PRECONDITIONED ITERATIVE REGULARIZATION FOR Ill-POSED PROBLEMS

By

Martin Hanke
James Nagy
and
Robert Plemmons

IMA Preprint Series # 1024
September 1992

Preconditioned Iterative Regularization for Ill–Posed Problems

Martin Hanke [*], *James Nagy* [†] *and Robert Plemmons* [‡]

Abstract.
A preconditioned iterative regularization scheme is proposed and tested for solving large scale structured linear systems $H\mathbf{f} = \mathbf{g} + \eta$, arising from the discretization of ill-posed inverse problems in the presence of noise. The case where H is a block Toeplitz matrix with Toeplitz blocks is considered. The preconditioned conjugate gradient method is applied, where H is approximated by a block circulant matrix with circulant blocks. This results in a fast 2-D FFT-based iterative scheme with applications, *e.g.*, in image restoration. Thus we precondition in the Fourier domain, while iterating in the spatial domain. Our main purpose is to show how the iterations can be effectively and efficiently regularized for solving ill-posed problems by using the spectral decomposition of the preconditioner.

Key Words. Toeplitz-block system, FFT–based preconditioners, ill-posed inverse problems, conjugate gradients, least squares, regularization, deconvolution, image restoration.

AMS(MOS) Subject Classifications. 65F10, 65F15, 43E10.

1 Introduction

In many engineering applications (*e.g.*, image restoration [12]) it is often necessary to numerically solve large scale structured $N \times N$ dimensional linear systems

$$H\mathbf{f} = \bar{\mathbf{g}}, \tag{1}$$

*Institut für Praktische Mathematik, Universität Karlsruhe, Englerstrasse 2, W-7500, Karlsruhe, Germany. Research supported by the Deutsche Forschungsgemeinschaft (DFG).

†Department of Mathematics, Southern Methodist University, Dallas, TX 75275-0461.

‡Department of Mathematics and Computer Science, Wake Forest University, Box 7388, Winston-Salem, NC 27109. Cray Y-MP time was provided by a grant from the North Carolina Supercomputing Center. Research supported by the US Air Force under grant no. AFOSR–91–0163 and NSF grant no. CCR–92–01105.

arising from discretization of ill-posed inverse problems. For example, (1) may result from the discretization of a first kind integral operator equation. Numerical methods for such problems can be highly ill-conditioned because the minimum singular value of H is close to zero. Moreover, for these ill-posed problems the matrix H is typically full. Very often, however, the H will have a Toeplitz, or Toeplitz related, block structure. An $n \times n$ matrix T is called a Toeplitz matrix if the entries of T satisfy $t_{i,j} = t_{i-j}$, *i.e.,* the entries of T are constant down each diagonal. If the entries of an $n \times n$ Toeplitz matrix C also satisfy $c_{n-j} = c_{-j}$, then C is called a circulant matrix.

In this paper we will consider solving (1) when the matrix H is an $m \times m$ block Toeplitz matrix, with each block being an $n \times n$ (point) Toeplitz matrix. Specifically, we will assume that

$$H = \begin{bmatrix} T_0 & T_{-1} & \cdots & T_{-m+1} \\ T_1 & T_0 & \cdots & T_{-m+2} \\ \vdots & \vdots & & \vdots \\ T_{m-1} & T_{m-2} & \cdots & T_0 \end{bmatrix}, \tag{2}$$

where each block T_k is an $n \times n$ point Toeplitz matrix. We call a matrix of this form a *block Toeplitz with Toeplitz blocks* matrix, and denote it by BTTB(m,n). Such matrices arise naturally from discretization of 2-D Fredholm integral equations of the first kind with displacement kernels. Since, in this case, the continuous problem is ill-posed (cf. Groetsch [8]), it follows then that the matrix H will be very ill-conditioned; we call (1) a *discrete ill-posed inverse problem.* As a consequence of the ill-posedness, the solution $\mathbf{f}$ will be very sensitive to perturbations in the right hand side. This is an important observation since in most applications $\mathbf{g}$ is either a recorded signal or image, or is obtained from measurements.

In general, what is known is not $\bar{\mathbf{g}} = H\mathbf{f}$, but

$$\mathbf{g} = H\mathbf{f} + \boldsymbol{\eta},$$

where $\boldsymbol{\eta}$ represents unknown (Gaussian) noise or measurement errors. Because of the above comments on the ill-conditioned nature of H, one cannot simply solve $H\mathbf{f} = \mathbf{g}$ and expect to compute an accurate approximation to $\mathbf{f}$. In particular, $\mathbf{g}$ can be dominated by noise at high frequencies. Thus if H is ill-conditioned, then the *noise term* will become amplified at the higher frequencies, since the power spectrum of the degraded image is typically highest at low frequencies and rolls off significantly for higher ones. Therefore the higher frequencies, which are usually of more interest to the user, are degraded the most in approximating $\mathbf{f}$.

Regularization is often used to alleviate the sensitivity of the inverse problem to the perturbations to the right hand side. Regularization has the effect of filtering out the noise at the expense of restricting the set of admissible solutions. Various regularization procedures include Tikhonov [8], truncated SVD [11], and iterative methods[9], especially conjugate gradient iterations [17]. In each case, one has the non-trivial problem of choosing a *regularization parameter*, which controls the degree of bias in the solution. For example, in conjugate gradients, the regularization parameter is the iteration number, *i.e.*, the number of iterations applied.

Here we consider using the preconditioned conjugate gradient algorithm as a regularization procedure. Moreover, we study the potential use of a *block circulant with circulant blocks* (BCCB) matrix C as a preconditioner. BCCB preconditioners have been shown to be effective in speeding the convergence of the conjugate gradient algorithm for solving well-conditioned BTTB systems of equations [2, 6]. In addition, the Fast Fourier Transform (FFT) can be used in the computations. These preconditioners were extended to least squares problems by R. Chan, Nagy and Plemmons [3, 4], where it is shown how the FFT–based computations can be applied to ill-posed problems when Tikhonov regularization is used.

We stress that to the best of our knowledge preconditioning effects have not yet been considered for iterative regularization methods; the topic is especially delicate, for either the preconditioner is as badly conditioned as is H, in which case each iterative step will already be unstable, or else C is a poor approximation of H, in which case the preconditioner will not be effective. We will show here how to adapt the BCCB preconditioner of [4, 6] to the context of discrete ill-posed equations arising from image restoration applications.

This paper is outlined as follows. In Section 2 we review the construction of BCCB preconditioners for BTTB matrices. Section 3 is concerned with the relationship between regularization and the singular value decomposition. In Section 4 we discuss regularization by iteration, with an emphasis on preconditioned conjugate gradient iterations. In Sections 5 and 6 we consider the problems of constructing a BCCB preconditioner for ill-conditioned BTTB matrices. Here we propose a particular way of adapting the BCCB preconditioner matrix using its FFT-based spectral decomposition. Finally, in Section 7 we provide some numerical results illustrating the features and the effectiveness of our scheme.

2 Preconditioned 2-D Problems

In this section we review conjugate gradient preconditioners for the iterative solution of some large-scale BTTB Toeplitz, *i.e.* 2-D, problems. Toeplitz related matrices arise in very many scientific applications, such as signal and image processing.

Traditional direct methods for solving point $n \times n$ Toeplitz systems of linear equations have $O(n^2)$ computational complexity, see *e.g.*, [7]. Fast direct Toeplitz solvers have complexity $O(n \log^2 n)$, but their numerical stability is not completely understood. In 1986, Gilbert Strang [16] addressed the question of whether iterative methods can compete with direct methods for solving symmetric positive definite Toeplitz systems. Strang proposed the use of circulant matrices to precondition conjugate gradient iterations for Toeplitz systems in order to accelerate the convergence. The reason why this approach is competitive with direct methods is clear. The use of circulant preconditioners for these Toeplitz problems allows the use of Fourier transforms throughout the computations, and these FFT-based iterations are not only numerically efficient, but also highly parallelizable.

Numerous articles have extended the Strang idea to more general point Toeplitz systems, and several types of circulant preconditioners have been suggested, *e.g.*, [5]. Recently, FFT-based preconditioners have been proposed for solving BTTB problems by T. Chan and Olkin [6] and, for the least squares case, by R. Chan, Nagy and Plemmons [3, 4].

The conjugate gradient (CG) method is an iterative method of fundamental importance for solving systems of linear equations (1), *cf.*, Golub and van Loan [7]. The classical CG method applies only to the case where H is symmetric positive definite. But if H is not symmetric positive definite one can still use the CG method; for example, by considering the factored form of the normal equations

$$H^*(\mathbf{g} - H\mathbf{f}) = \mathbf{0},$$

where $*$ denotes the conjugate transpose. This variant of CG is generally called CGNR.

The convergence rate of the conjugate gradient algorithm depends on the singular values of the coefficient matrix H. If the singular values cluster around a fixed point, convergence will be rapid, *cf.* van der Sluis and van der Vorst [18]. Thus, to make the algorithm a useful iterative method, one usually *preconditions* the system. The preconditioned conjugate gradient algorithm, with preconditioner C, uses the conjugate gradient method to solve

$$HC^{-1}\mathbf{y} = \mathbf{g},$$

with $C\mathbf{f} = \mathbf{y}$.

This leads to the preconditioned conjugate gradient (PCGNR) algorithm for general linear systems and for least squares problems. The version of the PCGNR algorithm we use is given in [1] and can be stated as follows:

Algorithm 1. PCGNR. *Let* $\mathbf{f}^{(0)}$ *be an initial vector, and let* C *be a given preconditioner. This algorithm computes the solution,* $\mathbf{f}$, *to* $H\mathbf{f} = \mathbf{g}$.

- $\mathbf{r}^{(0)} = \mathbf{g} - H\mathbf{f}^{(0)}$
- $\mathbf{p}^{(0)} = \mathbf{s}^{(0)} = C^{-*}H^*\mathbf{r}^{(0)}$
- $\gamma_0 = \|\mathbf{s}^{(0)}\|_2^2$
- **for** $k = 0, 1, 2, \ldots$

$$\begin{aligned}
&\mathbf{q}^{(k)} = HC^{-1}\mathbf{p}^{(k)}\\
&\alpha_k = \gamma_k / \|\mathbf{q}^{(k)}\|_2^2\\
&\mathbf{f}^{(k+1)} = \mathbf{f}^{(k)} + \alpha_k C^{-1}\mathbf{p}^{(k)}\\
&\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k \mathbf{q}^{(k)}\\
&\mathbf{s}^{(k+1)} = C^{-*}H^*\mathbf{r}^{(k+1)}\\
&\gamma_{k+1} = \|\mathbf{s}^{(k+1)}\|_2^2\\
&\beta_k = \gamma_{k+1}/\gamma_k\\
&\mathbf{p}^{(k+1)} = \mathbf{s}^{(k+1)} + \beta_k \mathbf{p}^{(k)}
\end{aligned}$$

We consider the T. Chan [5] circulant approximation for a square point Toeplitz matrix. His circulant approximation is optimal with respect to the Frobenius norm, is defined for arbitrary square matrices A, and is constructed by averaging the entries along certain pairs of diagonals of A. In particular, for a given generic

$n \times n$ matrix A, let $\mathcal{C}(A)$ be the $n \times n$ circulant approximation of A as defined in T. Chan [5], *i.e.*, C is the minimizer of $F(X) = \|A - X\|_F$ over all circulant matrices X. For the special case where $A \equiv T$, where T is Toeplitz, we have

$$C = \mathcal{C}(A) = (c_{j-\ell})_{j,\ell},$$

where the (j,ℓ)th entry of C is given by the diagonal

$$c_k = \begin{cases} \dfrac{(n-k)t_k + kt_{k-n}}{n} & 0 \leq k < n, \\ c_{n+k} & 0 < -k < n. \end{cases} \tag{3}$$

Throughout this paper we will use the T. Chan approximation $\mathcal{C}(\cdot)$ exclusively as a basic tool to construct our preconditioners for BTTB problems.

We now describe a BCCB preconditioner for BTTB, *i.e.*, 2-D, problems. It is based on second level approximations and will be called the *Level–2 preconditioner.* Our Level–2 preconditioner is based on circulant approximations at the Toeplitz-block level and again at a second, block Toeplitz level. The BTTB matrix H in (2) will be approximated by a second-level circulant preconditioner, denoted by $\mathcal{C}_2(H)$. The preconditioner is related to the one proposed by T. Chan and Olkin [6] for square Toeplitz-block matrices and extended to least squares problems by R. Chan, Nagy and Plemmons [4].

For the BTTB matrix H given by (2), we define a first-level circulant approximation by

$$\mathcal{C}_1(H) \equiv \begin{bmatrix} \mathcal{C}(T_0) & \mathcal{C}(T_{-1}) & \cdots & \mathcal{C}(T_{-m+1}) \\ \mathcal{C}(T_1) & \mathcal{C}(T_0) & \cdots & \mathcal{C}(T_{-m+2}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{C}(T_{m-1}) & \mathcal{C}(T_{m-2}) & \cdots & \mathcal{C}(T_0) \end{bmatrix}. \tag{4}$$

Now we apply another first-level circulant approximation $\tilde{\mathcal{C}}_1$ to $\mathcal{C}_1(H)$ to get the second-level circulant approximation $\mathcal{C}_2(H)$ that we want. The matrix $\tilde{\mathcal{C}}_1\{\mathcal{C}_1(H)\}$ is obtained by treating each block $\mathcal{C}(T_\mu)$ in $\mathcal{C}_1(H)$ as an entry of a matrix and then applying the formula (3) to this "point" matrix. More precisely, we form the block-circulant matrix

$$\tilde{\mathcal{C}}_1\{\mathcal{C}_1(H)\} = \begin{bmatrix} \tilde{C}_0 & \tilde{C}_{-1} & \cdots & \tilde{C}_{-m+1} \\ \tilde{C}_1 & \tilde{C}_0 & \cdots & \tilde{C}_{-m+2} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{C}_{m-1} & \tilde{C}_{m-2} & \cdots & \tilde{C}_0 \end{bmatrix}$$

where the $\tilde{C}_\mu$ are computed using (3). Since each $\mathcal{C}(T_\mu)$ is circulant, we see that the sums $\tilde{C}_\mu$ are circulant matrices. Moreover, by their definition, we see that $\tilde{C}_\mu = \tilde{C}_{\mu-m}, \quad 0 < \mu < m$. In particular, we see that $\tilde{\mathcal{C}}_1\{\mathcal{C}_1(H)\}$ is then a block-circulant matrix with circulant blocks of the form BCCB(m,n). We note further that $\mathcal{C}_2(H) = \tilde{\mathcal{C}}_1\{\mathcal{C}_1(H)\}$, as given in R. Chan and Jin [2, Theorem 3]. It is shown in [2] that $\mathcal{C}_2(H)$ is "optimal" in the sense that it minimizes $\|H - X\|_F$ over all BCCB matrices X. For our purposes, we need only find the first column of $\mathcal{C}_1(H)$, and then use that column to find the first column of the final BCCB matrix $\mathcal{C}_2(H)$.

We note that the eigenvalues of $\mathcal{C}_2(H)$ can be obtained by taking the 2-D FFT of the first column of $\mathcal{C}_2(H)$. Specifically, if $\mathbf{c}_2$ denotes the first column of $\mathcal{C}_2(H)$ and `fft2d` denotes the 2-D FFT operator, then the vector of eigenvalues of $\mathcal{C}_2(H)$ is given by `fft2d`($\mathbf{c}_2$).

The construction of our Level–2 BCCB preconditioner (*i.e.*, the eigenvalues of $C \equiv \mathcal{C}_2(H)$) for the matrix H in (2) is summarized in the following algorithm:

Algorithm 2. Construction of the circulant **Level–2** preconditioner for 2-D problems. *Let H be the BTTB(m,n) matrix given as in* (2). *This algorithm computes the eigenvalues of the BCCB preconditioner C, using the 2-D FFT, and stores them in the diagonal matrix Λ.*

Find the first column, $\mathbf{c}_1$*, of* $\mathcal{C}_1(H)$*, using* (3),

Find the first column, $\mathbf{c}_2$*, of* $C \equiv \mathcal{C}_2(H) = \tilde{\mathcal{C}}_1\{\mathcal{C}_1(H)\}$ *, using* (3),

Find Λ *from* $\Lambda\mathbf{1} =$ `fft2d`($\mathbf{c}_2$)*, using the 2-D FFT.*

Since H is $mn \times mn$, the cost of forming the Level–2 BCCB preconditioner C (*i.e.*, its eigenvalues in Λ), for the case where the matrix H is BTTB(m,n) , is $O(mn(\log n + \log m))$ operations. For implementation in the PCGNR algorithm, we need only the eigenvalues of C together with the 2-D FFT. We observe that for any vector $\mathbf{y}$, $C^{-1}\mathbf{y}$ can be computed by the formula

$$C^{-1}\mathbf{y} = \mathtt{ifft2d}[\Lambda^{-1}\mathtt{fft2d}(\mathbf{y})]$$

in

$$O(mn(\log n + \log m)) \tag{5}$$

operations by using `fft2d` and its inverse, `ifft2d`. FFT-based matrix-vector multiplications involving the BTTB matrices H and H^* in the PCGNR algorithm also have complexity (5), so (5) is the complexity per iteration of the PCGNR algorithm, implemented using the 2-D FFT.

We note that a convergence analysis of the PCGNR method with the BCCB preconditioner C is given in [4]. There it is shown that, for certain important classes of H, the singular values of the preconditioned matrix are clustered around one, leading to rapid convergence. It also follows from the analysis in [4] that the PCGNR iterations may be ineffective or even break down if H is very ill-conditioned. In the remaining sections of the paper, our purpose is to show how the iterations can be effectively and efficiently regularized for solving ill-posed problems by using the FFT-based spectral decomposition of C.

3 Regularization and the SVD

Let

$$H = U\ \Sigma\ V^T$$

be the singular value decomposition (SVD) of H, (cf. [7]) where the columns $\mathbf{u}_1, \ldots, \mathbf{u}_N$ of U and the columns $\mathbf{v}_1, \ldots, \mathbf{v}_N$ of V form orthonormal bases of $\mathbb{R}^N$, respectively, and Σ is a diagonal matrix containing the singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_N > 0$ (we assume that H is nonsingular).

In applications arising from integral equations of the first kind, σ_1 is a moderately sized number – more or less independent of the discretization – and the ill-conditioning of H stems from clustering of the smaller singular values at the origin. In convolution problems over $\mathbb{R}^N$, such as those we will consider, the singular values typically fill out the entire interval $(0, \sigma_1]$ with no particular gap in their spectrum.

Expanding $\mathbf{f}$ and $\boldsymbol{\eta}$ in terms of the SVD we have

$$\mathbf{f} = \sum_{i=1}^{N} f_i \mathbf{v}_i, \qquad \boldsymbol{\eta} = \sum_{i=1}^{N} \eta_i \mathbf{u}_i,$$

which gives

$$\mathbf{g} = \sum_{i=1}^{N} (\mathbf{g}^T \mathbf{u}_i)\mathbf{u}_i = \sum_{i=1}^{N} (\sigma_i f_i + \eta_i)\mathbf{u}_i.$$

It is not our aim to rigorously treat the theory of regularization, for this we rather refer to the forthcoming survey [10]. Instead, we prefer to derive an understanding of the preconditioning issues in CGNR-based regularization. With this in mind, a heuristic treatment of regularization seems to be more appropriate.

Our discussion is based on the following two main assumptions:

- The spectral coefficients $\bar{\mathbf{g}}^T \mathbf{u}_i = \sigma_i f_i$ of the unperturbed right–hand side $\bar{\mathbf{g}} = H\mathbf{f}$ decay in absolute value like the σ_i; indeed, we always have

  $$|\bar{\mathbf{g}}^T u_i| = \sigma_i |f_i| \leq \sigma_i ||\mathbf{f}||, \quad i = 1, 2, \ldots.$$

  This assumption reflects the *Picard condition* of the continuous problem (*cf.* [8]) which is even somewhat stronger. Following Hansen [11, Def. 3.3] we call this the *discrete Picard condition.*

- $\boldsymbol{\eta}$ originates from "Gaussian white noise"; that means its spectral components η_i are independent stochastic variables with mean zero and variance $\varepsilon^2 > 0$. We may therefore expect that

  $$|\eta_i| \approx \varepsilon, \qquad i = 1, 2, \ldots.$$

For practical reasons we will further on assume that the signal $\bar{\mathbf{g}}$ dominates the noise $\boldsymbol{\eta}$; no reasonable reconstructions are possible otherwise.

We separate the indices $i \in \{1, \ldots, N\}$ into three subsets: $\mathcal{I}_s$, corresponding to the *signal subspace* $\mathcal{U}_s := \text{span}\{\mathbf{u}_i : \varepsilon \ll \sigma_i\}$; $\mathcal{I}_n$, corresponding to the *noise subspace* $\mathcal{U}_n := \text{span}\{\mathbf{u}_i : \sigma \ll \varepsilon\}$; and the remaining ones $\mathcal{I}_t$, defining $\mathcal{U}_t := \text{span}\{\mathbf{u}_i : i \in \mathcal{I}_t\}$ as the *transition subspace*, where σ_i and ε have the same order of magnitude. While the first two sets constitute a highly simplifying distinction between "good" and "bad" singular values, the idea of $\mathcal{I}_t$ associated with the transition subspace is to cope with the inherent shortcomings of such a procedure.

Reconstructing naively, *i.e.*, without regularization, yields

$$H^{-1}\mathbf{g} = \sum_{i=1}^{N} \frac{\mathbf{g}^T\mathbf{v}_i}{\sigma_i}\,\mathbf{v}_i = \sum_{i=1}^{n} \frac{\sigma_i f_i + \eta_i}{\sigma_i}\,\mathbf{v}_i. \tag{6}$$

For the small indices $i \in \mathcal{I}_s$ there is a large number c with $\sigma_i > c\varepsilon$, and

$$\frac{\sigma_i f_i + \eta_i}{\sigma_i} = f_i + \frac{\eta_i}{\sigma_i} \approx f_i \pm \frac{1}{c}, \quad i \in \mathcal{I}_s,$$

where $1/c$ is small; these spectral components are indeed well reconstructed. For the larger indices $i \in \mathcal{I}_n$ we have instead $\sigma_i \ll \varepsilon$, and

$$\frac{\sigma_i f_i + \eta_i}{\sigma_i} = f_i + \frac{1}{\sigma_i}\,\eta_i, \quad i \in \mathcal{I}_n,$$

so that the noise contribution η_i is highly magnified.
We conclude that $H^{-1}\mathbf{g}$ is dominated by noise, and regularization has to deal with that, *e.g.*, by incorporating filter factors $\varphi_i \geq 0$ into the expansion (6):

$$\mathbf{f}_{\{\varphi_i\}} = \sum_{i=1}^{N} \varphi_i\,\frac{\sigma_i f_i + \eta_i}{\sigma_i}\,\mathbf{v}_i. \tag{7}$$

The filter factors φ_i should satisfy the following requirement: they should be close to 1 for $i \in \mathcal{I}_s$ – where stable reconstruction of f_i is possible – and they should be close to 0 for the critical $i \in \mathcal{I}_n$. It remains to discuss the impact of $i \in \mathcal{I}_t$. The corresponding φ_i should lie between 0 and 1, and the essential task of any regularizing algorithm is to choose these filter factors φ_i, $i \in \mathcal{I}_t$, in such a way that the approximation error

$$||\mathbf{f} - \mathbf{f}_{\{\varphi_i\}}||$$

is minimized. Note that optimal choice of these filter factors *a priori* is difficult (if not impossible) as it depends on the actual ratios f_i/η_i, $i \in \mathcal{I}_t$, which are unknown of course. Good *a posteriori* choices are necessary for a regularizing algorithm to be effective.

We emphasize that in image restoration applications, a similar notion of signal space, noise space and transition third space may be introduced in the Fourier domain: the blurred image $\bar{\mathbf{g}}$ is often close to being bandlimited, *i.e.* its Fourier coefficients are small outside a symmetric interval $\mathcal{F}_s = \{\omega : 0 \leq |\omega| \leq \omega_s\}$. The expected Fourier coefficients of the white noise $\boldsymbol{\eta}$, however, have the same magnitude for all frequencies, ε say, and they will dominate the Fourier coefficients of $\mathbf{g} = \bar{\mathbf{g}} + \boldsymbol{\eta}$ in $\mathcal{F}_n = \{\omega : |\omega| \geq \omega_n > \omega_s\}$. Finally, $\mathcal{F}_t = \{\omega : \omega_s < |\omega| < \omega_n\}$ goes between $\mathcal{F}_s$ and $\mathcal{F}_n$.

4 Regularizing Properties of PCGNR

We continue our heuristic investigation of regularization in the special case of (P)CGNR, see also [10, 17]. More rigorous but less illustrative analyses can be found in Nemirovskii [14] and Plato [15].

We first consider the case $C = I$ (CGNR). Observe that the k^{th} iterate $\mathbf{f}^{(k)}$ of Algorithm 1 assumes indeed a filtered SVD–expansion of form (7), *i.e.*

$$\mathbf{f}^{(k)} = \sum_{i=1}^{N} \varphi_i^{(k)} \frac{\sigma_i f_i + \eta_i}{\sigma_i} \mathbf{v}_i,$$

where the filter factors $\varphi_i^{(k)}$ are given by the so-called *Ritz polynomials* r_k via

$$\varphi_i^{(k)} = 1 - r_k(\sigma_i^2). \tag{8}$$

The Ritz polynomials in turn may alternatively be introduced as a certain sequence of orthogonal polynomials (*cf.* [17]), or as the unique sequence of polynomials minimizing the quadratic form

$$F[p_k] = \sum_{i=1}^{N} p_k^2(\sigma_i^2)\ (\sigma_i f_i + \eta_i)^2 \tag{9}$$

amongst all polynomials p_k of degree k, normalized so that

$$p_k(0) = 1, \qquad k = 0, 1, 2, \ldots$$

.

We find the second characterization (9) more convenient. Rewriting the minimizer $F(r_k)$ in terms of the filter factors (8) yields

$$F[r_k] = \min_{\substack{\deg p_k = k, \\ p_k(0)=1}} F[p_k] = \sum_{i=1}^{N} (1 - \varphi_i^{(k)})^2 (\sigma_i f_i + \eta_i)^2. \tag{10}$$

Consider now the "weights" of the quadratic form F,

$$w_i = (\sigma_i f_i + \eta_i)^2.$$

According to the discrete Picard condition for $\bar{\mathbf{g}}$ and the white noise property of $\boldsymbol{\eta}$, we observe that the weights w_i will be heavier for $i \in \mathcal{I}_s$; that is, for spectral components corresponding to the signal subspace. The corresponding filter factors will therefore be the first to approach 1 as k increases.

On the other hand, $r_k(0) = 1$, and the weights w_i with $i \in \mathcal{I}_n$ are just of the order of ε. Thus, $r_k(\lambda)$ will remain close to 1 in an entire neighborhood of $\lambda = 0$ for moderate values of k, and the filter factors $\varphi_i^{(k)}$ for $i \in \mathcal{I}_n$ will therefore be close to 0 in the beginning of the iteration process.

Consequently, the $\varphi_i^{(k)}$ fulfill the filtering requirements that we have imposed for (7) – at least, up to some iteration number k': at that point of the iteration CGNR has sufficient degrees of freedom when choosing the Ritz polynomials to cancel the contributions of $i \in \mathcal{I}_s$ to $F[r_k]$ in (10), and the contributions of $i \in \mathcal{I}_n$ start getting important in the minimization process (10) (a good way to think of these effects is in terms of an interpolation process; see for instance the analysis of Louis [13]). At this stage of the iteration, $r_k(\sigma_i^2)$, $i \in \mathcal{I}_n$ fixed, will significantly decrease and CGNR starts to reconstruct noise; the filtering requirements are violated and CGNR should ultimatively be stopped.

Again, we emphasize the role of $\mathcal{U}_t$: CGNR should not be stopped before the entire information from the signal subspace is reconstructed, but it must be stopped before it starts reconstructing the noise space components. When to stop exactly depends on the amount of signal information relative to noise that is contained in the transition subspace $\mathcal{U}_t$.

We now turn to the question of *preconditioning.* The standard idea of preconditioning is to speed up the convergence process; in the present context this should rather mean to reduce the number of iterations required to reconstruct the information from the signal subspace. Depending on the robustness of our stopping rule, we may or may not wish to accelerate the convergence process as it reconstructs information from the transition subspace.

The number of iterations required to reconstruct the component of $\mathbf{f}$ in $\mathcal{U}_s$ mainly depends on the spread and the clustering of the singular values σ_i with $i \in \mathcal{I}_s$; the more clustered the singular values, the faster the convergence. As mentioned before, discrete convolution equations usually have densely distributed wide-spread spectra with little clustering, hence CGNR will require many iterations before the stopping rule is satisfied.

The preconditioner C should therefore cluster the spectrum of the mapping HC^{-1} restricted to the signal subspace U_s. On the other hand, the entire spectrum of HC^{-1} must not be clustered because this would combine the signal subspace and noise subspace; in such a case all iterates, right from the beginning, would be contaminated by noise, and the regularizing property of CGNR would be lost. Rather, HC^{-1} should act like H on the noise subspace, or, in other words, C should act like the identity on $\mathcal{U}_n$.

The remaining question about the behavior of C on the transition subspace $\mathcal{U}_t$ is delicate. It seems to us as if $C \approx I$ on $\mathcal{U}_t$ is the most appropriate for our purposes and applications: Due to the non-clustered part $\{\sigma_i : i \in \mathcal{I}_t\}$ of the spectrum of H, PCGNR will reach and pass the optimal stopping point quite slowly so that termination rules will work robustly. Of course, however, if the signal subspace $\mathcal{U}_s$ has been underestimated, then potential speedup will be lost with such a choice of C; we refer to the experimental section for examples.

We emphasize that the choice of C should depend on the actual noise level, because the definitions of $\mathcal{U}_s$, $\mathcal{U}_t$ and $\mathcal{U}_n$ depend on the noise variance ε^2. For smaller noise levels we can allow "more preconditioning" in the sense of "more clustering"; for large noise levels we can tolerate "little preconditioning" only.

The standard Level–2 preconditioner of Section 2 would correspond to "full preconditioning", giving as much clustering as possible with C. According to what we have said before, this choice of preconditioner is totally out of order for ill–posed problems because this would mix up the noise and signal subspaces.

In the next section we will show how the Level–2 preconditioner can be adapted to manipulate the level of preconditioning that we actually want. We also prove a result which essentially states that the singular values of the preconditioned matrix corresponding to the signal subspace cluster at 1, while the singular values corresponding to the noise subspace still cluster at 0.

5 Signal Subspace and the Truncation Parameter

As has been discussed in the previous section, a critical point of our preconditioning approach is the specification of the – so far only vaguely defined – signal subspace $\mathcal{U}_s$. Roughly speaking, the signal subspace should contain a major part of that right-hand side's fraction which comes from $\bar{\mathbf{g}}$, rather than from $\boldsymbol{\eta}$. Once $\mathcal{U}_s$ is specified, we use it to choose a *truncation parameter* τ. This parameter τ is chosen so that the eigenvalues of C satisfying $|\lambda_i| \geq \tau$ correspond to $\mathcal{U}_s$, with the remaining eigenvalues corresponding to $\mathcal{U}_n$ and $\mathcal{U}_t$. This truncation parameter is then used to construct our modified preconditioner, which we denote as C_τ, from C. We will postpone the specific description of C_τ until the next section, while concentrating here on how the signal subspace $\mathcal{U}_s$ is specified and how, from it, the truncation parameter τ is chosen.

As the discussion in Section 3 implies, a convenient subspace might be defined with Fourier vectors. At this point we recall the final remark of Section 3: it suggests to compute the discrete Fourier transform of the right-hand side vector $\mathbf{g}$; its Fourier coefficients will decay to about the level of ε as the frequency increases (cf. Figures 3 and 5 for an example).

We suggest to choose ω_s as that particular frequency for which all Fourier coefficients relative to frequencies $|\omega| > \omega_s$ are less than ε in absolute value. As in Section 3, the signal space $\mathcal{U}_s$ is taken to be that space which is spanned by the Fourier vectors with frequencies of at most ω_s.

To compute our truncation parameter, τ, we thus compute the discrete Fourier transform of the right hand side $\mathbf{g}$, and find the point ε at which these coefficients level off. The index where this stagnation first begins indicates where the random errors start to dominate the right hand side. That is, we obtain a specification of the signal subspace. Thus, if p_τ is the index where the Fourier coefficients begin to level off, we then take the magnitude of the p_τ^{th} eigenvalue of C, $|\lambda_{p_\tau}|$, as the truncation parameter τ.

6 Modified Level–2 Preconditioner

In this section we describe how the Level–2 preconditioner can be modified to incorporate the amount of preconditioning we want; and, in addition, we provide some theoretical results on the effectiveness of the preconditioner. It was suggested in Section 4 how $C \equiv \mathcal{C}_2(H)$ is to be modified, but it is useful to state an algorithm which explicitly shows how the modified Level–2 preconditioner is to be constructed.

First, to obtain an explicit description of our preconditioner, recall that C is a BCCB matrix, and hence can be easily diagonalized using the 2-D FFT. In fact, when using C as a preconditioner, one needs only its eigenvalues (see Algorithm 2.) Now recall further from Section 4, that for ill-posed problems we want the modified preconditioner to act like C on $\mathcal{U}_s$, and like the identity on $\mathcal{U}_n$ and on $\mathcal{U}_t$.

This is equivalent to locating the eigenvalues of C which correspond to $\mathcal{U}_n$ and to $\mathcal{U}_t$ and replacing them with ones.

As discussed in Section 5, we choose a truncation parameter τ such that eigenvalues of C satisfying $|\lambda_i| \geq \tau$ correspond to $\mathcal{U}_s$, with the remaining eigenvalues corresponding to $\mathcal{U}_n$ and $\mathcal{U}_t$. To simplify our discussion, we assume that the eigenvalues of C are λ_i with $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{mn}|$, and we define the *truncation index* p_τ to be the largest integer p such that $|\lambda_p| \geq \tau$. Moreover, since our preconditioner depends on the parameter τ, we denote it by C_τ.

Now, as the above remarks suggest, given the tolerance τ, to construct C_τ we first find the eigenvalues λ_i of C. Then C_τ is simply the BCCB matrix with eigenvalues

$$\tilde{\lambda}_i = \begin{cases} \lambda_i & \text{if } |\lambda_i| \geq \tau \\ 1 & \text{if } |\lambda_i| < \tau \end{cases} .$$

This discussion is summarized in the following algorithm.

Algorithm 3. Construction of the modified **Level–2** preconditioner for 2-D problems. *Let H be the $BTTB(m,n)$ matrix given as in* (2), *and let τ be chosen by the method described in Section* 5. *This algorithm computes the eigenvalues of the BCCB preconditioner C_τ, using the 2-D FFT, and stores them in the diagonal matrix Λ.*

Use Algorithm 2 to compute the eigenvalues λ_i of C.

If $|\lambda_i| < \tau$, replace λ_i with 1.

The cost of finding C_τ, and for solving systems with C_τ, is on the same order as that for C (see remarks following Algorithm 2.)

We now turn to the convergence analysis of the PCGNR method with C_τ used as a preconditioner. As mentioned earlier, the convergence rate of the PCGNR depends on the distribution of the singular values of HC_τ^{-1}. In particular, if the singular values of HC_τ^{-1} corresponding to the signal subspace $\mathcal{U}_s$ cluster around one, then the convergence will be rapid.

Here, as in [4], we assume that the entries $h_{i-j,k-l}$ of H are obtained from a doubly indexed sequence $\{h_{\mu,\nu}\}$, and that

$$\sum_{\mu=-\infty}^{\infty} \sum_{\nu=-\infty}^{\infty} |h_{\mu,\nu}| \leq K < \infty. \tag{11}$$

Thus we immediately have

$$||H||_2^2 \leq ||H||_1 ||H||_\infty \leq K^2, \tag{12}$$

where, as usual, $||H||_1$ and $||H||_\infty$ denote the column and row sum norms of H, respectively.

The theoretical analysis to follow concerns the limiting process where m and n tend to infinity, while the noise variance ε^2 is assumed to be fixed. Such an analysis reflects the process of increasing the level of discretization. Note that essentially, our choice of τ only depends on ε but not on m or n.

The following lemma relating C and H is proved in [4].

Lemma 6.1. *Let H satisfy* (11). *Then for all $\delta > 0$, there exist $M, N > 0$, such that for all $m > M$ and $n > N$,*

$$H - C = U + V$$

where

$$\text{rank}(U) = O(m) + O(n)$$

and

$$||V||_2 < \delta.$$

We now consider the modified preconditioner C_τ. The following lemma establishes how well C_τ approximates H.

Lemma 6.2. *Let H satisfy* (11), *$\tau > 0$, and C_τ be constructed as above. Then for all $\delta > 0$, there exist $M, N > 0$, such that for all $m > M$ and $n > N$,*

$$H - C_\tau = U + V + W - P_\tau$$

where P_τ is the orthogonal projector onto the direct sum $\mathcal{U}_n \oplus \mathcal{U}_t$,

$$\text{rank}(U) = O(m) + O(n),$$

$$||V||_2 < \delta,$$

and W is a matrix satisfying $W = P_\tau W P_\tau$, and

$$||W||_2 < \tau.$$

Proof: For the spectral decomposition of BCCB matrices we make use of the 2–D Fourier transform matrix $\mathcal{F}$ [12], and we assume without loss of generality that the columns of $\mathcal{F}^*$ contain the eigenvectors of C in such an ordering that the corresponding eigenvalues of C are in non-increasing order. In particular, if p_τ is the truncation index corresponding to τ, the final $mn - p_\tau$ columns of $\mathcal{F}^*$ contain an orthonormal basis for $\mathcal{U}_n \oplus \mathcal{U}_t$.

Observe from Lemma 6.1 that we can write

$$H - C_\tau = H - C + C - C_\tau = U + V + W - P_\tau$$

where U and V are from Lemma 6.1, and

$$W = \mathcal{F}^*\text{diag}(0, \ldots, 0, \lambda_{p+1}, \ldots, \lambda_{mn})\mathcal{F},$$

$$P_\tau = \mathcal{F}^*\text{diag}(0, \ldots, 0, 1 \ldots, 1)\mathcal{F}.$$

From this it follows that P_τ is the orthogonal projector onto $\mathcal{U}_n \oplus \mathcal{U}_\tau$, and obviously $W = P_\tau W P_\tau$. Furthermore, as $\mathcal{F}$ is a unitary matrix,

$$||W||_2 \leq \lambda_{p+1} < \tau.$$

The properties of U and V now follow from Lemma 6.1. □

We now present our clustering result for the singular values of the preconditioned matrix HC_τ^{-1}.

Theorem 6.1. *Let H satisfy* (11), $\tau > 0$ *and* C_τ *defined as above. Then given* $\delta > 0$, *there exist* $M, N > 0$, *such that for all* $m > M$ *and* $n > N$,

$$HC_\tau^{-1} - (I - P_\tau) = U + V + W$$

where U and V satisfy

$$\mathrm{rank}(U) = O(m) + O(n)$$

and

$$||V||_2 < \delta, \qquad ||W||_2 < \tau.$$

Thus at most $O(m) + O(n)$ of the singular values of the preconditioned matrix

$$HC_\tau^{-1}$$

lie outside the two intervals $(1 - \delta - \tau, 1 + \delta + \tau) \cup [0, \delta + \tau)$.

Proof: First, notice that from Lemma 6.2 we can write

$$HC_\tau^{-1} - (I - P_\tau) = (H - C_\tau)C_\tau^{-1} + P_\tau = (\tilde{U} + \tilde{V} + \tilde{W})C_\tau^{-1} + P_\tau(I - C_\tau^{-1}),$$

where $\tilde{U}$, $\tilde{V}$ and $\tilde{W}$ are given by Lemma 6.2. By definition, $P_\tau(I - C_\tau^{-1}) = 0$, so that we obtain the identity

$$HC_\tau^{-1} - (I - P_\tau) = U + V + W$$

where we have let $U = \tilde{U}C_\tau^{-1}$, $V = \tilde{V}C_\tau^{-1}$ and $W = \tilde{W}C_\tau^{-1}$. Note that the particular form of $\tilde{W}$ derived in Lemma 6.2 implies that $W = \tilde{W}$ and hence W satisfies the given assertion. Furthermore,

$$\mathrm{rank}(U) = \mathrm{rank}(\tilde{U}) \leq O(m) + O(n).$$

Now observe that, from the definition of C_τ, its minimum eigenvalue is bounded by $\min\{1, \tau\}$. Thus, assuming $\tau \leq 1$ and using Lemma 6.2, we obtain

$$||V||_2 \leq ||\tilde{V}||_2 ||C_\tau^{-1}||_2 \leq \tilde{\delta}/\tau.$$

Now since the $\tilde{\delta}$ can be chosen arbitrarily small, it follows that $||V||_2 < \delta$.

Then, as $I - P_\tau$ is an orthogonal projector with all its singular values being either 0 or 1, it follows from [19] that at most $O(m) + O(n)$ singular values of HC_τ^{-1} lie outside the union of the intervals $(1 - \delta - \tau, 1 + \delta + \tau) \cup [0, \delta + \tau)$. □

We note that using a standard error analysis of the conjugate gradient method, it can be shown (see, *e.g.*, [4]) that our method here will give the best regularized approximation to $\mathbf{f}$ in at most $O(m) + O(n)$ steps for m and n sufficiently large. Also it follows from Theorem 6.1 that $HC_\tau^{-1} \approx I$ on $\mathcal{U}_s$ whereas $HC_\tau^{-1} \approx H \approx 0$ on $\mathcal{U}_n$; in other words, C_τ is a good preconditioner on the signal subspace and acts like the identity on the noise subspace, as desired. We also emphasize that for moderately ill-posed problems with little noise, where we can choose $\tau \approx 0$ we can make use of "full preconditioning" and then we obtain convergence in at most $O(m) + O(n)$ steps.

7 Numerical Results for 2-D Problems

In this section we provide some numerical experiments illustrating the effectiveness of our methods. The examples we use are model problems in image restoration computations [12], and are of the form

$$\mathbf{g} = H\mathbf{f} + \boldsymbol{\eta},$$

where H is a BTTB matrix representing the blurring in the image. All computations were performed on the Cray Y-MP at the North Carolina Supercomputing Center.

We begin by generating a known 64×64 image shown in Figure 1, and form the vector $\mathbf{f}$ of dimension 4096 by row ordering (see [12]) the original image. We then choose the ill-conditioned 4096×4096 blurring matrix H as

$$h_{i-j,k-l} = \begin{cases} \exp\{-0.1((i-j)^2 + (k-l)^2)\} & -5 \leq i-j, k-l \leq 5 \\ 0 & \text{otherwise} \end{cases}, \tag{13}$$

and construct the blurred, noisy image by forming the vector $\mathbf{g} = H\mathbf{f} + \boldsymbol{\eta}$ (see Figure 2). The entries of the vector $\boldsymbol{\eta}$ are generated randomly from a normal distribution with a mean of 0.0 and a variance of $\varepsilon^2 = 1.0$, and scaled to obtain specific noise levels. Two noise levels are considered, $\|\boldsymbol{\eta}\|_2/\|H\mathbf{f}\|_2 = 10^{-3}$ and 10^{-2}, *i.e.* 0.1% and 1% noise.

To choose the truncation parameter τ, we use the scheme suggested in Section 5. Specifically, we compute the discrete Fourier transform of the right hand side $\mathbf{g}$, and find the level ε at which these coefficients level off. The index where this stagnation first begins indicates where the random errors start to dominate the right hand side. That is, we obtain a specification of the signal subspace. Thus, if i is the index where the Fourier coefficients begin to level off, we then take the magnitude of the p^{th} eigenvalue of C, $|\lambda_p|$, as the truncation parameter τ used in Section 6 for constructing the preconditioner C_τ.

Example 1: 0.1% noise.

In this first example, we consider 0.1% noise. Figure 3 shows a plot of the eigenvalues of C along with a plot of the Fourier coefficients of the right hand side. Here we see that the Fourier coefficients begin to stagnate when $p \approx 1450$. Thus, we take $|\lambda_{1450}| \approx 0.0032$ as the truncation parameter τ.

In Figure 4, we show the relative errors for several iterations of PCGNR with our preconditioner and with no preconditioner. The minimal relative error for PCGNR, with our preconditioner, is attained at $k = 30$ iterations, with a relative error value of ≈ 0.1698. On the other hand, when no preconditioner is used, then the minimal relative error occurs at $k = 189$ iterations, with a relative error value of ≈ 0.1654. These results are summarized in Table 1. Note from Figure 4 that the convergence rate of PCGNR, with our preconditioner, is significantly faster than the CGNR with no preconditioning. Each method produces a similar minimal relative error.

	0.1% noise		1% noise	
	no prec.	prec.	no prec.	prec.
iterations	189	30	43	6
min. error	0.1654	0.1698	0.2236	0.2276

Table 1: Summary of numerical results from Examples 1 and 2.

Example 2: 1% noise.

In this example we generate 1% noise. Again, to choose the truncation parameter, we consider the plots of the eigenvalues of C and the corresponding Fourier coefficients of the right hand side, shown in Figure 5. From these plots, we see that the Fourier coefficients begin to level off when $p \approx 500$.

Using $|\lambda_{500}| \approx 0.0233$ as the truncation parameter τ, the relative errors for several iterations of PCGNR with our preconditioner, and with no preconditioner, are shown in Figure 6. In this case, the minimal relative error computed by PCGNR occurs at $k = 6$ iterations with a relative error value of ≈ 0.2276, whereas the minimum relative error for CGNR occurs at $k = 43$ iterations with a relative error value of ≈ 0.2236. These results are also summarized in Table 1.

Choosing an "optimal" truncation parameter is critical to the effectiveness of our preconditioners. We illustrate this in the following two examples, where we use the data from Example 2 with 1% noise.

Example 3: Tolerance too large.

In this example, we choose the truncation parameter to be $\tau = |\lambda_{250}| \approx 0.1539$. This has the effect of underestimating the signal subspace (see Figure 5), and, as remarked in Section 4, we expect the potential speedup of the preconditioner to be lost. This is easily seen in Figure 7, where we show the relative error for each iteration of the PCGNR with this preconditioner (dashed line) as well as with no preconditioner (solid line). In this case, the minimal relative error is reached at $k = 41$ iterations, with a value of 0.2238. This should be compared to Figure 6 where the optimal truncation parameter was used.

Example 4: Tolerance too small.

In this example we take the truncation parameter to be $|\lambda_{1000}| \approx 0.0077$. In this case we have, effectively, overestimated the signal subspace (see Figure 5). Thus we would expect the noise to dominate, and hence we will not be able to obtain as accurate a solution as in Example 2. This can also be seen in Figure 7, where the dash-dot line is a plot of the relative errors for each iteration of the PCGNR with the preconditioner of this example. Here we reach the minimum relative error at iteration 9, with a large relative error value of 0.3005. Clearly, the optimal truncation parameter in Example 2 produces a much more accurate solution with fewer iterations.

The restored image for Example 1 using our algorithm is given in Figure 8. Our examples illustrate advantages, as well as some problems, with choosing a good truncation parameter for our preconditioner. We see that if the tolerance is chosen well, then we get a significant speed-up in convergence to a minimal relative error solution. But the above examples also indicate, as was eluded to in Section 6, that the relative error of the approximate solution obtained from the PCGNR iterations can be very sensitive to the choice of the truncation parameter due to the ill-conditioned nature of C_τ. In particular, if the truncation parameter τ is underestimated then we may not be able to compute an accurate approximation to the exact solution. This indicates that one should choose $C \approx I$ on $\mathcal{U}_t$, as was suggested in Section 6. Moreover, if it is not clear where the signal subspace ends, then it appears to be best to overestimate the truncation tolerance τ, rather than underestimate it.

Finally, we remark that we have not addressed the very important, and difficult, problem of deciding when to stop the PCGNR iterations. There has recently been much research devoted to this topic, and we refer the interested reader to the discussion in [10]. We emphasize that if one has a good stopping criteria for the PCGNR, then our approach to preconditioning may offer significant improvement on convergence rate for ill-posed problems of the type discussed in this paper.

Acknowledgement. Much of this work was done while the authors were participating in the *Year on Applied Linear Algebra* at the Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN 55455.

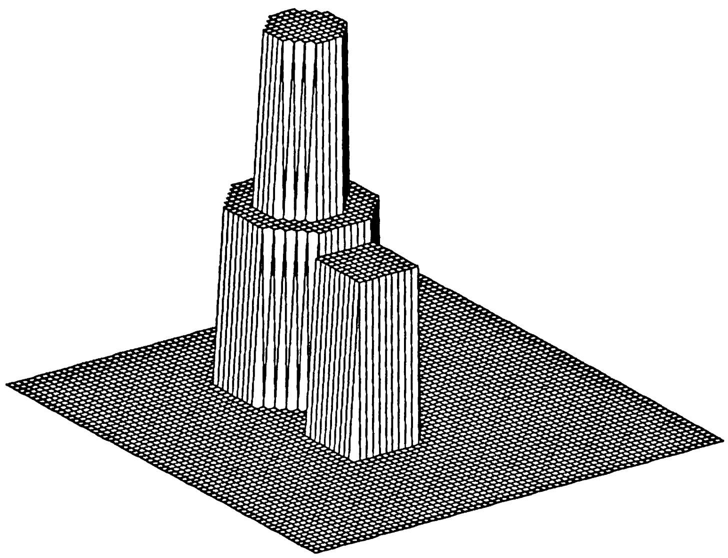

Figure 1: Original image.

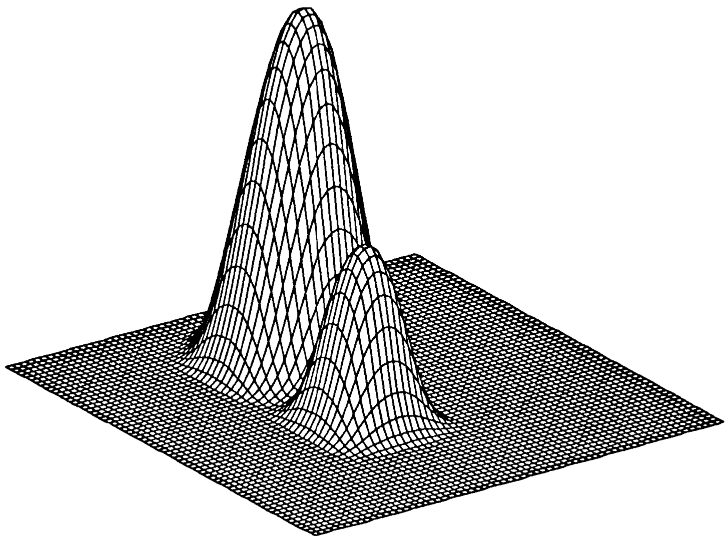

Figure 2: Blurred image with 0.1% noise.

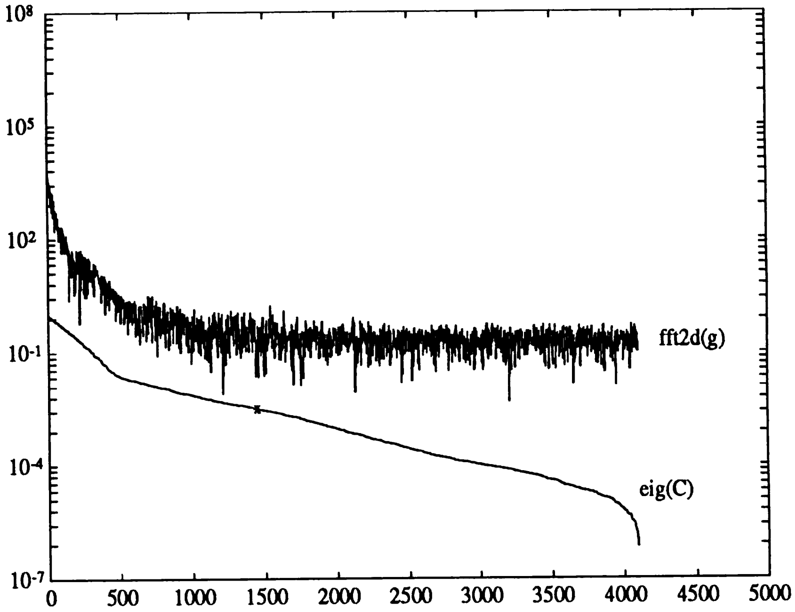

Figure 3: Eigenvalues of C and Fourier transform of g for 0.1% noise.

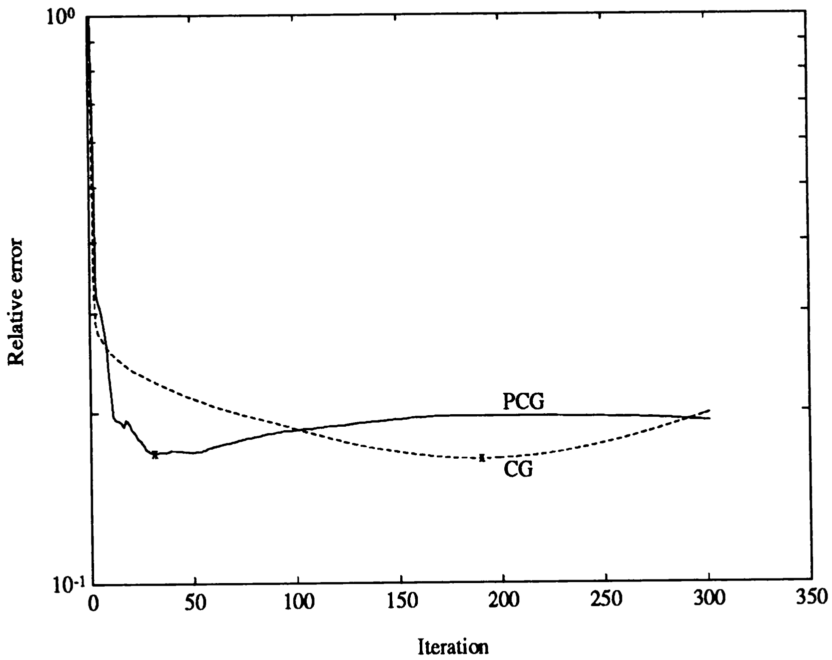

Figure 4: Errors for 0.1% noise.

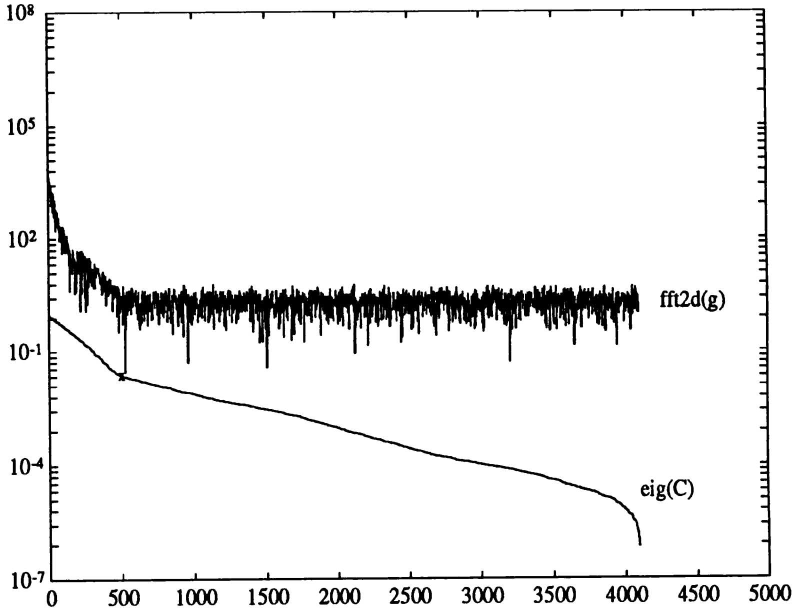

Figure 5: Eigenvalues of C and Fourier transform of $\mathbf{g}$ for 1% noise.

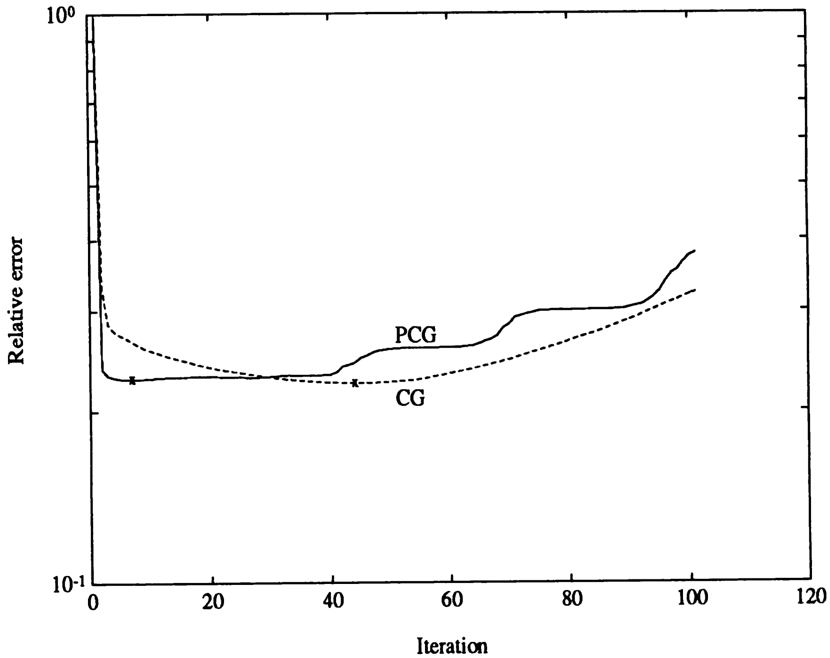

Figure 6: Errors for 1% noise.

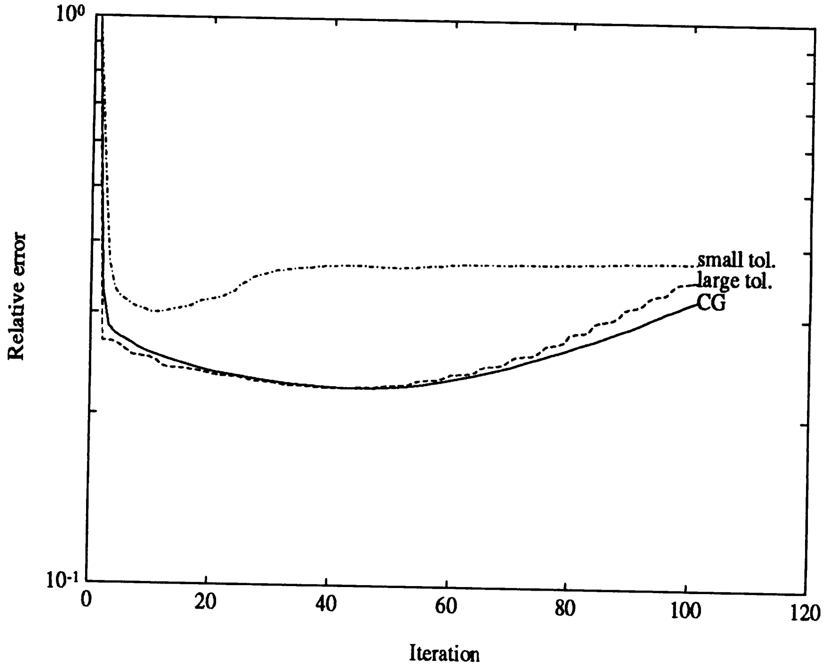

Figure 7: Results when tolerance τ is chosen too large or too small.

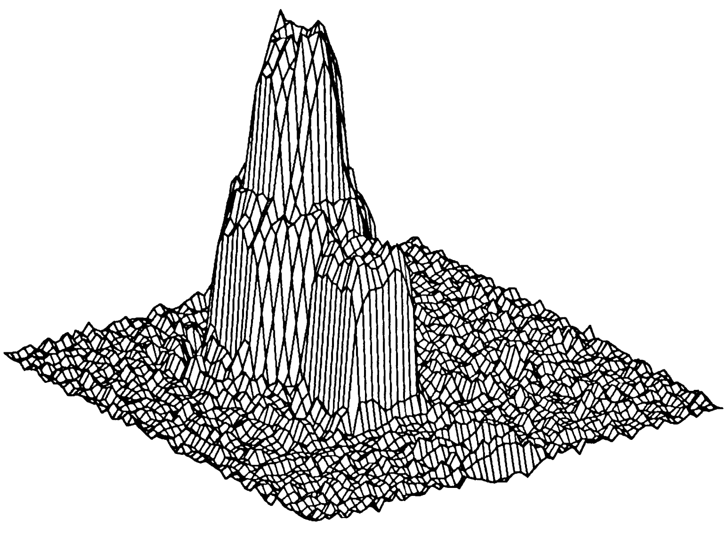

Figure 8: Restored image in Example 1.

References

[1] A. Bjorck, *Least Squares Methods*, in Handbook of Numerical Methods, ed. P. Ciarlet and J. Lions, Elsevier/North Holland Vol. 1, 1989.

[2] R. Chan, X. Jin, *A Family of Block Preconditioners for Block Systems*, SIAM J. Sci. Statist. Comput., to appear.

[3] R. Chan, J. Nagy and R. Plemmons, *Circulant Preconditioned Toeplitz Least Squares Iterations*, preprint (1991), SIAM J. Matrix Anal. Appl., to appear.

[4] R. Chan, J. Nagy and R. Plemmons, *FFT-Based Preconditioners for Toeplitz-Block Least Squares Problems*, preprint (1992), SIAM J. Numerical Anal., to appear.

[5] T. Chan, *An Optimal Circulant Preconditioner for Toeplitz Systems*, SIAM J. Sci. Statist. Comput., V9 (1988), pp. 766–771.

[6] T. Chan and J. Olkin, *Preconditioners for Toeplitz-block matrices*, preprint, (1992).

[7] G. Golub and C. Van Loan, *Matrix Computations*, Johns Hopkins Univ. Press, Baltimore, MD, 2nd Ed., 1989.

[8] C. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman Publishing, Boston, 1984.

[9] M. Hanke, *Accelerated Landweber Iterations for the Solution of Ill-posed Problems*, Numer. Math., V60 (1991), pp. 341–373.

[10] M. Hanke and P. C. Hansen, *Regularization Methods for Large-Scale Problems.* Technical University of Denmark, UNI•C Report UNIC-92-04, August 1992.

[11] P. C. Hansen, *Truncated SVD Solutions to Discrete Ill-Posed Problems with Ill-Determined Numerical Rank*, SIAM J. Sci. Stat.. Comp., V11 (1990), pp. 503-518.

[12] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Engelwood Cliffs, NJ 1989.

[13] A. K. Louis, *Convergence of the Conjugate Gradient Method for Compact Operators*, in Inverse and Ill-Posed Problems, H. W. Engl and C. W. Groetsch (Eds.), Academic Press, Boston, New York, London, 1978, pp. 177–183.

[14] A. S. Nemirovskii, *The Regularizing Properties of the Adjoint Gradient Method in Ill-Posed Problems*, USSR Comput. Math. Math. Phys., V26,2 (1986), pp. 7–16.

[15] R. Plato, *Optimal Algorithms for Linear Ill-Posed Problems Yield Regularization Methods*, Numer. Funct. Anal. Optim., V11 (1990), pp. 111–118.

[16] G. Strang, *A Proposal for Toeplitz Matrix Calculations*, Stud. Appl. Math., V74 (1986), pp. 171–176.

[17] A. Van der Sluis and H. Van der Vorst, *SIRT- & CG-Type Methods for the Iterative Solution of Least Squares Problems*, Lin. Alg. Appl., V130 (1990), pp. 257–302.

[18] A. Van der Sluis and H. A. Van der Vorst, *The Rate of Convergence of Conjugate Gradients*, Numer. Math., V48 (1986), pp. 543–560.

[19] J. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.

Recent IMA Preprints

#	Author/s	Title
960	**Richard A. Brualdi, Shmuel Friedland and Alex Pothen**, Sparse bases, elementary vectors and nonzero minors of compound matrices	
961	**J.W. Demmel**, Open problems in numerical linear algebra	
962	**James W. Demmel and William Gragg**, On computing accurate singular values and eigenvalues of acyclic matrices	
963	**James W. Demmel**, The inherent inaccuracy of implicit tridiagonal QR	
964	**J.J.L. Velázquez**, Estimates on the $(N-1)$-dimensional Hausdorff measure of the blow-up set for a semilinear heat equation	
965	**David C. Dobson**, Optimal design of periodic antireflective structures for the Helmholtz equation	
966	**C.J. van Duijn and Joseph D. Fehribach**, Analysis of planar model for the molten carbonate fuel cell	
967	**Yongzhi Xu, T. Craig Poling and Trent Brundage**, Source localization in a waveguide with unknown large inclusions	
968	**J.J.L. Velázquez**, Higher dimensional blow up for semilinear parabolic equations	
969	**E.G. Kalnins and Willard Miller, Jr.**, Separable coordinates, integrability and the Niven equations	
970	**John M. Chadam and Hong-Ming Yin**, A diffusion equation with localized chemical reactions	
971	**A. Greenbaum and L. Gurvits**, Max-min properties of matrix factor norms	
972	**Bei Hu**, A free boundary problem arising in smoulder combustion	
973	**C.M. Elliott and A.M. Stuart**, The global dynamics of discrete semilinear parabolic equations	
974	**Avner Friedman and Jianhua Zhang**, Swelling of a rubber ball in the presence of good solvent	
975	**Avner Friedman and Juan J.L. Velázquez**, A time-dependence free boundary problem modeling the visual image in electrophotography	
976	**Richard A. Brualdi, Hyung Chan Jung and William T. Trotter, Jr.**, On the poset of all posets on n elements	
977	**Ricardo D. Fierro and James R. Bunch**, Multicollinearity and total least squares	
978	**Adam W. Bojanczyk, James G. Nagy and Robert J. Plemmons**, Row householder transformations for rank-k Cholesky inverse modifications	
979	**Chaocheng Huang**, An age-dependent population model with nonlinear diffusion in R^n	
980	**Emad Fatemi and Farouk Odeh**, Upwind finite difference solution of Boltzmann equation applied to electron transport in semiconductor devices	
981	**Esmond G. Ng and Barry W. Peyton**, A tight and explicit representation of Q in sparse QR factorization	
982	**Robert J. Plemmons**, A proposal for FFT-based fast recursive least-squares	
983	**Anne Greenbaum and Zdenek Strakos**, Matrices that generate the same Krylov residual spaces	
984	**Alan Edelman and G.W. Stewart**, Scaling for orthogonality	
985	**G.W. Stewart**, Note on a generalized sylvester equation	
986	**G.W. Stewart**, Updating URV decompositions in parallel	
987	**Angelika Bunse-Gerstner, Volker Mehrmann and Nancy K. Nichols**, Numerical methods for the regularization of descriptor systems by output feedback	
988	**Ralph Byers and N.K. Nichols**, On the stability radius of generalized state-space systems	
989	**David C. Dobson**, Designing periodic structures with specified low frequency scattered in far-field data	
990	**C.-T. Pan and Kermit Sigmon**, A bottom-up inductive proof of the singular value decomposition	
991	**Ricardo D. Fierro and James R. Bunch**, Orthogonal projection and total least squares	
992	**Chiou-Ming Huang and Dianne P. O'Leary**, A Krylov multisplitting algorithm for solving linear systems of equations	
993	**A.C.M Ran and L. Rodman**, Factorization of matrix polynomials with symmetries	
994	**Mike Boyle**, Symbolic dynamics and matrices	
995	**A. Novick-Cohen and L.A. Peletier**, Steady states of the one-dimensional Cahn-Hilliard spaces	
996	**Zhangxin Chen**, Large-scale averaging analysis of single phase flow in fractured reservoirs	
997	**Boris Mordukhovich**, Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis	
998	**Yongzhi Xu**, CW mode structure and constraint beamforming in a waveguide with unknown large inclusions	
999	**R.P. Gilbert and Yongzhi Xu**, Acoustic waves and far-field patterns in two dimensional oceans with porous-elastic seabeds	
1000	**M.A. Herrero and J.J.L. Velázquez**, Some results on blow up for semilinear parabolic problems	
1001	**Pierre-Alain Gremaud**, Numerical analysis of a nonconvex variational problem related to solid-solid phase transitions	
1002	**Izchak Lewkowicz**, Stability robustness of state space systems inter-relations between the continuous and discrete time cases	
1003	**Kenneth R. Driessel and Wasin So**, Linear operators on matrices: Preserving spectrum and displacement structure	

1004 **Carolyn Eschenbach**, Idempotence for sign pattern matrices
1005 **Carolyn Eschenbach, Frank J. Hall and Charles R. Johnson**, Self-inverse sign patterns
1006 **Marc Moonen, Paul Van Dooren and Filiep Vanpoucke**, On the QR algorithm and updating the SVD and URV decomposition in parallel
1007 **Paul Van Dooren**, Upcoming numerical linear algebra issues in systems and control theory
1008 **Avner Friedman and Juan J.L. Velázquez**, The analysis of coating flows near the contact line
1009 **Stephen J. Kirkland and Michael Neumann**, Convexity and concavity of the Perron root and vector of Leslie matrices with applications to a population model
1010 **Stephen J. Kirkland and Bryan L. Shader**, Tournament matrices with extremal spectral properties
1011 **E.G. Kalnins, Willard Miller, Jr. and Sanchita Mukherjee**, Models of q-algebra representations: Matrix Elements of $U_q(su_2)$
1012 **Zhangxin Chen and Bernardo Cockburn**, Error estimates for a finite element method for the drift-diffusion semiconductor device equations
1013 **Chaocheng Huang**, Drying of gelatin asymptotically in photographic film
1014 **Richard E. Ewing and Hong Wang**, Eulerian-Lagrangian localized adjoint methods for reactive transport in groundwater
1015 **Bing-Yu Zhang**, Taylor series expansion for solutions of the Korteweg-de Vries equation with respect to their initial values
1016 **Kenneth R. Driessel**, Some remarks on the geometry of some surfaces of matrices associated with Toeplitz eigenproblems
1017 **C.J. Van Duijn and Peter Knabner**, Flow and reactive transport in porous media induced by well injection: Similarity solution
1018 **Wasin So**, Rank one perturbation and its application to the Laplacian spectrum of a graph
1019 **G. Baccarani, F. Odeh, A. Gnudi and D. Ventura**, A critical review of the fundamental semiconductor equations
1020 **T.R. Hoffend Jr.**, Magnetostatic interactions for certain types of stacked, cylindrically symmetric magnetic particles
1021 **IMA Summer Program for Graduate Students**, Mathematical Modeling
1022 **Wayne Barrett, Charles R. Johnson, and Pablo Tarazaga**, The real positive definite completion problem for a simple cycle
1023 **Charles A. McCarthy**, Fourth order accuracy for a cubic spline collocation method
1024 **Martin Hanke, James Nagy, and Robert Plemmons**, Preconditioned iterative regularization for I11-posed problems
1025 **John R. Gilbert, Esmond G. Ng, and Barry W. Peyton**, An efficient algorithm to compute row and column counts for sparse Cholesky factorization
1026 **Xinfu Chen**, Existence and regularity of solutions of a nonlinear nonuniformly elliptic system arising from a thermistor problem
1027 **Xinfu Chen and Weiqing Xie**, Discontinuous solutions of steady state, viscous compressible Navier-Stokes equations
1028 **E.G. Kalnins, Willard Miller, Jr., and Sanchita Mukherjee**, Models of q-algebra representations: Matrix elements of the q-oscillator algebra
1029 **W. Miller, Jr. and Lee A. Rubel**, Functional separation of variables for Laplace equations in two dimensions
1030 **I. Gohberg and I. Koltracht**, Structured condition numbers for linear matrix structures
1031 **Xinfu Chen**, Hele-Shaw problema nd area preserved curve shortening motion
1032 **Zhangxin Chen and Jim Douglas, Jr.** Modelling of compositional flow in naturally fractured reservoirs
1033 **Harald K. Wimmer**, On the existence of a least and negative-semidefinite solution of the discrete-time algebraic Riccati equation
1034 **Harald K. Wimmer**, Monotonicity and parametrization results for continuous-time algebraic Riccati equations and Riccati inequalities
1035 **Bart De Moor, Peter Van Overschee, and Geert Schelfhout**, H_2 model reduction for SISO systems
1036 **Bart De Moor**, Structured total least squares and L_2 approximation problems
1037 **Chjan Lim**, Nonexistence of Lyapunov functions and the instability of the Von Karman vortex streets
1038 **David C. Dobson and Fadil Santosa**, Resolution and stability analysis of an inverse problem in electrical impedance tomography – dependence on the input current patterns
1039 **C.N. Dawson, C.J. van Duijn, and M.F. Wheeler**, Characteristic-Galerkin methods for contaminant transport with non-equilibrium adsorption kinetics
1040 **Bing-Yu Zhang**, Analyticity of solutions of the generalized Korteweg-de Vries equation with respect to their initial values
1041 **Neerchal K. Nagaraj and Wayne A. Fuller**, Least squares estimation of the linear model with autoregressive errors
1042 **H.J. Sussman & W. Liu**, A characterization of continuous dependence of trajectories with respect to the input for control-affine systems