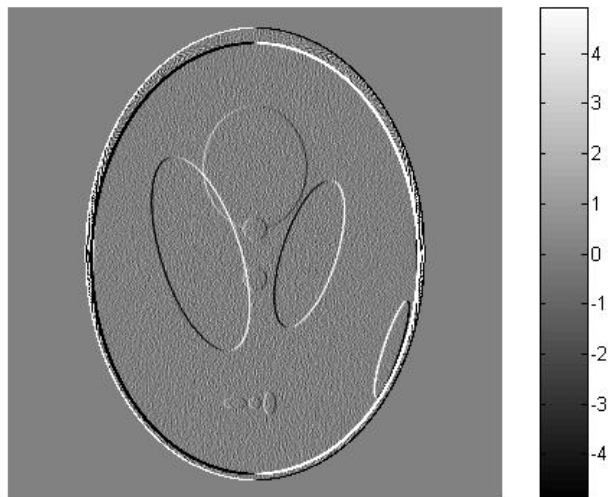Fig. 4: Smoothed derivative in x_1 direction of the image in Figure 3.

Fig. 5: Application of the above algorithm for the direct computation of the derivative in x_1 - direction.

Fig. 6: Application of the above algorithm for the direct computation of the derivative in x_2 - direction.
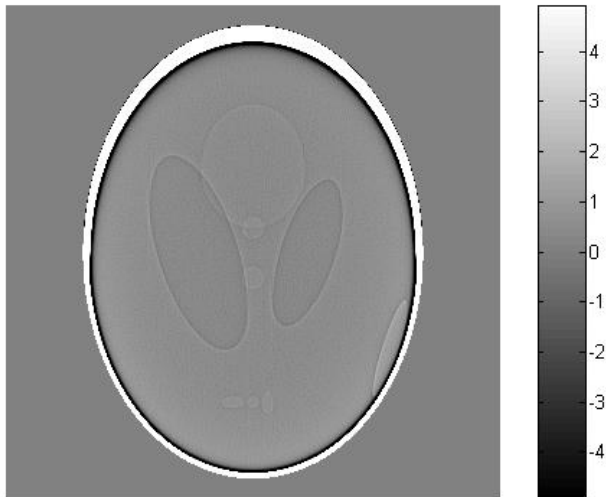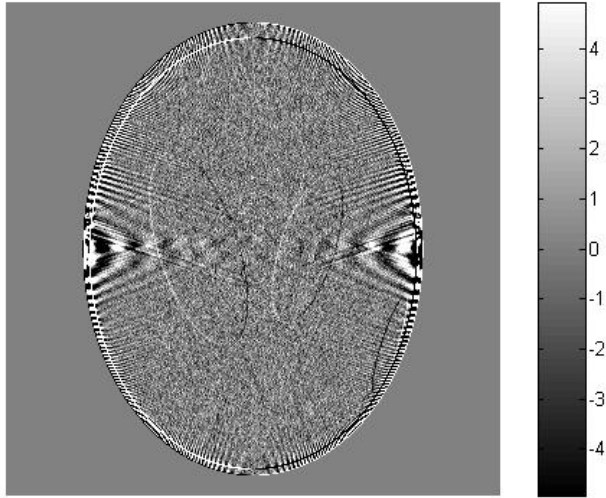

Fig. 7: Λ tomography, [10].

Fig. 8: Reconstruction of the same data set with $\beta = b$ replaced by $\beta = 2b$ in the second filter.

Figure 8 shows that if we don't follow the theoretically motivated strategy of parameter selection as above we get much worse reconstructions. If we choose a smaller γ then we loose resolution. As consequence we note that it pays off to combine the two steps of image reconstruction and image analysis wherever possible.

Finally the Figures 9 and 10 show $|\mathbf{L}_1 f| + |\mathbf{L}_2 f|$ where the colortable is changed such that the highest values are black, and the same window is used for both images. Figure 9 is produced with the here presented method, Figure 10 with first reconstructing f and then the differentiation is performed on the smoothed reconstructions as in Figure 4.
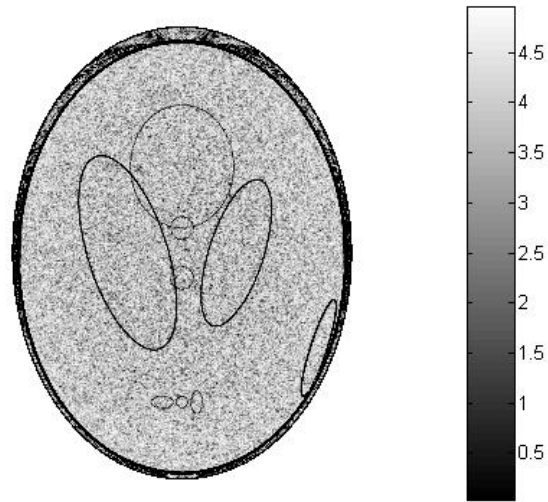
Fig. 9: Display of $|\mathbf{L}_1 f| + |\mathbf{L}_2 f|$ with the here presented method.
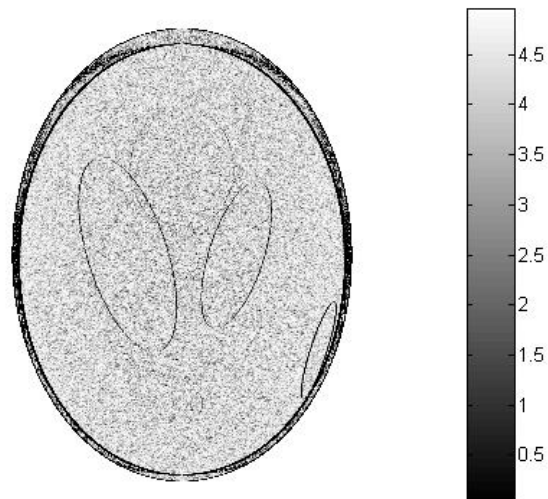

Fig. 10: Display of $|\mathbf{L}_1 f| + |\mathbf{L}_2 f|$ when first the function was reconstructed and then the derivatives are computed from the smoothed image, compare Fig. 4.

What the computation time is concerned this approach differs only by the filter selection from the standard filtered backprojection, so it is as fast as this, and the additional computing time for the differentiation of the reconstruction is not needed. When taking into account that the backprojection step in the calculation is, due to the determination which detector position for a combination of direction and reconstruction point is to be used, even the calculation of three images, namely the density itself and the two derivatives, is almost as fast as the reconstruction of the image itself if it is performed in one program.

REFERENCES

[1] G. Backus and F. Gilbert *The Resolving Power of Growth Earth Data.*, Geophys. J. Roy. Astron. Soc. 16(1968), pp. 169-205

[2] J. Boman and E. T. Quinto *Support theorems for real analytic Radon transforms*, Duke Math. J. 55(1987), pp. 943-948.

[3] S. Bonnet, F. Peyrin, F. Turjman, and R. Prost *Multiresolution Reconstruction in Fan-Beam Tomography* IEEE Transactions on Image Processing, 11 (2002),pp. 169-176.

[4] J. F. Canny *A computational approach to edge detection*, IEEE TPAMI 8 (1986), pp. 679–698.

[5] M. Cetin and W. C. Karl *Feature-enhanced synthetic aperture radar image formation based nonquadratic regularization* IEEE TIP 10 (2001) pp. 623–631

[6] B. Chalmond *Modeling and Inverse Problems in Image Analysis*, Springer, Berlin, 2003

[7] M. E. Davison and F. A. Grünbaum *Tomographic reconstruction with arbitrary directions*, IEEE TNS 26 (1981), pp. 77–120

[8] U. Eckhardt *Zur numerischen Behandlung inkorrekt gestellter Aufgaben*, Computing 17, (1976) pp. 193–206

[9] H. W. Engl, M. Hanke M and A. Neubauer *Regularization of Inverse Problems*, (Kluwer, Dordrecht), 1996

[10] A. Faridani, D. V. Finch, E. L. Ritman, and K. T. Smith *Local Tomography II*, SIAM J. Appl. Math. 57 (1997), pp. 1095–1127.

[11] B. Jähne *Image Processing for Scientific Applications*, CRC Press, Boca Raton, 2nd ed. 2004

[12] P. Jonas and A. K. Louis *A Sobolev space analysis of linear regularization methods for ill–posed problems* J. Inv. Ill-Posed Probl. 9 (2001), pp. 59-74

[13] A. Katsevich *Improved cone beam local tomography*, Inverse Problems, 22 (2006), pp. 627–643

[14] I. Loris, G. Nolet, I. Daubechies and F. A. Dahlen *Tomographic inversion using ℓ_1 - norm regularization of wavelet coefficients* . Geophysical Journal International 170 (2007) pp. 359–370

[15] A. K. Louis *Acceleration of convergence for finite element solutions of the Poisson equation*, Numer. Math. 33 (1979) pp. 43–53

[16] ———*Approximate inverse of the 3D Radon transform*, Math. Meth. Appl. Sci. 5 (1983), pp. 176–185.

[17] ———*Inverse und schlecht gestellte Probleme*, Teubner, Stuttgart, 1989.

[18] ———*Approximate inverse for linear and some nonlinear problems* Inverse Problems, 12 (1996), pp. 175-190

[19] ———*A Unified approach to regularization methods for linear ill–posed problems*, Inverse Problems, 15 (1999), pp. 489–498.

[20] A. K. Louis and P. Maass *A mollifier method for linear operator equations of the first kind*, Inverse Problems, 6 (1990), pp. 427–440

[21] ———*Contour reconstruction in 3-D X-Ray CT*, IEEE Transactions on Medical Imaging, 12 (1993), pp. 764–769

[22] A. K. Louis, P. Maass and A. Rieder *Wavelets*, 2nd ed., Teubner, Stuttgart, 1989 and Wiley, Chichester, 1997 (English translation)

[23] D. A. Murio *The mollification method and the numerical solution of ill-posed problems*, Wiley, Chichester, 1993

[24] F. Natterer *Error bounds for Tikhonov regularization in Hilbert spaces*, Appl. Anal., 18, 29 − 37, 1984

[25] ———*The mathematics of computerized tomography*(Wiley and Teubner), Stuttgart, 1986

[26] S. Oeckl, T. Schön, A. Knauf and A. K. Louis *Multiresolution 3D-computerized tomography and its application to NDT*, Proc. ECNDT, 9 (2006)

[27] J.S. Ovall *Asymptotically exact functional error estimates on superconvergent gradient recovery*, Numer. Math. 102 (2006) pp. 543–558

[28] E. T. Quinto *Radon Transforms, Differential Equations, and Microlocal Analysis*, Contemporary Mathematics, 278(2001), pp. 57-68.

[29] R. Ramlau and W. Ring *A Mumford-Shah level-set approach for the inversion and segmentation of X-ray tomography data* J. Comput. Phys. 221 (2007) pp. 539–557