

Projected Landweber method and preconditioning

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1997 Inverse Problems 13 441

(<http://iopscience.iop.org/0266-5611/13/2/016>)

[The Table of Contents](#) and [more related content](#) is available

Download details:

IP Address: 130.251.61.251

The article was downloaded on 13/10/2009 at 13:40

Please note that [terms and conditions apply](#).

Projected Landweber method and preconditioning

M Piana and M Bertero

INFN and Dipartimento di Fisica dell'Università di Genova, via Dodecaneso 33, I-16146 Genova, Italy

Received 12 July 1996

Abstract. The projected Landweber method is an iterative method for solving constrained least-squares problems when the constraints are expressed in terms of a convex and closed set \mathcal{C} . The convergence properties of the method have been recently investigated. Moreover, it has important applications to many problems of signal processing and image restoration. The practical difficulty is that the convergence is too slow. In this paper we apply to this method the so-called *preconditioning* which is frequently used for increasing the efficiency of the conjugate gradient method. We discuss the significance of preconditioning in this case and we show that it implies a modification of the original constrained least-squares problem. However, when the original problem is ill-posed, the approximate solutions provided by the preconditioned method are similar to those provided by the standard method if the preconditioning is suitably chosen. Moreover, the number of iterations can be reduced by a factor of 10 and even more. A few applications to problems of image restoration are also discussed.

1. Introduction

In the theory of linear inverse problems, the use of *a priori* information is basic for defining approximate solutions which are stable and physically sound. In many instances this information can be exploited by requiring that the solution belongs to a closed and convex set \mathcal{C} .

Assume that the linear inverse problem consists in solving the equation

$$Af = g \quad (1.1)$$

where $A : X \rightarrow Y$ is a linear and continuous operator, X and Y are Hilbert spaces and $g \in Y$ is the datum of the problem, corrupted by noise or experimental errors. In other words g has the following structure:

$$g = Af^{(0)} + w \quad (1.2)$$

where $f^{(0)}$ is the function to be estimated and w is a term describing the noise. If the range of A , $\mathcal{R}(A)$, is not closed, then problem (1.1), in general, has no solution when g has the structure (1.2), because w does not belong to $\mathcal{R}(A)$. Moreover, if $f^{(0)}$ belongs to some closed and convex set \mathcal{C} , the use of linear regularization methods does not provide approximate solutions which belong to \mathcal{C} .

In such a case one can reformulate the original problem (1.1) as a constrained least-squares problem:

$$\|Af - g\|_Y = \text{minimum} \quad f \in \mathcal{C}. \quad (1.3)$$

When $\mathcal{R}(A)$ is not closed this problem can still be ill-posed. The existence of solutions depends on the set \mathcal{C} and we have the following cases. Let $A(\mathcal{C})$ be the image of \mathcal{C} ($A(\mathcal{C})$ is also a convex set).

- Problem (1.3) has a solution for any $g \in Y$ if and only if the set $A(\mathcal{C})$ is closed. This condition is satisfied, for instance, when the set \mathcal{C} is bounded, a case considered in one of the pioneering papers on ill-posed problems [1].

- If $A(\mathcal{C})$ is not closed, then problem (1.3) has a solution if and only if the convex projection of g onto $\overline{A(\mathcal{C})}$ belongs to $A(\mathcal{C})$. If \mathcal{C} is, for instance, the set of non-negative functions in $L^2(\mathbb{R})$, then $A(\mathcal{C})$, in general, is not closed.

It may be interesting to note that the problem of constrained regularized solutions, i.e.

$$\|Af - g\|_Y^2 + \mu \|f\|_X^2 = \text{minimum} \quad f \in \mathcal{C} \quad (1.4)$$

can also be formulated as a constrained least-squares problem. In fact, if $Z = Y \oplus X$ and $A_\mu : X \rightarrow Z$ is defined by $A_\mu f = \{Af, \sqrt{\mu}f\}$, then (1.4) can be rewritten as follows:

$$\|A_\mu f - g\|_Z = \text{minimum} \quad f \in \mathcal{C} \quad (1.5)$$

where $g = \{g, 0\} \in Z$. This problem, which has been investigated by Neubauer [2], has interesting applications in image restoration, as proved by Lagendijk *et al* [3], if one wishes to reduce the ringing effects, typical of linear methods, by the use of the positivity constraint. Since the operator A_μ has a bounded inverse, problem (1.5) is well-posed.

The projected Landweber method is an iterative method which can be used for approximating the solutions of the above-mentioned problems. Its convergence and regularization properties have been investigated by Eicke [4]. It can also be easily implemented if the projection operator over the convex and closed set \mathcal{C} , $P_{\mathcal{C}}$, can be easily computed. A list of computable convex projections which occur frequently in many signal-processing applications and image restoration problems is provided by Youla and Webb [5].

It follows from many numerical simulations and practical applications (we will also give examples in this paper) that this method can provide much better estimates than the usual linear regularization methods. The practical difficulty is that the convergence is too slow, i.e. too many iterations are required to obtain the best approximation. For this reason we investigate the possibility of applying to the projected Landweber method the so-called *preconditioning* which is frequently used for increasing the efficiency of the conjugate gradient method [6]. We will find that, in fact, the use of preconditioning increases the convergence speed even if it implies a modification of the original least-squares problem.

The plan of the paper is as follows. In section 2 we summarize the main results and open problems concerning the projected Landweber method. In section 3 we briefly describe a method introduced by Strand [7] for increasing the speed of convergence of the Landweber method, we show that this is one of the various possible forms of preconditioning and we prove that this can also be considered as the application of the Landweber method to a modified least-squares problem. However, in the absence of constraints, this modified problem is equivalent to the original one. This result is no longer true in the case of constraints, so that the projected Landweber method when applied to the modified least-squares problem can provide different results. This point is discussed in section 4 where we also give two rather natural forms of preconditioning. We show by means of numerical examples that, in such a way, one gets results similar to those provided by the standard projected Landweber method but with a much smaller number of iterations. Finally, in section 5 we discuss a few applications to problems in image restoration.

2. The projected Landweber method: state of the art and open problems

Let us consider first the non-constrained least-squares problem associated with the first-kind equation (1.1)

$$\|Af - g\|_Y = \text{minimum.} \tag{2.1}$$

Let us denote by $\mathcal{N}(A)$ the null space of A . If $\mathcal{R}(A)$ is not closed, then the problem (2.1) is ill-posed. Least-squares solutions exist if and only if $g \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$. In such a case the set of least-squares solutions coincides with the set of the solutions of the Euler equation

$$A^*Af = A^*g \tag{2.2}$$

where A^* denotes the adjoint operator of A . The minimal norm solution is the so-called *generalized solution* f^\dagger which is also the unique least-squares solution belonging to $\mathcal{N}(A)^\perp$. The *generalized inverse* A^\dagger is the linear operator with domain $\mathcal{D}(A^\dagger) = \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$ and range $\mathcal{R}(A^\dagger) = \mathcal{N}(A)^\perp$, defined by [8]

$$f^\dagger = A^\dagger g. \tag{2.3}$$

If we introduce the nonlinear operator $G : X \rightarrow X$

$$G(f) = f + \tau(A^*g - A^*Af) \tag{2.4}$$

then it is obvious that the set of solutions of problem (2.1) coincides with the set of the fixed points of the operator G . Moreover, if the relaxation parameter τ satisfies the conditions

$$0 < \tau < \frac{2}{\|A\|^2} \tag{2.5}$$

then the operator G is non-expansive, that is

$$\|G(f_1) - G(f_2)\|_X \leq \|f_1 - f_2\|_X. \tag{2.6}$$

The method of Landweber is precisely the method of successive approximations for determining the fixed points of the operator G , that is

$$f_{n+1} = G(f_n). \tag{2.7}$$

From general theorems on the fixed points of non-expansive operators [9] it is only possible to prove that, when the set of the fixed points of G is not empty, then the sequence $\{f_n\}_{n=1}^\infty$ weakly converges to a fixed point of G .

By means of spectral methods it is possible to prove the following stronger result [10].

Theorem 2.1. *If $g \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$ and τ satisfies conditions (2.5) then $\{f_n\}_{n=1}^\infty$ is strongly convergent for any initial approximation f_0 and*

$$\lim_{n \rightarrow \infty} f_n = f^\dagger + P_N f_0 \tag{2.8}$$

where P_N is the orthogonal projection onto $\mathcal{N}(A)$. In particular, the limit is f^\dagger when $f_0 = 0$.

In the case of noisy data the method, in general, does not converge because g does not satisfy the condition required by the theorem. In such a case, however, the method has a property of semiconvergence. Let $\|f_n - f^{(0)}\|_X$ be the restoration error, where $f^{(0)}$ is the unknown function defined in equation (1.2). Then, as n increases, the restoration error is decreasing first and increasing afterwards.

The projected Landweber method is a modification of the Landweber method which can be used for solving the constrained least-squares problem (1.3). We introduce a few notations. We denote by \mathcal{I} the closure of the set $A(\mathcal{C})$, $\mathcal{I} = \overline{A(\mathcal{C})}$. This is also a convex set. Then we denote by $P_{\mathcal{I}}$ the convex projection onto \mathcal{I} . As already stated in the introduction, problem (1.3) has a solution if and only if $P_{\mathcal{I}}g \in A(\mathcal{C})$. In such a case the set of the constrained least-squares solutions is closed and convex. Therefore, there exists a constrained least-squares solution of minimal norm, denoted by $f_{\mathcal{C}}^{\dagger}$. It is possible to prove [4, 11] that the set of the constrained least-squares solutions coincides with the set of the fixed points of the nonlinear operator

$$P_{\mathcal{C}}G(f) = P_{\mathcal{C}}[f + \tau(A^*g - A^*Af)] \quad (2.9)$$

where $P_{\mathcal{C}}$ is the convex projection onto the set \mathcal{C} . In general, $P_{\mathcal{C}}$ is nonlinear. Since $P_{\mathcal{C}}$ is non-expansive, the operator $P_{\mathcal{C}}G$ is also non-expansive if the relaxation parameter τ satisfies conditions (2.5). Then the projected Landweber method is just the method of successive approximations for determining the fixed points of the operator $P_{\mathcal{C}}G$:

$$\bar{f}_{n+1} = P_{\mathcal{C}}G(\bar{f}_n). \quad (2.10)$$

Using general results on the approximation of fixed points of non-expansive operators, as well as peculiar properties of the operator G , the following result has been proved [4].

Theorem 2.2. *If $P_{\mathcal{I}}g \in A(\mathcal{C})$, then the sequence $\{\bar{f}_n\}_{n=1}^{\infty}$ weakly converges to a fixed point of $P_{\mathcal{C}}G$ and therefore to a solution of problem (1.3).*

It is reasonable to conjecture that this result is not the best possible. In fact, assume, for simplicity, that the operator A is injective and that $g = Af^{(0)}$ with $f^{(0)} \in \mathcal{C}$. Then both the sequence $\{f_n\}_{n=1}^{\infty}$ defined by equation (2.7) and the sequence $\{\bar{f}_n\}_{n=1}^{\infty}$ defined by equation (2.10) converge to $f^{(0)}$ (one can take, for instance, $f_0 = 0$). But while one proves that the sequence f_n is strongly convergent, one is only able to prove that the sequence $\{\bar{f}_n\}_{n=1}^{\infty}$ is weakly convergent. This is against the evidence provided by numerical simulations. The restoration error $\|\bar{f}_n - f^{(0)}\|$ is always smaller (in many cases much smaller) than the restoration error $\|f_n - f^{(0)}\|$. See, for example, figure 1(a), where we have plotted the root-mean-squared (rms) errors for the restoration of a Gaussian function with $\sigma = 0.1$ when A is the convolution operator with the characteristic function of the interval $[-1, 1]$. In this case we use data not corrupted by noise.

Moreover, in the case of noisy data, the property of semiconvergence (the restoration error is decreasing first and increasing afterwards) is proved for the sequence $\{f_n\}_{n=1}^{\infty}$ but not for the sequence $\{\bar{f}_n\}_{n=1}^{\infty}$. As expected, the sequence $\{\bar{f}_n\}_{n=1}^{\infty}$ also has this property in the case of simulated data. This is shown in figure 1(b) in the case of the same reconstruction of figure 1(a), but when the data are affected by 5% Gaussian noise.

Another question concerns the dependence of the limit on the initial approximation \bar{f}_0 . As far as we know, it has not been proved that, when $\bar{f}_0 = 0$, the limit of \bar{f}_n is precisely the constrained least-squares solution of minimal norm $f_{\mathcal{C}}^{\dagger}$.

Finally, a few words about the case of the constrained regularized solution, problem (1.5). The operator A_{μ} is such that $A_{\mu}^*A_{\mu}$ has a bounded inverse so that the spectrum of

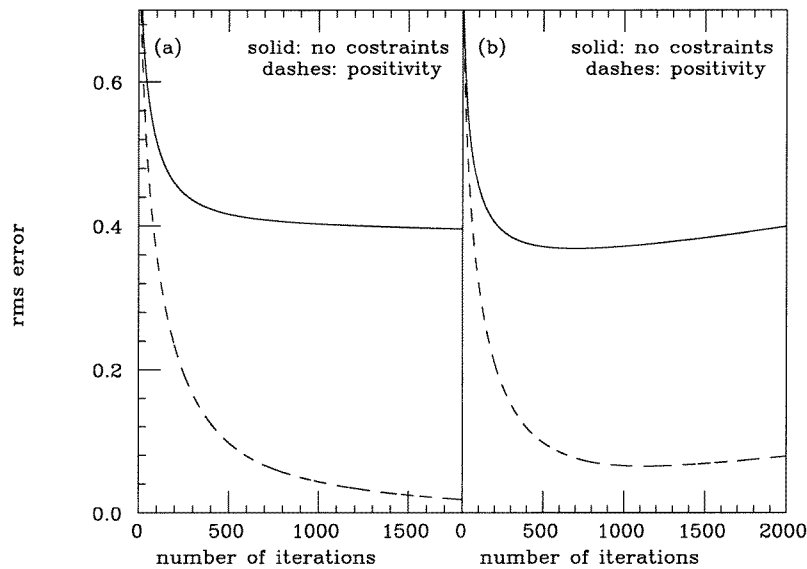


Figure 1. Restoration of a Gaussian function ($\sigma = 0.1$) convolved with the characteristic function of the interval $[-1, 1]$. (a) Behaviour of the restoration error as a function of the number of iterations in the case of exact data; the full curve corresponds to the non-constrained method, the dashed curve to the projected method with positivity constraint. (b) Behaviour of the restoration error in case of data corrupted by 5% Gaussian noise; full and dashed curves have the same meaning as in case (a). Both curves show a minimum, even if this minimum is rather flat.

$A_\mu^* A_\mu$ is interior to a bounded and positive interval, $\sigma(A_\mu^* A_\mu) \subset [\|A_\mu\|^{-2}, \|A_\mu\|^2]$. Then the operator G_μ , defined as the operator G with A replaced by A_μ both in equation (2.4) and (2.5), is a contraction

$$\|G_\mu(f_1) - G_\mu(f_2)\|_X \leq \rho_\mu \|f_1 - f_2\|_X \tag{2.11}$$

with

$$\rho_\mu = \sup_{\lambda \in \sigma(A_\mu^* A_\mu)} |1 - \tau \lambda| < 1. \tag{2.12}$$

Therefore, the operator $P_C G_\mu$ is also a contraction and the projected Landweber method, applied to problem (1.5), converges for any g , to the unique constrained regularized solution.

3. Strand method and preconditioning

The main problem concerning the applications of the projected Landweber method is that, in general, it needs a large number of iterations to provide a satisfactory approximation of the unknown function $f^{(0)}$, equation (1.2) (see, for instance, figure 1(b), where the minimum is reached after 957 iterations).

As this feature is characteristic also of the non-projected algorithm, first we consider the acceleration of the Landweber iteration (2.7), since this case provides insight into the more important problem of the acceleration of the projected Landweber method.

Let us note first that the n th iterate can be written as a filtered version of the generalized solution. If we denote by $\{E_\lambda\}$ the spectral family of the operator $A^* A$, whose spectrum is

interior to the interval $[0, \|A\|^2]$, then, in the case $f_0 = 0$, we have [10]

$$f_n = \int_0^{\|A\|^2} \{1 - (1 - \tau\lambda)^n\} \frac{1}{\lambda} dE_\lambda A^* g. \quad (3.1)$$

This expression clearly shows why the Landweber method requires a large number of iterations. If we assume, for simplicity, that $\|A\| = 1$ and $\tau = 1$, then we see that, in order to recover a spectral component corresponding to $\lambda = 10^{-3}$ with 10% error, we need a number of iterations given by $n = -1/\log_{10}(0.999) \simeq 2300$.

A method for accelerating the Landweber iterations was proposed by Strand [7] in the case of compact operators. His analysis is based on the singular value decomposition of the operator A , but it can be easily extended to the case of an operator A^*A with a more general spectrum.

The basic idea consists in modifying the Euler equation (2.2) as follows. Let $D : X \rightarrow X$ be a linear operator satisfying the following conditions:

- (i) D is continuous;
- (ii) D is positive definite and bounded from below so that it has a continuous inverse;
- (iii) D commutes with the operator A^*A . Then equation (2.2) is replaced by the following,

$$DA^*Af = DA^*g \quad (3.2)$$

and the corresponding Landweber method is given by

$$f_{n+1} = f_n + \tau D(A^*g - A^*Af_n) \quad (3.3)$$

with τ satisfying the following conditions:

$$0 < \tau < \frac{2}{\|DA^*A\|}. \quad (3.4)$$

A few remarks about conditions (i)–(iii). Condition (i) is obvious if one wishes to define the product DA^*A without restrictions. Condition (ii), combined with condition (iii), implies that the operator DA^*A is still positive definite while the existence of D^{-1} implies that the set of the solutions of equation (2.2) coincides with the set of the solutions of equation (3.2). Finally, as we will see, condition (iii) allows a simple discussion of the convergence and regularization properties of the algorithm.

In fact, condition (iii) can be expressed by saying that there exists a positive valued function $F(\lambda)$ defined on the interval $[0, \|A\|^2]$ (with a positive lower bound in order to satisfy condition (ii)) such that

$$D = F(A^*A). \quad (3.5)$$

Strand assumes that $F(\lambda)$ is a polynomial or a rational function. For future use we notice that, in these cases, one can easily prove that the following relationship holds true:

$$F(A^*A)A^* = A^*F(AA^*). \quad (3.6)$$

By extending the analysis which applies to the usual Landweber method one can prove the following.

Proposition 3.1. *If the operator D is given by equation (3.5) then, in the case $f_0 = 0$, the n th iterate f_n as given by equation (3.3) has the following representation*

$$f_n = \int_0^{\|A\|^2} \{1 - [1 - \tau\lambda F(\lambda)]^n\} \frac{1}{\lambda} dE_\lambda A^* g \quad (3.7)$$

where τ satisfies conditions (3.4) and $\{E_\lambda\}$ is the spectral family of the operator A^*A .

From this proposition it is straightforward to prove that the Bialy theorem, i.e. theorem 2.1, holds true for the sequence $\{f_n\}$ defined by equation (3.3) and that this sequence has the semiconvergence property in the case of noisy data. Moreover, if we also require that the new algorithm has to recover first the largest spectral components, then equation (3.7) makes clear that a further assumption on the function $F(\lambda)$ is convenient:

(iv) the function $\phi(\lambda)$ defined by

$$\phi(\lambda) = \lambda F(\lambda) \tag{3.8}$$

is non-decreasing on the interval $[0, \|A\|^2]$.

The effect of the choice of the function $F(\lambda)$ on the convergence of the method is now clear: if $F(\lambda)$ is such that for $\lambda \geq \lambda_0$ the function $\tau\phi(\lambda)$ is approximately equal to 1, then all the spectral components with $\lambda \geq \lambda_0$ are rapidly recovered.

Finally, a remark which will be useful in the next section. Let $\tilde{D} : Y \rightarrow Y$ be the operator defined by

$$\tilde{D} = F(AA^*) \tag{3.9}$$

then, in the data Hilbert space Y , we can introduce the scalar product

$$(g_1, g_2)_{\tilde{Y}} = (\tilde{D}g_1, g_2)_Y. \tag{3.10}$$

The new scalar product is equivalent to the original one and we will denote by \tilde{Y} the Hilbert space Y when equipped with this scalar product.

If we denote by \tilde{A}^* the adjoint operator of A in \tilde{Y} , then from equation (3.6) one obtains the following relationship between \tilde{A}^* and A^* :

$$\tilde{A}^* = A^*\tilde{D} = DA^*. \tag{3.11}$$

From this relationship it is easy to derive that

Proposition 3.2. *Equation (3.2) is the Euler equation associated with the least-squares problem*

$$\|Af - g\|_{\tilde{Y}}^2 = (\tilde{D}(Af - g), Af - g)_Y = \text{minimum} \tag{3.12}$$

and equation (3.3) is the Landweber method for the solution of this problem.

It is obvious that problem (3.12) and problem (2.1) are equivalent in the sense that they have exactly the same solutions. This equivalence, however, is no longer true in the case of constrained problems.

The modification of equation (2.1) as given by equation (3.2) is a form of preconditioning. Preconditioning is a method which has been introduced for increasing the efficiency of the conjugate gradient method but, of course, it can be applied also to other iterative methods. More precisely, that expressed by equation (3.2) can be called a left preconditioning, while it is possible to introduce a right preconditioning by means of the change of variables

$$f = Du \tag{3.13}$$

so that equation (2.1) becomes

$$A^*ADu = A^*g. \tag{3.14}$$

If the operator D satisfies the conditions (i)–(iii) stated above, then the Landweber method applied to this equation coincides with that defined in equation (3.3). In fact, from the equation

$$u_{n+1} = u_n + \tau(A^*g - A^*ADu_n) \tag{3.15}$$

with τ satisfying conditions (3.4), if one defines f_n as $f_n = Du_n$, then the f_n satisfy equation (3.3).

Since the operator $D^{\frac{1}{2}}$ exists, thanks to condition (ii), it is also possible to introduce the usual form of preconditioning [6] as follows

$$D^{\frac{1}{2}}A^*AD^{\frac{1}{2}}v = D^{\frac{1}{2}}A^*g \quad (3.16)$$

with $f = D^{\frac{1}{2}}v$. The Landweber method applied to this equation provides again the iterative scheme (3.3).

4. Acceleration of the projected Landweber method

We consider now the projection of the Landweber method (3.3). We obtain

$$\bar{f}_{n+1} = P_{\mathcal{C}}(\bar{f}_n + \tau D(A^*g - A^*\bar{f}_n)) \quad (4.1)$$

with τ satisfying conditions (3.4). Since equation (3.2) is the Euler equation associated with the least-squares problem (3.12), it is obvious that equation (4.1) defines the projected Landweber method for this problem.

In order to establish the convergence properties of this sequence let us observe that, since the scalar product of Y is equivalent to the scalar product of \tilde{Y} , then the closure of $A(\mathcal{C})$ in Y coincides with the closure of $A(\mathcal{C})$ in \tilde{Y} . We still denote by \mathcal{I} this closed and convex set and we denote by $\tilde{P}_{\mathcal{I}}$ the convex projection onto \mathcal{I} in \tilde{Y} . Then from proposition 3.2 and theorem 2.2 it immediately follows that

Theorem 4.1. *If $\tilde{P}_{\mathcal{I}}g \in A(\mathcal{C})$, then the sequence $\{\bar{f}_n\}_{n=1}^{\infty}$ defined by equation (4.1) weakly converges (both in \tilde{Y} and in Y) to a solution of the constrained least-squares problem*

$$(\tilde{D}(Af - g), Af - g)_Y = \text{minimum} \quad f \in \mathcal{C}. \quad (4.2)$$

A few remarks about the limit of this sequence. If the operator A is injective and if $g = Af^{(0)}$, with $f^{(0)} \in \mathcal{C}$, then both problem (1.3) and problem (4.2) have the unique solution $f^{(0)}$ and therefore the sequences (2.10) and (4.1) have the same limit $f^{(0)}$, independently of the initial approximation f_0 .

The limits can be different, however, if the conditions stated above are not satisfied. In fact, the limit of the sequence (2.10) is a solution of the equation

$$Af = P_{\mathcal{I}}g \quad (4.3)$$

while the limit of the sequence (4.1) is a solution of the equation

$$Af = \tilde{P}_{\mathcal{I}}g. \quad (4.4)$$

What is the relationship between the operators $P_{\mathcal{I}}$ and $\tilde{P}_{\mathcal{I}}$? The following result gives some indications in this direction.

Proposition 4.2. *If $\tilde{\mathcal{I}} = \tilde{D}^{\frac{1}{2}}\mathcal{I}$ and $P_{\tilde{\mathcal{I}}}$ is the convex projection onto $\tilde{\mathcal{I}}$ in Y , then $\tilde{P}_{\mathcal{I}}$ is given by*

$$\tilde{P}_{\mathcal{I}} = \tilde{D}^{-\frac{1}{2}}P_{\tilde{\mathcal{I}}}\tilde{D}^{\frac{1}{2}}. \quad (4.5)$$

Proof. If $g \in Y$ and $\|\cdot\|_{\tilde{Y}}$ is the norm induced by the scalar product (3.10), then

$$\|\tilde{D}^{-\frac{1}{2}}P_{\tilde{\mathcal{I}}}\tilde{D}^{\frac{1}{2}}g - g\|_{\tilde{Y}} = \|\tilde{D}^{\frac{1}{2}}(\tilde{D}^{-\frac{1}{2}}P_{\tilde{\mathcal{I}}}\tilde{D}^{\frac{1}{2}} - g)\|_Y = \|P_{\tilde{\mathcal{I}}}\tilde{D}^{\frac{1}{2}}g - \tilde{D}^{\frac{1}{2}}g\|_Y. \tag{4.6}$$

But, by definition of $P_{\tilde{\mathcal{I}}}$,

$$\|P_{\tilde{\mathcal{I}}}\tilde{D}^{\frac{1}{2}}g - \tilde{D}^{\frac{1}{2}}g\|_Y \leq \|\tilde{D}^{\frac{1}{2}}g' - \tilde{D}^{\frac{1}{2}}g\|_Y \quad \forall g' \in \mathcal{I} \tag{4.7}$$

and then

$$\|\tilde{D}^{-\frac{1}{2}}P_{\tilde{\mathcal{I}}}\tilde{D}^{\frac{1}{2}}g - g\|_{\tilde{Y}} \leq \|g' - g\|_{\tilde{Y}} \quad \forall g' \in \mathcal{I}. \tag{4.8}$$

□

This result does not provide a direct relationship between $\tilde{P}_{\mathcal{I}}$ and $P_{\mathcal{I}}$ but between $\tilde{P}_{\mathcal{I}}$ and the convex projection onto $\tilde{D}^{\frac{1}{2}}\mathcal{I}$ in Y . However, by means of this result, it is not difficult to provide examples in \mathbb{R}^2 showing that the two projections, in general, do not coincide.

We point out the implications of this fact. When problem (1.3) is well-posed, in the sense that $A(\mathcal{C})$ is closed, then problem (4.2) is also well-posed but it is a different problem: its solutions do not coincide with the solutions of problem (1.3) except when $g \in A(\mathcal{C})$. Therefore, in general, the iterative method (4.1) cannot be used for approximating the solutions of problem (1.3). This remark applies, for instance, to the case where the set \mathcal{C} is bounded or to the case of the constrained regularized solutions, equation (1.4).

The situation may be different when problem (1.3) is ill-posed, in the sense that $A(\mathcal{C})$ is not closed. In that case, also problem (4.2) is ill-posed. If both iterative methods have the semiconvergence property and if the data g are given by equation (1.2) then we expect that the two methods give (different) approximations of $f^{(0)}$ but that the method (4.1) is more rapidly convergent if the operator D is properly chosen.

Unfortunately, as already stated in section 2, estimates of the restoration error are not yet available for the projected Landweber method, so that it is not possible to support our previous statement by means of theoretical arguments. However, we have some experimental evidence in this direction even if it seems that the cost to be paid for accelerating the algorithm is a small loss of accuracy. In order to discuss this point we provide a few very simple examples of preconditioning.

As already observed in the previous section, a suitable choice of the preconditioner D is a continuous function of A^*A , $F(A^*A)$, such that the function $\phi(\lambda) = \lambda F(\lambda)$ in some way mimicks the behaviour of the Heaviside function. Examples of operators D having this property are already provided by Strand [7]. In this paper, seventh-order polynomials obtained by a mean square approximation to the Heaviside step function are used. However, a very simple example, closely related to the Tikhonov regularization method is suggested by Sanz and Huang [12]. In fact the operator D is given by

$$D = (A^*A + \alpha I)^{-1} \tag{4.9}$$

where α is a positive real parameter. In this case the function $\phi(\lambda)$, equation (3.8), is given by

$$\phi(\lambda) = \frac{\lambda}{\lambda + \alpha} \tag{4.10}$$

and, as this is a monotonically increasing function, conditions (3.4) on the relaxation parameter become

$$0 < \tau < \frac{2(\|A\|^2 + \alpha)}{\|A\|^2}. \tag{4.11}$$

We point out that the choice $\tau = 1$ is compatible with these conditions.

The convergence speed is directly ruled by the value of the Tikhonov parameter α , as the first derivative of $\phi(\lambda)$ in zero is given by

$$\phi'(0) = \frac{1}{\alpha}. \quad (4.12)$$

Nevertheless, in presence of real data, the choice of a too small value for α is not recommended. In fact, in the case $\bar{f}_0 = 0$ and $\tau = 1$, the first iterate, as given by equation (4.1) is

$$\bar{f}_1 = P_C[(A^*A + \alpha I)^{-1}A^*g] \quad (4.13)$$

and therefore it is simply the projection onto \mathcal{C} of the Tikhonov regularized solution associated with α . If α is too small, this solution is completely corrupted by the errors affecting the data g and it cannot be a good starting point for the iteration procedure. This remark is confirmed by numerical simulations. However, good choices of α may be much smaller than the values provided by the methods used for estimating optimum values of the regularization parameter.

Another simple way of accelerating the projected Landweber method consists in projecting onto \mathcal{C} not at each iteration but, for instance, every N iterations. In fact this procedure can also be described in terms of an operator D which, as we will see, is a polynomial of degree $N - 1$.

If \bar{f}_n is the n th iterate, let us investigate the effect on \bar{f}_n of N unconstrained iterations which we assume to perform before projecting again onto \mathcal{C} . From equation (2.4), if τ_0 is a value of the relaxation parameter satisfying conditions (2.5), we obtain

$$G^N(\bar{f}_n) = \tau_0[I + (I - \tau_0A^*A) + \cdots + (I - \tau_0A^*A)^{N-1}]A^*g + (I - \tau_0A^*A)^N\bar{f}_n. \quad (4.14)$$

If we put

$$D = \tau_0[I + (I - \tau_0A^*A) + \cdots + (I - \tau_0A^*A)^{N-1}] \quad (4.15)$$

and if we observe that, by means of simple algebra, one gets

$$I - DA^*A = (I - \tau_0A^*A)^N \quad (4.16)$$

then we can write equation (4.14) as follows

$$G^N(\bar{f}_n) = \bar{f}_n + D(A^*g - A^*A\bar{f}_n). \quad (4.17)$$

Therefore, if \bar{f}_{n+1} is obtained by applying P_C to $G^N(\bar{f}_n)$

$$\bar{f}_{n+1} = P_C G^N(\bar{f}_n) \quad (4.18)$$

we see that this iterative scheme is equivalent to (4.1) with D given by equation (4.15) and $\tau = 1$.

In order to prove that the choice $\tau = 1$ is permitted, we observe that the function $\phi(\lambda)$, defined in equation (3.8), is now given by

$$\phi(\lambda) = 1 - (1 - \tau_0\lambda)^N \quad (4.19)$$

as follows from equation (4.16). Since this function is increasing on the interval $[0, \|A\|^2]$ when $\tau_0 \leq 1/\|A\|^2$, while it has a maximum at $\lambda_0 = 1/\tau_0$ when $\|A\|^{-2} < \tau_0 < 2\|A\|^{-2}$ (and the maximum value is one), from equation (3.4) one finds the following conditions for the relaxation parameter τ of equation (4.1)

$$0 < \tau < \frac{2}{\psi(\tau_0\|A\|^2)} \quad (4.20)$$

where

$$\psi(\tau_0 \|A\|^2) = \begin{cases} 1 - (1 - \tau_0 \|A\|^2)^N & \tau_0 \leq 1/\|A\|^2 \\ 1 & \tau_0 > 1/\|A\|^2. \end{cases} \quad (4.21)$$

Since the right-hand side of equation (4.20) is always greater than 1, the value $\tau = 1$ is compatible with these inequalities.

The convergence rate is ruled now by the value of N , because from equation (4.19) one also has

$$\phi'(0) = \tau_0 N. \quad (4.22)$$

Also in this case it is not convenient to choose too large a value of N in the presence of noisy data, even if a good choice could be a value of N much larger than that obtained from the stopping rules of the usual Landweber method [13].

We conclude this section by showing the result of a numerical experiment. The problem is the same as that considered in section 2. We generate data by blurring a Gaussian function ($\sigma = 0.1$) with the characteristic function of the interval $[-1, 1]$ and by corrupting the result with multiplicative Gaussian noise. The constraint considered is that of positivity and we have investigated the acceleration of the method determined by Tikhonov preconditioning (4.9).

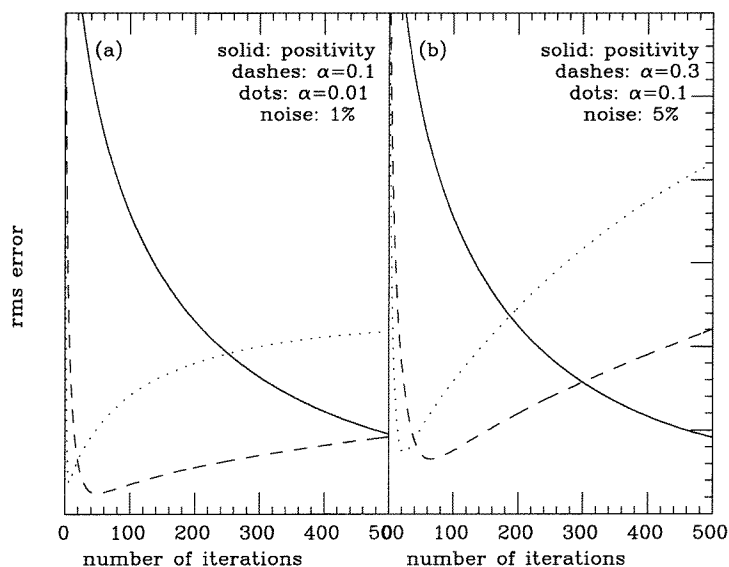


Figure 2. Restoration of a Gaussian function ($\sigma = 0.1$) convolved with the characteristic function of the interval $[-1, 1]$. (a) Restoration error in the case of 1% noise: the full curve corresponds to the method without preconditioning, the dashed and dotted curves to the method with Tikhonov preconditioning with $\alpha = 0.1$ and $\alpha = 0.01$, respectively. (b) The same as in (a) but in the case of 5% noise. The values of α are now 0.3 and 0.1, respectively.

In figure 2(a) we give the results in the case of 1% noise. The method without preconditioning has a minimum for $n = 2072$ and the minimum restoration error is 2.5%. If we use the Tikhonov preconditioning with $\alpha = 0.1$, the minimum is reached after 49 iterations and the minimum error is again 2.5%. When $\alpha = 0.01$ the minimum is reached after seven iterations but the minimum value is now 3.8%. Anyway, this is still an acceptable value.

In figure 2(b) we give the results obtained in the case of 5% noise. In the absence of preconditioning the minimum is now at $n = 957$ and the minimum error is 5.9%. In the case of preconditioning with $\alpha = 0.3$ the minimum occurs at $n = 67$ and the minimum error is 6.6% while in the case of $\alpha = 0.1$ the minimum occurs at $n = 24$ and the minimum error is 7.4%.

These results indicate that Tikhonov preconditioning allows a considerable acceleration of the projected Landweber method at the cost of a moderate increase of the minimum restoration error. We also note that the acceleration depends on the amount of noise affecting the data; in particular, the optimum value of n decreases when the noise increases.

5. Applications to image restoration

In this section we discuss a few applications of the projected Landweber method to problems of image restoration. We illustrate our remarks with the results of some numerical simulations.

5.1. Generalization of the Gerchberg algorithm and super-resolution

The so-called Gerchberg method [14] is a very simple iterative method, based on the Fourier transform, for the extrapolation of band-limited signals which are known on a bounded time interval. The same method can be applied to the problem of out-of-band extrapolation of space-limited images. It can be shown that this method is equivalent to the application of the Landweber method to the inversion of the appropriate Slepian operator [15]. We will show that it can also be viewed as a projected Landweber method and we generalize it to the case of convolution equations. Since in all these cases we have linear projections, the Bialy theorem applies to these iterative methods.

We first summarize the basic facts about the problem of out-of-band extrapolation. Let \mathbb{D} be a bounded, measurable (with non-zero measure) domain of \mathbb{R}^2 and let f be a square-integrable function whose support is contained in \mathbb{D} . Such a function is also called *space-limited*. The Fourier transform of f , \hat{f} , is an entire function of exponential type. Then the problem of out-of-band extrapolation can be formulated as follows. Determine a function f , with support in \mathbb{D} , being given its Fourier transform on a bounded, measurable domain \mathbb{B} (again with non-zero measure). The solution of this problem is unique thanks to the uniqueness of analytic continuation. The problem however is ill-posed as concerns existence and continuous dependence of the solution on the data. In fact, if \hat{g} is the data function, defined on \mathbb{B} , the problem of determining f such that $\hat{f}(\omega) = \hat{g}(\omega)$ when $\omega \in \mathbb{B}$, does not have a solution if \hat{g} is an arbitrary square-integrable function.

The previous problem can be reformulated as a constrained least-squares problem. We define two projection operators. The first is the orthogonal projection onto the subspace of square-integrable band-limited functions whose band is contained in \mathbb{B} . This operator, denoted by $P^{(\mathbb{B})}$, can be written as a convolution operator

$$(P^{(\mathbb{B})} f)(\mathbf{x}) = (H_{\mathbb{B}} * f)(\mathbf{x}) \quad (5.1)$$

where $H_{\mathbb{B}}$ is the inverse Fourier transform of the characteristic function of the domain \mathbb{B}

$$H_{\mathbb{B}}(\mathbf{x}) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \chi_{\mathbb{B}}(\omega) \exp(i\mathbf{x} \cdot \omega) d\omega. \quad (5.2)$$

The second, denoted by $P^{(\mathbb{D})}$, is the orthogonal projection onto the subspace of the square-integrable functions whose support is contained in \mathbb{D}

$$(P^{(\mathbb{D})} f)(\mathbf{x}) = \chi_{\mathbb{D}}(\mathbf{x}) f(\mathbf{x}) \tag{5.3}$$

where $\chi_{\mathbb{D}}$ is the characteristic function of the domain.

Then the problem of out-of-band extrapolation of a space-limited function is equivalent to the following constrained least-squares problem: find a function $f \in L^2(\mathbb{R}^2)$ such that

$$\|P^{(\mathbb{B})} f - g\| = \text{minimum} \quad P^{(\mathbb{D})} f = f \tag{5.4}$$

where g is the inverse Fourier transform of the function which coincides with \hat{g} over \mathbb{B} and is zero elsewhere. Since $(P^{(\mathbb{B})})^2 = P^{(\mathbb{B})}$ and $P^{(\mathbb{B})} g = g$, the Euler equation associated with this problem is given by

$$P^{(\mathbb{B})} f = g. \tag{5.5}$$

On the other hand, the subspace of the square-integrable functions supported by \mathbb{D} is isomorphic to $L^2(\mathbb{D})$. If we denote by A the integral operator which is isomorphic to $P^{(\mathbb{B})} P^{(\mathbb{D})}$ when restricted to the previous subspace:

$$(Af)(\mathbf{x}) = \int_{\mathbb{D}} H_{\mathbb{B}}(\mathbf{x} - \mathbf{x}') f(\mathbf{x}') d\mathbf{x}' \tag{5.6}$$

then the constrained least-squares problem (5.4) is equivalent to the non-constrained problem (2.1) with A defined as in equation (5.6) and $f \in L^2(\mathbb{D})$. The operator $A : L^2(\mathbb{D}) \rightarrow L^2(\mathbb{R}^2)$ is compact and its singular functions are related to the generalized prolate spheroidal functions introduced by Slepian [15] while its singular values are the square roots of the eigenvalues associated with these functions.

As we have noted, the problem of out-of-band extrapolation is ill-posed. Anyway, it has been shown that some significant extrapolation is possible when there is a certain relationship between the size of \mathbb{B} and that of \mathbb{D} . If \mathbb{B} is contained in a square centred at the origin with size 2Ω and, similarly, \mathbb{D} is contained in a square centred at the origin with size $2X$, then out-of-band extrapolation is feasible when the parameter $c = \Omega X$ is not much larger than 1. In such a case, one also says that super-resolution has been achieved because it is possible to achieve a resolution better than the Nyquist sampling distance π/Ω . Estimates of the improvement can be obtained by means of the singular system of the operator (5.6) [16].

Since the formulations (5.4) and (5.6) of the problem of out-of-band extrapolation are equivalent, the projected Landweber method for (5.4) is equivalent to the Landweber method for (5.6). The largest singular value of A , σ_1 , is less than 1 so that we can choose $\tau = 1$ in equation (2.10) and we obtain, for problem (5.4)

$$f_{n+1} = P_{\mathbb{D}}[g + (I - P_{\mathbb{B}})f_n] \tag{5.7}$$

with $f_0 = 0$. The implementation of this method is precisely the Gerchberg algorithm:

- compute $f_1 = P_{\mathbb{D}}g$;
- from f_n compute \hat{f}_n ;
- compute $\hat{g}_{n+1}(\omega) = \hat{g}(\omega) + [1 - \chi_{\mathbb{B}}(\omega)]\hat{f}_n(\omega)$;
- from \hat{g}_{n+1} compute g_{n+1} ;
- compute $f_{n+1} = \chi_{\mathbb{D}}g_{n+1}$.

On the other hand, from the Landweber method applied to the integral operator (5.6) one derives in a standard way the singular functions expansion of the n th iterate of the Gerchberg method

$$f_n = \sum_{k=0}^{\infty} \{1 - (1 - \sigma_k^2)^n\} \frac{(g, v_k)}{\sigma_k} u_k. \tag{5.8}$$

It is obvious that, while the Gerchberg algorithm is easily computable for arbitrary shapes of \mathbb{B} and \mathbb{D} , this is not true for the expansion (5.8). On the other hand, the latter clearly indicates the filtering effect of the iterative process, which is not evident in the scheme previously described. We think important to point out that the algorithm can also be easily applied to the case where \mathbb{D} is the union of a finite set of disjoint domains $\mathbb{D}_1, \dots, \mathbb{D}_N$, each one having linear dimensions of the order of the Nyquist distance π/Ω but such that the distance between any pair is much greater than π/Ω . A function whose support satisfies these conditions corresponds essentially to the *nearly black objects* considered by Donoho *et al* [17] and also in this case super-resolution can be achieved. However, in order to compare the results provided by the Gerchberg method with those provided by maximum entropy [17] it should be necessary to consider the Gerchberg method with the additional constraint of positivity. This additional constraint will be discussed in the next section.

The extension of the Gerchberg method to the case of blurred images, already investigated by Gori in the case of band-limited blurring [18], is now easy. Suppose that we wish to restore a function f , with support contained in \mathbb{D} , a blurred image g of f being known, the point spread function (PSF) of the imaging system being K . Then we have to solve the problem

$$\|K * f - g\| = \text{minimum} \quad P_{\mathbb{D}}f = f \quad (5.9)$$

or equivalently the problem associated with the integral operator

$$(Af)(x) = \int_{\mathbb{D}} K(x - x')f(x') dx'. \quad (5.10)$$

Again $A : L^2(\mathbb{D}) \rightarrow L^2(\mathbb{R}^2)$ is a compact operator if $\hat{K}(\omega)$ is bounded.

Since $\|A\| \leq \sup_{\omega} |\hat{K}(\omega)|$, by using a value of the relaxation parameter τ which satisfies the conditions (2.5) we have, for the projected Landweber method applied to (5.9)

$$f_{n+1} = P_{\mathbb{D}}[f_n + \tau(K^* * g - K^* * K * f_n)] \quad (5.11)$$

whose implementation, in the case $f_0 = 0$, is as follows:

- compute $f_1 = \tau P_{\mathbb{D}}(K^* * g)$;
- from f_n compute \hat{f}_n ;
- compute $\hat{g}_{n+1}(\omega) = \tau \hat{K}(\omega) + (1 - \tau |\hat{K}(\omega)|^2) \hat{f}_n(\omega)$;
- from \hat{g}_{n+1} compute g_{n+1} ;
- compute $f_{n+1} = P_{\mathbb{D}}g_{n+1}$.

It is obvious that, also in this case, one needs the computation of one direct and one inverse Fourier transform at each iteration step and therefore one can achieve super-resolution by performing simultaneously deblurring and out-of-band extrapolation without using the singular functions of the operator (5.10) whose computation can be very difficult.

In order to provide examples of the method and investigate the acceleration achievable with Tikhonov preconditioning, we consider (for simplicity) one-dimensional problems. More precisely, we consider the two following examples, one corresponding to a space-limited blur and the other to a band-limited blur.

(1) The PSF of the blur is due to uniform motion of the image [19], i.e. it is the characteristic function of an interval $[0, a]$. By a change of variables it is always possible to reduce this interval to the interval $[-1, 1]$. Then

$$K(x) = \chi_{[-1,1]}(x) \quad (5.12)$$

whose Fourier transform is given by

$$\hat{K}(\omega) = 2 \frac{\sin \omega}{\omega}. \quad (5.13)$$

The typical feature of this function is that it has a countable set of zeros at the frequencies $\omega_k = k\pi$, ($k = \pm 1, \pm 2, \dots$). Then, linear methods, such as Tikhonov or Landweber without constraint on the support, provide estimates of the solution whose Fourier transform is zero at the frequencies ω_k . As a consequence, these methods do not provide information about the solution in a neighbourhood of the zeros of $\hat{K}(\omega)$.

(2) The PSF of an ideal imaging system in the case of incoherent illumination [18]:

$$K(x) = \left(\frac{\sin \pi x}{\pi x} \right)^2 \tag{5.14}$$

whose Fourier transform is given by

$$\hat{K}(\omega) = \left(1 - \frac{|\omega|}{2\pi} \right) \chi_{[-2\pi, 2\pi]}(\omega). \tag{5.15}$$

This PSF is band-limited and its band coincides with the interval $[-2\pi, 2\pi]$. According to the usual definitions, the Rayleigh resolution distance associated with this PSF is 1 while the Nyquist distance associated with the band is 1/2. Again linear methods without constraint on the support provide estimates of the solution which are band-limited, i.e. without information about its Fourier transform outside the band.

In the case of the first PSF we have considered the restoration of a Gaussian with $\sigma = 0.1$. The image has been perturbed with multiplicative Gaussian noise. We have used Tikhonov preconditioning with various values of α and we have assumed $\mathbb{D} = [-0.5, 0.5]$ (outside this interval the Gaussian is essentially zero). For each iteration we have computed the relative restoration error

$$\delta_n = \frac{\|f_n - f^{(0)}\|}{\|f^{(0)}\|} \tag{5.16}$$

where $f^{(0)}$ is the Gaussian mentioned above and the relative residual

$$\epsilon_n = \frac{\|Af_n - g\|}{\|g\|} \tag{5.17}$$

where g is the noisy image. For each numerical experiment the optimum number of iterations, n_{opt} , has been determined as the value of n minimizing δ_n . The results are reported in table 1.

We see that the acceleration obtained by decreasing α is considerable and that the minimum restoration error does not depend significantly on α . This result is not surprising because the projected method we are considering is essentially a way for implementing the Landweber method for the compact operator defined by the constraint on the support. As a consequence, the preconditioned method is precisely the Strand method.

We note that for fixed values of α the optimum number of iterations n_{opt} decreases when the amount of the noise increases. Moreover, the corresponding residuals ϵ_{opt} are always of the order of the error on the data.

An example of restoration is given in figures 3 and 4. In figure 3 we plot the object function, the data function obtained by convolving the object with the PSF (5.12) and by affecting the result with 5% Gaussian noise and the respective Fourier transforms. Figure 4 contains the reconstructions provided by the Landweber method and by the constrained preconditioned method for the value of α which gives the fastest convergence ($\alpha = 0.01$, $n_{\text{opt}} = 3$). We point out again that the Landweber method without constraints acts as a linear filter so that the Fourier transform of the regularized solution is characterized by holes in correspondence with the zeros of the Fourier transform of the kernel. In contrast, the use of the constraint on the support allows the restoration of the Fourier transform of $f^{(0)}$ also

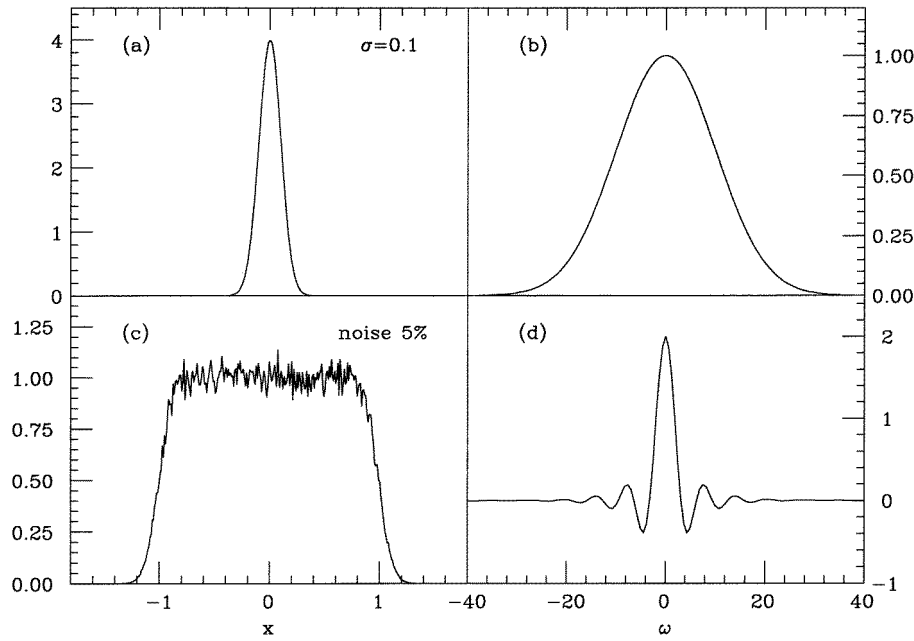


Figure 3. Restoration of a Gaussian function ($\sigma = 0.1$) from the blurred and noisy image provided by the PSF represented by the characteristic function of the interval $[-1, 1]$. (a) Input function; (b) Fourier transform of the input function; (c) data function affected by 5% Gaussian noise; (d) Fourier transform of the data function.

Table 1. Results for the restoration of a Gaussian blurred by the kernel (5.12). δ_{opt} and ϵ_{opt} are the values of δ_n and ϵ_n corresponding to $n = n_{\text{opt}}$. The upper part corresponds to 1% multiplicative Gaussian noise, the central part to 3% multiplicative Gaussian noise and the lower part to 5% multiplicative Gaussian noise. The parameter α corresponds to the Tikhonov preconditioner, equation (4.9).

Noise		$\alpha = 1$	$\alpha = 0.5$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$
1%	n_{opt}	351	196	44	23	6
	δ_{opt}	2.6	2.7	2.7	2.8	3.2
	ϵ_{opt}	0.9	0.9	0.9	0.9	0.9
3%	n_{opt}	218	122	28	15	4
	δ_{opt}	5.5	5.6	5.8	6.0	7.1
	ϵ_{opt}	2.8	2.8	2.8	2.8	2.8
5%	n_{opt}	173	97	22	12	3
	δ_{opt}	7.6	7.7	8.1	8.3	10.3
	ϵ_{opt}	4.6	4.6	4.6	4.6	4.7

in the neighbourhoods of these zeros. This result is not surprising because this constraint implies the analyticity of the Fourier transform.

In the case of the second PSF, equation (5.14), we have considered the restoration of a function $f^{(0)}$ which is the sum of two Gaussians with $\sigma = 0.1$ and centred at the points $x_1 = -0.4$ and $x_2 = 0.4$. We report in table 2 the results we have obtained.

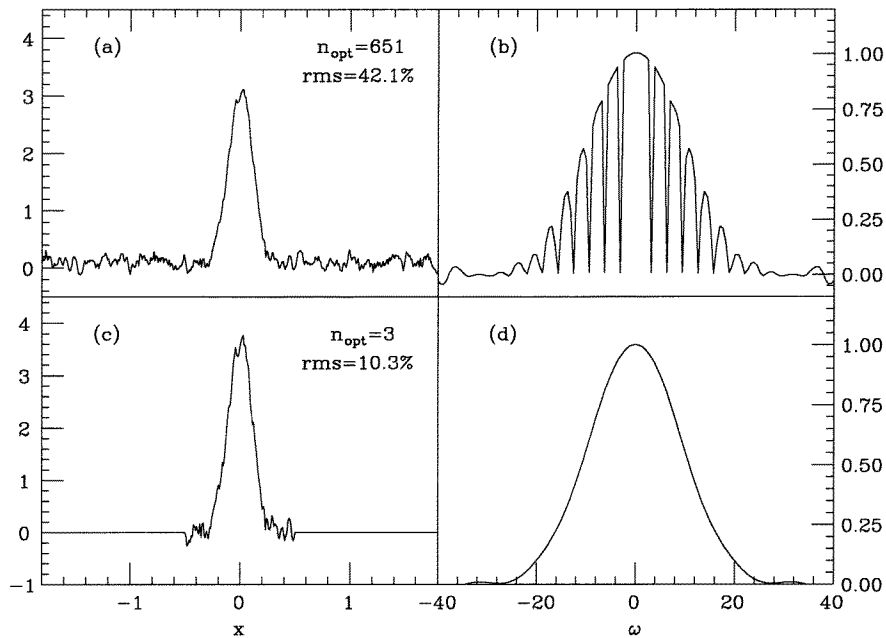


Figure 4. Restoration of a Gaussian function ($\sigma = 0.1$) from the blurred and noisy image of figure 3(c). (a) Regularized solution provided by the Landweber method without constraints; (b) Fourier transform of the regularized solution in (a). The holes are a consequence of the linearity of the method and correspond to the zeros of the Fourier transform of the PSF. (c) Regularized solution provided by the projected and preconditioned Landweber method; the support is the interval $[-0.5, 0.5]$ and Tikhonov preconditioning is adopted with $\alpha = 0.01$. (d) Fourier transform of the constrained regularized solution. The interpolating effect of the constraint is evident.

Table 2. Results for the restoration of two Gaussian functions blurred by the kernel (5.14). The quantities have the same meaning as in table 1.

Noise		$\alpha = 10^{-3}$	$\alpha = 10^{-4}$	$\alpha = 10^{-5}$	$\alpha = 10^{-6}$	$\alpha = 10^{-7}$	$\alpha = 10^{-8}$
1%	n_{opt}	1245	12 359	203	76	156	8
	δ_{opt}	38.5	35.3	36.5	34.7	27.6	28.1
	ϵ_{opt}	0.9	0.9	1.1	1.5	1.4	1.0
3%	n_{opt}	772	536	105	36	19	3
	δ_{opt}	41.4	40.3	42.9	48.8	32.0	31.6
	ϵ_{opt}	2.8	2.9	3.6	4.3	4.4	3.9
5%	n_{opt}	590	425	68	3	13	2
	δ_{opt}	45.0	46.2	48.5	55.2	36.8	34.4
	ϵ_{opt}	4.7	4.9	6.0	6.3	7.3	6.4

In this example we can use much smaller values of α for obtaining reliable reconstructions, although the restoration errors are significantly larger than in the previous case. If we look at the columns of table 2, we notice the phenomenon already observed at the end of section 4: for a given α , n_{opt} decreases as the amount of noise increases. An example of restoration is given in figures 5 and 6. We see that the constraint on the support

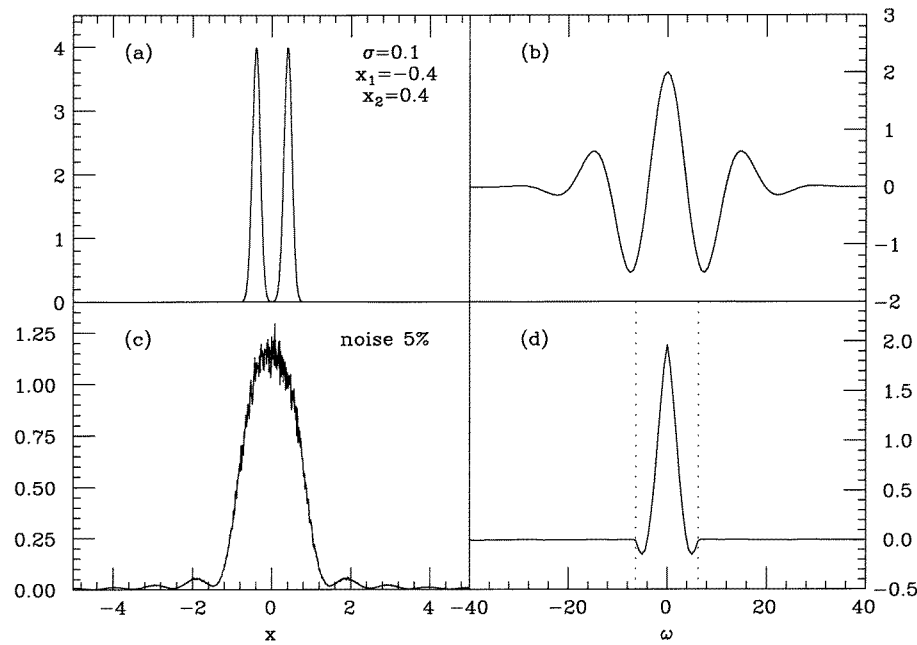


Figure 5. Restoration of two Gaussians ($\sigma = 0.1$), centred in ± 0.4 from the blurred and noisy image provided by the PSF represented by the squared sinc function, equation (5.13). (a) Input function; (b) Fourier transform of the input function; (c) data function affected by 5% Gaussian noise; (d) Fourier transform of the data function. The band is indicated by the dotted vertical lines.

allows a moderate extrapolation of the Fourier transform of $f^{(0)}$ outside the band of the PSF. This can be again justified by referring to the analyticity of the Fourier transform.

5.2. Positivity and super-resolution

The Gerchberg method and its extensions investigated in the previous subsection are still linear methods because the projection operator $P_{\mathbb{D}}$ is linear. However, although the constraint on the support can improve the resolution in image restoration, as demonstrated by the examples provided above, it still allows the so-called ringing effects [3] which are typical of all linear methods. In figure 4(c) and in figure 6(c), the ringing is due to the negative values of the solution which should be positive. It has been suggested [3] that the constraint of positivity can be used for reducing and eliminating these effects. In such a case the projected Landweber method is a genuine nonlinear iterative method.

If \mathcal{C}_+ is the closed and convex set of the non-negative functions in $L^2(\mathbb{R}^2)$, then the projection operator P_+ onto \mathcal{C}_+ is given by

$$(P_+f)(x) = \begin{cases} f(x) & f(x) \geq 0 \\ 0 & f(x) < 0 \end{cases} \quad (5.18)$$

and therefore it is easily computable. Non-negative least-squares solutions of the image restoration problem can be computed as in the scheme described in the previous subsection simply by replacing $P_{\mathbb{D}}g_{n+1}$ by P_+g_{n+1} at the last step. Therefore, also in this case one needs only the computation of one direct and one inverse Fourier transform at each iteration step.

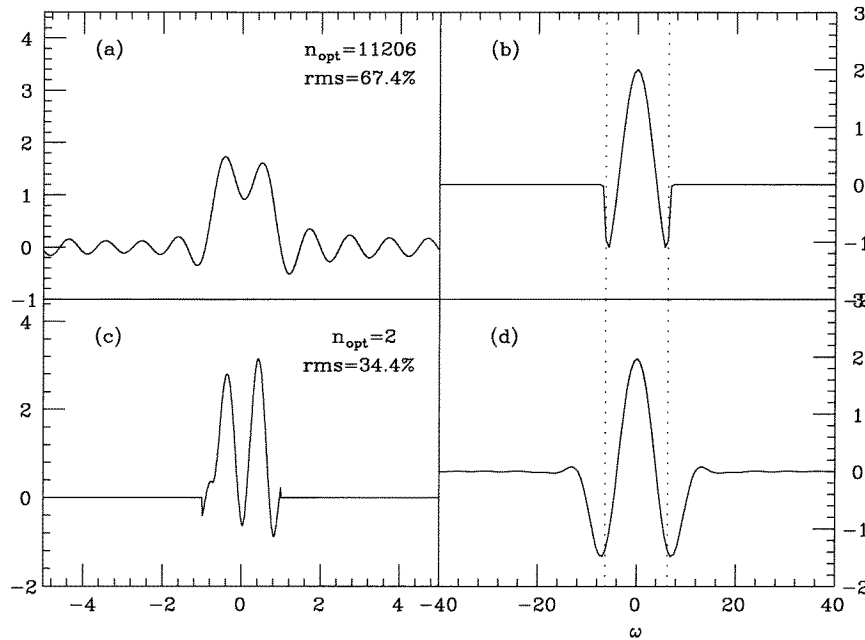


Figure 6. Restoration of the two Gaussian peaks of figure 5(a) from the data of figure 5(c). (a) Regularized solution provided by the Landweber method without constraints; (b) Fourier transform of the regularized solution in (a); (c) regularized solution provided by the projected Landweber method with the constraint of bounded support, interval $[-1, 1]$; Tikhonov preconditioning is adopted with $\alpha = 10^{-6}$; (d) Fourier transform of the constrained regularized solution. The use of the constraint implies a moderate out-of-band extrapolation. The band is indicated by the dotted vertical lines.

Since the method can be easily implemented, it is an important tool for investigating the effect of positivity in image restoration. The conclusion which can be derived from our numerical experiments is that positivity can have both interpolating and extrapolating effects in the Fourier domain, similar to those obtained in the case of a constraint on the support.

Examples of the performances of the Landweber method with positivity are given in figures 7 and 8. In particular, figures 7(a) and 7(b) contain the optimal approximate solution and its Fourier transform in the same case considered in figures 3 and 4, while figures 8(a) and 8(b) refer to the same case considered in figures 5 and 6. It is clear that the use of positivity allows one to eliminate the ringing effects which appear in the reconstructions given both by the non-constrained and constrained Landweber methods. Moreover, in the case of uniform motion blur, the nonlinearity of the method with positivity implies strong interpolating effects (figure 7(b)): the holes in the Fourier transform of the regularized solution provided by the non-constrained Landweber method and corresponding to the zeros of the Fourier transform of the PSF are completely filled here. Finally, the use of the positivity constraint in the case of incoherent illumination implies a significant out-of-band extrapolation (figure 8(b)). Nevertheless, while the super-resolution effects provided by the Gerchberg algorithm can be easily explained in terms of the analyticity of the Fourier transform, here the reasons for this extrapolation are less clear, so that we can say that super-resolution in the case of the Landweber method with positivity is still an open problem.

In view of practical applications, we want to point out that the large number of iterations necessary to obtain the reconstructions of figures 7 and 8 ($n_{\text{opt}} = 957$ for figure 7 and $n_{\text{opt}} = 35\,195$ for figure 8) corresponds to extremely flat minima of the restoration error as a function of the iteration number. For example, in the case of the reconstruction of the Gaussian function, a restoration error of 11.6% is obtained after 400 iterations while a restoration error of 7.5% is obtained after 600 iterations. This behaviour is even more significant for the reconstruction of the two Gaussian peaks. In this case a restoration error of 30.8% is obtained after 2000 iterations while 7000 iterations corresponds to 20.2% error. This means that sufficiently reliable reconstructions can be obtained after a number of iterations much smaller than n_{opt} .

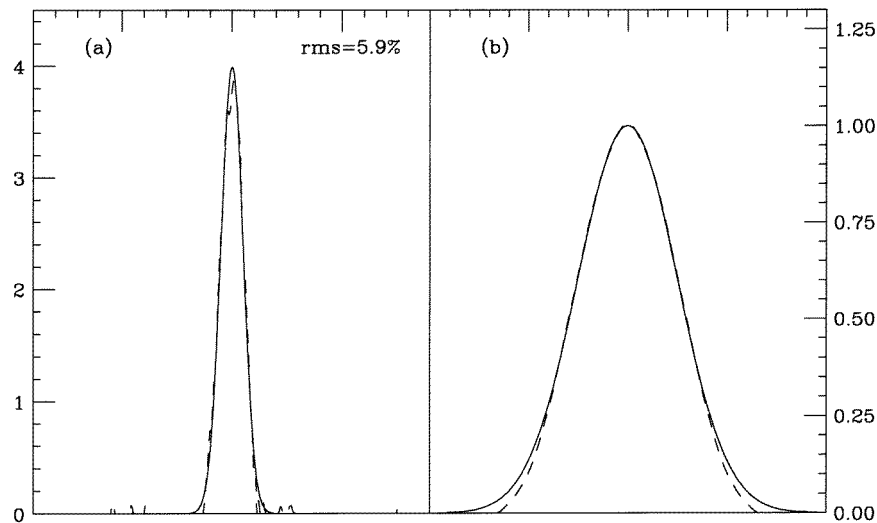


Figure 7. Example of the performances of the Landweber method with the positivity constraint. (a) Optimal approximation (dashed curve) of the Gaussian function ($\sigma = 0.1$) of figure 3(a) (full curve) from the blurred and noisy image of figure 3(c). (b) Fourier transforms of the functions of (a).

Anyway, the convergence rate of the algorithm can be increased by the use of Tikhonov preconditioning (4.9). In table 3 we give the results of the restoration of the Gaussian function when the blur is given by the characteristic function of the interval $[-1, 1]$. These results must be compared with those contained in table 1. We find that the behaviour of the Landweber method with positivity is substantially equivalent to the behaviour of the Landweber method with the constraint of bounded support. In fact we have a considerable acceleration by decreasing α without a significant increase of the restoration error. Moreover, at least for the smallest values of α , positivity seems to be more efficient than the bounded support constraint.

Things are different if we consider the restoration of the two Gaussian peaks from the blurred image obtained by the convolution with the squared sinc function (5.14). From table 4 it follows that we have again acceleration but the restoration error increases too much for small values of α and the correspondent values of the residuals become intolerably large. However, when α is not too small the restoration error obtained by using only the positivity constraint is much smaller than the restoration error obtained by using the support constraint. We note that for too small values of α the restoration error may be larger than

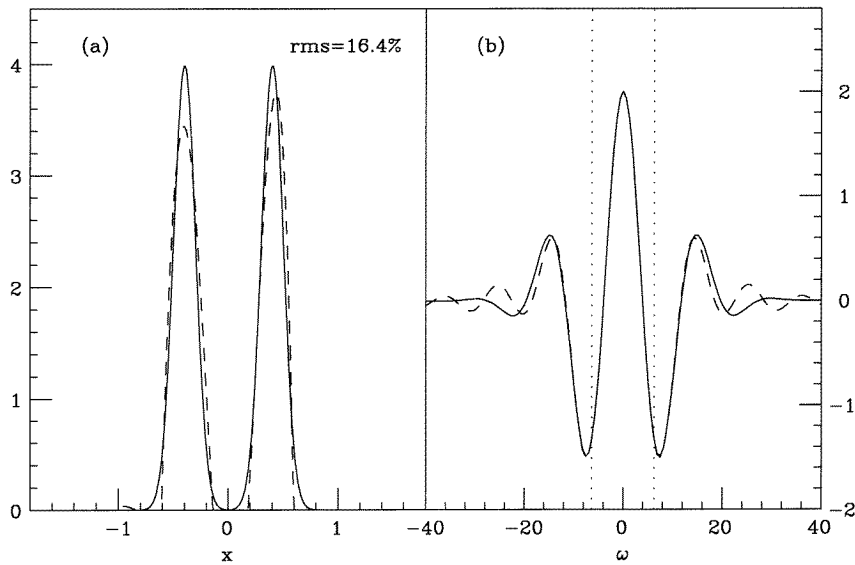


Figure 8. Example of the performances of the Landweber method with the positivity constraint. (a) Optimal approximation (dashed curve) of the input function of figure 5(a) (full curve) from the blurred and noisy image of figure 5(c). (b) Fourier transforms of the functions of (a). The band is indicated by the dotted vertical lines.

Table 3. Results for the restoration of a Gaussian blurred by the kernel (5.12). δ_{opt} and ϵ_{opt} are the values of δ_n and ϵ_n corresponding to $n = n_{opt}$. The upper part corresponds to 1% multiplicative Gaussian noise, the central part to 3% multiplicative Gaussian noise and the lower part to 5% multiplicative Gaussian noise. The parameter α corresponds to the Tikhonov preconditioner.

Noise		$\alpha = 1$	$\alpha = 0.5$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$
1%	n_{opt}	410	226	49	26	7
	δ_{opt}	2.5	2.5	2.5	2.6	3.8
	ϵ_{opt}	0.9	0.9	1.0	1.2	2.6
3%	n_{opt}	236	131	30	16	4
	δ_{opt}	4.7	4.8	5.2	5.8	9.1
	ϵ_{opt}	2.8	2.8	3.1	3.6	7.6
5%	n_{opt}	190	106	24	13	4
	δ_{opt}	6.2	6.4	7.4	8.2	13.4
	ϵ_{opt}	4.6	4.7	5.2	5.9	11.8

100% and is an always increasing function of the number of iterations. The boxes in table 4 corresponding to these cases have been left empty.

Finally, we remark that the conditions of positivity and compact support can be used simultaneously. In this case, since the operators P_+ and $P^{(D)}$ commute, the projection operator is $P_C = P_+ P_D$ and the process can be again preconditioned in order to accelerate the convergence. It is interesting to note that in this case we can increase the convergence speed by decreasing the Tikhonov parameter without increasing the restoration error which is still of the order of the best restoration error obtained with positivity alone. For example, in

Table 4. Results for the restoration of two Gaussian functions blurred by the kernel (5.14). The quantities have the same meaning as in table 3.

Noise		$\alpha = 10^{-2}$	$\alpha = 10^{-3}$	$\alpha = 10^{-4}$	$\alpha = 10^{-5}$	$\alpha = 10^{-6}$	$\alpha = 10^{-7}$
1%	n_{opt}	8917	2425	3043	268	15	1
	δ_{opt}	15.1	15.7	15.4	25.2	51.4	73.9
	ϵ_{opt}	0.9	1.1	2.1	9.0	32.7	63.6
3%	n_{opt}	3094	1015	855	41	1	
	δ_{opt}	16.4	20.0	22.4	49.5	75.7	
	ϵ_{opt}	2.8	3.4	6.5	26.4	56.7	
5%	n_{opt}	2066	589	365	3		
	δ_{opt}	18.1	24.6	30.7	64.1		
	ϵ_{opt}	4.6	5.6	11.0	34.1		

the case of uniform motion blur and reconstruction of the Gaussian function, the restoration error corresponding to $\alpha = 0.01$ and 5% noise is 8.7% after four iterations (13.4% after four iterations when the Landweber method with positivity is applied, see table 3) while in the case of an incoherently illuminated imaging system and reconstruction of two Gaussian functions the restoration error for $\alpha = 10^{-4}$ and 5% noise is 29.7% after 162 iterations (if positivity only is adopted the error is 30.7% after 365 iterations, see table 4).

6. Concluding remarks

In this paper we have shown that the projected Landweber method can be very useful in problems of image restoration and it can also be an efficient numerical method by an appropriate use of preconditioning techniques. For this reason we think that this method deserves more attention than in the past. In particular, it should be interesting to prove that it is a genuine regularization method, i.e. strong convergence of the iterates in the case of noise-free data and semiconvergence of the iterates in the case of noisy data. The results of Eicke [4] are rather preliminary in this direction.

Other questions raised by our analysis are the following. The results obtained by means of the Landweber method with positivity exhibit considerable interpolation and extrapolation properties in the Fourier domain. These properties have been confirmed by a recent application to an inverse problem in seismology [20] and are similar to properties of the CLEAN algorithm [21] which is used in the astronomical community. Therefore, these properties seem to be specific to the positivity constraint and not to the particular algorithms used for implementing this constraint. It should be important to justify theoretically such a result.

Finally, in the case of the preconditioned method, since we have focused on the so-called Tikhonov preconditioning, we have essentially considered the case where a one-parameter family of preconditioners is used. In such a case the dependence of the restoration error and of the optimum number of iterations on the preconditioning parameter should be theoretically understood. Moreover, some objective criteria to choose this parameter should be introduced. These problems are clearly related to the unsolved problem of the regularization properties of the projected Landweber method.

References

- [1] Ivanov V K 1962 On linear problems which are not well-posed *Sov. Math. Dokl.* **3** 981–3
- [2] Neubauer A 1988 Tikhonov-regularization of ill-posed linear operator equations on closed convex sets *J. Approx. Theor.* **53** 304–20
- [3] Lagendijk R, Biemond J and Boeke D 1988 Regularized iterative image restoration with ringing reduction *IEEE Trans. Acoust. Speech Signal Process.* **36** 1874–88
- [4] Eicke B 1992 Iteration methods for convexly constrained ill-posed problems in Hilbert space *Num. Funct. Anal. Opt.* **13** 413
- [5] Youla D C and Webb H 1989 Image restoration by the method of convex projections, Part 1. Theory *IEEE Trans. Med. Imaging* **MI-1** 81–94
- [6] Golub G H and Van Loan C F 1989 *Matrix Computations* (Baltimore, MD: Johns Hopkins)
- [7] Strand O N 1974 Theory and methods related to the singular-function expansion and Landweber's iteration for integral equations of the first kind *SIAM J. Num. Anal.* **11** 798
- [8] Nashed M Z (ed) 1976 *Generalized Inverses and Applications* (New York: Academic)
- [9] Youla D C 1987 Mathematical theory of image restoration by the method of convex projections *Image Recovery: Theory and Applications* (New York: Academic)
- [10] Bialy H 1959 Iterative behandlung linearer funktionalgleichungen *Arch. Rat. Mech. Anal.* **4** 166–76
- [11] Vasin V V and Ageev A L 1995 *Ill-posed Problems with a priori Information* (Utrecht: VSP)
- [12] Sanz J C and Huang T S 1983 Unified Hilbert space approach to iterative least-squares linear signal restoration *J. Opt. Soc. Am.* **73** 1455–65
- [13] Defrise M and De Mol C 1987 A note on stopping rules for iterative regularization methods and filtered SVD *Inverse Problems: An Interdisciplinary Study (Adv. Electr. Electr. Phys., Suppl. 19)* ed P C Sabatier (New York: Academic)
- [14] Gerchberg R W 1974 Super-resolution through error energy reduction *Opt. Acta* **21** 709–20
- [15] Slepian D 1964 Prolate spheroidal wave functions, Fourier analysis and uncertainty, IV. Extensions to many dimensions, generalized prolate spheroidal functions *Bell. Syst. Tech. J.* **43** 3009–57
- [16] Bertero M and Pike E R 1982 Resolution in diffraction-limited imaging, a singular value analysis, I. The case of coherent illumination *Opt. Acta* **29** 727–46
- [17] Donoho D L, Johnstone I M, Hoch J C and Stern A S 1991 Maximum entropy and the nearly black object *J. R. Stat. Soc. B* **54** 41–82
- [18] Gori F 1975 On an iterative method for super-resolution *Digest of the International Optical Computing Conf. (Washington, DC, 1975)* (IEEE Catalog No 75 CH0941) pp 137–41
- [19] Slepian D 1967 Restoration of photographs blurred by image motion *Bell Syst. Tech. J.* **46** 2353–62
- [20] Bertero M, Bindi D, Boccacci P, Cattaneo M, Eva C and Lanza V 1997 Application of the projected Landweber method to the estimation of the source time function in seismology *Inverse Problems* **13** 465–86
- [21] Fried D L 1995 Analysis of the CLEAN algorithm and implication of superresolution *J. Opt. Soc. Am. A* **12** 853–60