Mathematica Balkanica

Mathematical Society of South-Eastern Europe
A quarterly published by
the Bulgarian Academy of Sciences – National Committee for Mathematics

---

Mathematica Balkanica

New series Vol. 4, 1990, Fasc. 4

Preconditioned Iterative Solution of Spline Finite Element Systems

Victor Eijkhout[+], *Svetozar Margenov*[++]

Presented by V. Popov

In this paper we analyse the convergence behaviour of the preconditioned conjugate gradient method applied to a finite element system of equations where spline elements are used, and where the preconditioner is derived by incomplete factorization of the stiffness matrix obtained from piecewise linear basis functions. As the spline stiffness matrix can be proved to be spectrally equivalent to the linear elements one, such an iterative method will have the same asymptotic convergence speed as the iterative solution of the linear elements system using the same preconditioner. We present experimental results indicating that a very favourable proportionality constant is involved, and we give a Fourier analysis of the preconditioned system.

1. Introduction

Let us consider the numerical solution of an elliptic partial differential equation

$$\begin{aligned} \mathcal{L} &= -\nabla\cdot(k\nabla u)+qu=f, && (x,y)\in\Omega \\ u &= 0, && (x,y)\in\Gamma=\partial\Omega \end{aligned} \tag{1.1}$$

where $k\geqq k_0>0$, $q\geqq 0$, k and $q\in C(\bar{\Omega})$, and $\bar{\Omega}=[a,b]\times[c,d]$ is a rectangular domain.

We are interested in the efficient iterative solution of the linear algebraic systems that arise from application of the finite element method to the problem (1.1). In particular, we pay attention to methods that are suitable for vector and parallel computers. Many papers in recent times have considered the problem of obtaining such methods: several powerful theoretical results are obtained about variants of the domain decomposition method ([BW], [MK]), in the field of incomplete factorization methods ([AE], [AP], [EV]), and for a number of other similar methods ([AV]). These results are supported by promising numerical tests. Most of these methods are well studied when linear or bilinear elements are used, and often they are based on special properties of the matrices involved (for example M-matrix properties).

From another angle, the quadratic splines from $C^1(\Omega)$ are very suitable for approximating the solution of (1.1). Their main advantage is the high order of accuracy, combined with the relatively small number of unknowns involved. We recall the fact that quadratic splines from $C^1(\Omega)$ do not belong to the class of nodal finite element basis functions. Nevertheless, we can apply standard finite element techniques using the local basis of B-splines.

It seems an attractive idea to apply the iterative methods mentioned above in the context of quadratic splines for solving (1.1). In this paper, we show how to realise this idea using spectral equivalence of the stiffness matrices of quadratic splines and (bi-)linear elements.

Remark 1.1. The use of spectral equivalence in the case of Lagrangian nodal finite elements was proposed in [AB].

Remark 1.2. The results presented in this paper have additional meaning in connection with the application of spline macro elements when Ω is a rectangular polygon or a more complicated domain [M].

2. Quadratic splines from $C^1(\Omega)$ and their application in FEM

Let $\Delta_n=\{x_0=a<x_1<\ldots<x_n=b\}$ be a mesh on the interval $[a,b]$. We denote by $S_2(\Delta_n)$ the space of quadratic splines, where

$$S_2(\Delta_n)=\{s(x): s(x)\in P_2[x_i, x_{i+1}],\ i=0,\ldots,n-1;\ s(x)\in C^1[a,b]\}.$$

The dimension of $S_2(\Delta_n)$ is $n+2$.

We will denote by $B_{2,k}$ the quadratic B-splines, where

$$B_{2,k}(x)=3\sum_{r=k+1}^{k+3}(x_r-x)_+^2/\omega'_{k,2}(x_r) \tag{2.1}$$

where

$$(x_r-x)_+=\begin{cases}x_r-x & \text{when } x\leqq x_r\\ 0 & \text{when } x\geqq x_r\end{cases}$$

and

$$\omega_{k,2}(x)=(x-x_k)(x-x_{k+1})(x-x_{k+2})(x-x_{k+3}).$$

Now let $\bar{\Delta}_n\supset\Delta_n$ be an extension of the mesh Δ_n into the form $\bar{\Delta}_n=\{x_{-2}<x_{-1}<x_0=a<x_1<\cdots<x_n=b<x_{n+1}<x_{n+2}\}$. Under the above assumptions it is known that $\{B_{2,k}(x): k=-2,\ldots,n-1\}$ is a basis for $S_2(\Delta_n)$.

We let $\mathring{S}_2(\Delta_n)=\{s(x)\in S_2(\Delta_n): s(a)=s(b)=0\}$ and let $\{\alpha_i(x): i=1,\ldots,n\}$ be the local basis for $\mathring{S}(\Delta_n)$ defined by the formulas

$$\begin{aligned}&\alpha_1(x)=B_{-1}(x)-B_{-2}(x)\\ &\alpha_i(x)=B_{i-2}(x) \qquad\qquad i=2,\ldots,n-1\\ &\alpha_n(x)=B_{n-2}(x)-B_{n-1}(x)\end{aligned} \tag{2.2}$$

In a similar manner we define the mesh Δ_m in the interval $[c,d]$, its extension $\bar{\Delta}_m\supset\Delta_m$ and the space of B-splines $\mathring{S}_2(\Delta_m)$, equipped with a local basis denoted by $\{\beta_j(y): j=1,\ldots,m\}$.

We now apply the FEM to the solution of problem (1.1) with the element space $\mathring{S}_2=\mathring{S}_2(\Delta_n\times\Delta_m)=\mathring{S}_2(\Delta_n)\times\mathring{S}_2(\Delta_m)$. Let us denote the numerical solution of (1.1) by $u^h(x,y)\in\mathring{S}_2$. It follows that u^h can be written in the form

$$u^h(x,y)=\sum_{i=1}^{n}\sum_{j=1}^{m}\xi_{i,j}\alpha_i(x)\beta_j(y). \tag{2.3}$$

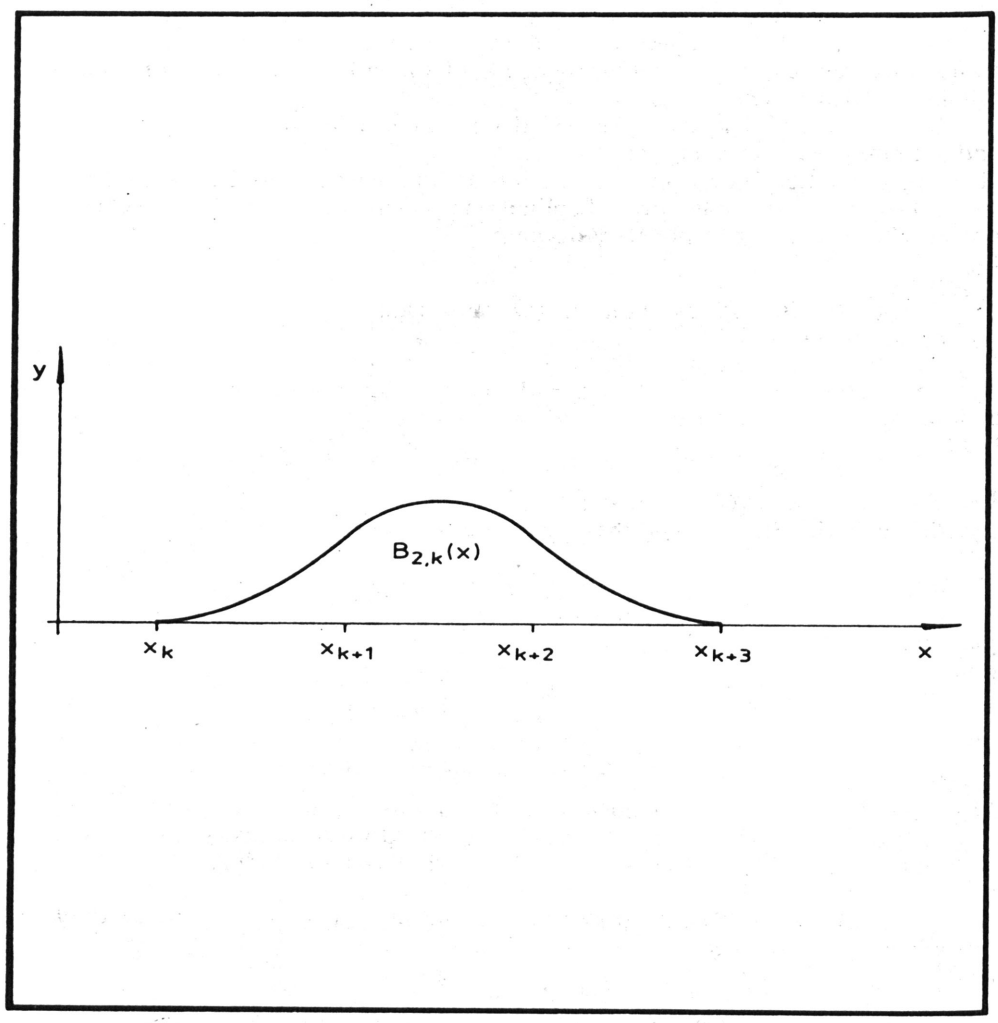

Fig. 2.1. A quadratic B-spline determined by notes x_k, x_{k+1}, x_{k+2}, x_{k+3}

The vector of unknowns ξ is the solution of the system

$$K\xi = F, \tag{2.4}$$

where K is the stiffness matrix of the problem.

We emphasize the following error estimate:

$$\|u - u^h\|_1 \leqq ch^2 \|u\|_3 \tag{2.5}$$

where $u(x, y)$ is the exact solution of (1.1) and $\|\cdot\|_s$ stands for the usual Sobolev

norm. The corresponding esimate using linear or bilinear finite elements would be $O(h)$. This means that we can obtain a more accurate solution using splines from $C^1(\Omega)$, while at the same time needing practically the same number of unknowns. An approximation based on piecewise quadratic finite elements, which is also of order h^2, would have either a double mesh width, or, when using the same h, four times as many points.

We concentrate on methods for the solution of the system (2.4). The matrix K is symmetric positive definite, and has a pentadiagonal block structure,

$$K = \text{pentadiag}(A_{i,i-2}, A_{i,i-1}, A_{i,i}, A_{i,i+1}, A_{i,i+2}), \tag{2.6}$$

where $A_{i,j}, j = i-2, \ldots, i+2$ are pentadiagonal submatrices of order m. It follows that K is a band matrix, but note that the band is relatively wide and contains a great many zeros. Taking this into account, direct methods for the solution of (2.4) should be avoided. In the next section we propose an approach to the iterative solution of (2.4) using the preconditioned conjugate gradient method.

3. Spectral equivalence of stiffness matrices

In this section we will prove the spectral equivalence of finite element matrices based on splines and based on linear or bilinear elements. From this we conclude that preconditioners for the latter type of matrices can also be used for the former type, and we derive an upper bound for the ratio of the numbers of iterations for both kinds of coefficient matrices when using a conjugate gradient method preconditioned by an incomplete factorization on the linear elements matrix.

For simplicity of presentation we consider from now on the Poisson model problem, i.e.

$$\begin{aligned} -u_{xx} - u_{yy} &= f, \quad (x, y) \in \Omega = (0,1) \times (0,1) \\ u &= 0, \qquad (x, y) \in \Gamma = \partial\Omega \end{aligned} \tag{3.1}$$

Moreover, let $\Delta = \Delta_n \times \Delta_m$, where Δ_n is a uniform mesh in $[0,1]$, i.e., $\Delta_n = \{x_i = ih : i = 0, \ldots, n;\ h = 1/n\}$. We also assume that $\bar{\Delta}_n$ is a uniform extension of Δ_n. It follows that in this case we can write the matrix K as a sum of tensor products. To this purpose we need the matrices

$$\begin{aligned} C &= \{\int_0^1 \alpha_i(x)\alpha_j(x)dx : i,j = 1, \ldots, n\} \\ B &= \{\int_0^1 \alpha_i'(x)\alpha_j'(x)dx : i,j = 1, \ldots, n\} \ . \end{aligned} \tag{3.2}$$

The system of linear algebraic equations corresponding to (3.1) is then

$$K\xi = F, \quad K = C \otimes B + B \otimes C, \tag{3.3}$$

where the sign $\otimes$ stands for the tensor product of matrices defined by the relation

$$P \otimes Q = \begin{pmatrix} p_{11}Q \cdots p_{1n}Q \\ \vdots \qquad\quad \vdots \\ p_{n1}Q \cdots p_{nn}Q \end{pmatrix}.$$

Note that the matrices C and B are symmetric positive definite and pentadiagonal.

We denote by $\bar{K}$ the stiffness matrix corresponding to (3.1) when bilinear finite elements based on $\Delta_{n+1,n+1}=\Delta_{n+1}\times\Delta_{n+1}$ are used. The uniform mesh Δ_{n+1} is defined as above

$$\Delta_{n+1}=\{x_i=ih : i=0,\dots,n+1;\ h=1/(n+1)\}.$$

The basic result in this section is the following theorem.

Theorem 3.1. *There exist positive constants* $0<d_1\leqq d_2$ *such that*

$$d_1 x^t\bar{K}x\leqq x^tKx\leqq d_2 x^t\bar{K}x \tag{3.4}$$

for every vector $x\in\mathsf{R}^{n^2}$. *Moreover,* $d\equiv d_2/d_1$ *is independent of* h.

Proof. We will use an auxiliary mesh $\tilde{\Delta}_{n,n}$, defined by

$$\tilde{\Delta}_{n,n}=\tilde{\Delta}_n\times\tilde{\Delta}_n$$

$$\tilde{\Delta}_n=\{x_0=0,\ x_1=3h/2,\ x_i=x_1+h,\ i=2,\dots,n-1,\ x_n=1;\ h=1/(n+1)\}.$$

Let us denote by I the set of the indices of the finite elements e_i corresponding to the mesh $\Delta_{n,n}$, and let the set I be split up in the following way:

$$I=I_1\cup I_2\cup I_3,\quad \forall_{i\neq j} I_i\cap I_j=\emptyset$$

$$I_1=\{i : e_i\cap\Gamma=\emptyset\}$$

$$I_2=\{i : e_i\cap\Gamma \text{ is exactly one of the sides of } e_i\}.$$

We denote by k_j the element stiffness matrices corresponding to the elements of $\Delta_{n,n}$ and to the basis $\{\alpha_j\beta_j : i,j=1,\dots,n\}$. It is easy to see that, if we use a suitable local ordering of unknowns, all element stiffness matrices corresponding to elements from the same subset I_j are identical. The dimensions of these matrices are respectively

$$\dim(k_1)=9\times 9$$

$$\dim(k_2)=6\times 6$$

$$\dim(k_3)=4\times 4.$$

Let now $\bar{k}_j$ be the element stiffness matrices corresponding to the elements of $\tilde{\Delta}_{n,n}$ and to the bilinear basis defined on $\Delta_{n+1,n+1}$. We assume that the same splitting of I and the same local ordering of the unknowns are used as above. Then $\dim(\bar{k}_j)=\dim(k_j)$ for $i=1,2,3$ and the following equalities hold

$$\begin{aligned} x^tKx&=\sum_{j=1}^{3}\ \sum_{i\in I_j} x_i^tk_jx_i\\ x^t\bar{K}x&=\sum_{j=1}^{3}\ \sum_{i\in I_j} x_i^t\bar{k}_jx_i, \end{aligned} \tag{3.5}$$

where x_i are the corresponding subvectors of x. It suffices to prove that k_j and $\bar{k}_j$ are spectrally equivalent, as (3.4) then follows immediately from (3.5).

$$k_1=\gamma_1\begin{pmatrix}
12 & 10 & -2 & 10 & -13 & -7 & -2 & -7 & -1\\
10 & 60 & 10 & -13 & -14 & -13 & -7 & -26 & -7\\
-2 & 10 & 12 & -7 & -13 & 10 & -1 & -7 & -2\\
10 & -13 & -7 & 60 & -14 & -26 & 10 & -13 & -7\\
-13 & -14 & -13 & -14 & 108 & -14 & -13 & -14 & -13\\
-7 & -13 & 10 & -26 & -14 & 60 & -7 & -13 & 10\\
-2 & -7 & -1 & 10 & -13 & -7 & 12 & 10 & -2\\
-7 & -26 & -7 & -13 & -14 & -13 & 10 & 60 & 10\\
-1 & -7 & -2 & -7 & -13 & 10 & -2 & 10 & 12
\end{pmatrix}$$

$$\bar{k}_1=\gamma_2\begin{pmatrix}
2 & 1 & 0 & 1 & -4 & & & & \\
1 & 16 & 1 & -4 & -10 & -4 & & & \\
0 & 1 & 2 & 0 & -4 & 1 & & & \\
1 & -4 & 0 & 16 & -10 & 0 & 1 & -4 & \\
-4 & -10 & -4 & -10 & 56 & -10 & -4 & -10 & -4\\
 & -4 & 1 & 0 & -10 & 16 & 0 & -4 & 1\\
 & & & 1 & -4 & 0 & 2 & 1 & 0\\
 & & & -4 & -10 & -4 & 1 & 16 & 1\\
 & & & & -4 & 1 & 0 & 1 & 2
\end{pmatrix}$$

$$k_3=\gamma_1\begin{pmatrix}
204 & 36 & 36 & 0\\
36 & 52 & 0 & 12\\
36 & 0 & 52 & 12\\
0 & 12 & 12 & 12
\end{pmatrix},\qquad \bar{k}_3=\gamma_2\begin{pmatrix}
54 & -3 & -3 & -4\\
-3 & 12 & -4 & 1\\
-3 & -4 & 12 & 1\\
-4 & 1 & 1 & 2
\end{pmatrix}.$$

As we would like to obtain explicit values for the spectral equivalence bounds d_1 and d_2 we consider the stiffness element matrices k_j and $\bar{k}_j$ in their respective representations. By way of illustration we show $k_1, \bar{k}_1, k_3$, and $\bar{k}_3$, with $\gamma_1=1/(360\,h^4)$, $\gamma_2=1/48$.

The matrices k_1 and $\bar{k}_1$ are symmetric and positive semi-definite. Moreover

$$\left.\begin{aligned} k_1 e=\bar{k}_1 e=0 \quad & e^t=(1,\ldots,1)\\ x^t k_1 x>0 \quad & \\ x^t \bar{k}_1 x>0 \quad & \end{aligned}\right\} \quad \forall x\neq e\in \mathbb{R}^9. \tag{3.6}$$

Because of the boundary conditions, the matrices $k_2, \bar{k}_2, k_3$ and $\bar{k}_3$ are symmetric and positive definite.

Let us now consider the generalized eigenvalue problems

$$\gamma_2 k_j x=\lambda_j \gamma_1 \bar{k}_j x, \quad j=1,2,3. \tag{3.7}$$

The eigenvalues $\lambda_j^{(k)}$ are independent of h. They are real and can be ordered as follows

$$j=1 \Rightarrow 0<\lambda_1^{(2)} \leqq \cdots \leqq \lambda_1^{(9)}$$
$$j=2 \Rightarrow 0<\lambda_2^{(1)} \leqq \cdots \leqq \lambda_2^{(6)}$$
$$j=3 \Rightarrow 0<\lambda_3^{(1)} \leqq \cdots \leqq \lambda_3^{(4)}.$$

Note that, as the unit vector is mapped onto zero by both k_1 and $\bar{k}_1$, the eigenvalue $\lambda_1^{(1)}$ can be chosen arbitrarily. We denote the minimal and maximal eigenvalues by $\lambda_{\min}=\min(\lambda_1^{(2)}, \lambda_2^{(1)}, \lambda_3^{(1)})$ and $\lambda_{\max}=\max(\lambda_1^{(9)}, \lambda_2^{(6)}, \lambda_3^{(4)})$ respectively. It follows from (3.5), (3.6) and (3.7) that

$$\lambda_{\min}\,\gamma_1 x^t \bar{K} x = \sum_{j=1}^{3} \sum_{i \in I_j} (\lambda_{\min}\,\gamma_1 x_1^t\, \bar{k}_j x_i)$$
$$\leqq \sum_{j=1}^{3} \sum_{i \in I_j} (\gamma_2 x_i^t k_j x_i) = \gamma_2 x^t K x \tag{3.8}$$

and

$$\lambda_{\max}\,\gamma_1 x^t \bar{K} x = \sum_{j=1}^{3} \sum_{i \in I_j} (\lambda_{\max}\,\gamma_1 x_i^t \bar{k}_j x_i)$$
$$\geqq \sum_{j=1}^{3} \sum_{i \in I_j} (\gamma_2 x_i^t k_j x_i) = \gamma_2 x^t K x. \tag{3.9}$$

Thus we obtain

$$\lambda_{\min}\,(\gamma_1/\gamma_2) x^t \bar{K} x \leqq x^t K x \leqq \lambda_{\max}(\gamma_1/\gamma_2) x^t \bar{K} x$$

i.e., we have (3.4) with $d \equiv (\lambda_{\max}/\lambda_{\min})$, which is independent of the mesh parameter h.

$\lambda_1^{(2)}=0{,}4724$	$\lambda_1^{(9)}=8.2093$
$\lambda_2^{(1)}=0.8140$	$\lambda_2^{(6)}=12.05$
$\lambda_3^{(1)}=1.272$	$\lambda_3^{(4)}=11.84$

Table 3.1

The extreme eigenvalues of the generalised eigenproblems (3.7) are given in Table 3.1. We conclude that $d \approx 25.508$.

We can take another step in finding spectrally equivalent matrices. Let us denote by $\bar{\bar{K}}$ the stiffness matrix corresponding to a linear approximation on the mesh $\Delta_{n+1,n+1}$.

Theorem 3.2. *The inequalities*

$$x^t \bar{\bar{K}} x \leqq x^t \bar{K} x \leqq 3 x^t \bar{\bar{K}} x \tag{3.10}$$

hold for every vector $x \in \mathsf{R}^{n^2}$.

Proof. The proof is straightforward, and can be performed using the techniques found in [AB]. ●

In conclusion of this section we will comment briefly on the above results. There exist several preconditioners for the solution of the linear systems arising from the approximate solution of (3.1) by linear or bilinear finite elements. Let C be such a preconditioner, and let $\bar{n}(\varepsilon)$ and $\tilde{n}(\varepsilon)$ be the respective numbers of iterations needed for the solution of these linear systems by a thus preconditioned conjugate gradient method to some relative accuracy ε. It follows from theorems 3.1 and 3.2 that, if we use the same preconditioner for the solution of the system (3.3), the number of iterations $n(\varepsilon)$ needed can be estimated by

$$n(\varepsilon) < 5.1\bar{n}(\varepsilon)$$
$$n(\varepsilon) < 8.8\tilde{n}(\varepsilon).$$

The results in the next section show that the constants involved are more favourable than this. We conclude that we can use all preconditioners for linear and bilinear element (the properties of which are well studied) for the efficient solution of the spline system (3.3).

Remark 3.3. The results given in this section can be generalized to non-equally spaced domain. In that case one must use the so-called normalized B-splines, defined by the formula

$$N_{2,k}(x) = (1/3h)(x_{k+3} - x_k)B_{2,k}(x); \quad h = 1/(n+1), \tag{3.11}$$

then the first relation of (3.6) holds and the proof follows along the same lines.

4. The spectrum of the preconditioned system

We have used a number of preconditioners to solve systems with spline finite element matrices. In this section we will define the preconditioners (this is largely the presentation in [AE]), and give a model Fourier analysis of the spectrum of the resulting preconditioned system.
All preconditioners used were incomplete block factorizations; they can be described as follows. Let

$$A = \begin{pmatrix} X & B^t \\ B & E \end{pmatrix} = \begin{pmatrix} X & 0 \\ B & E - BX^{-1}B^t \end{pmatrix} \begin{pmatrix} I & X^{-1}B^t \\ 0 & I \end{pmatrix}$$

be a typical step in a block factorization, where X is the current block pivot, then the incomplete factorization can be written as

$$C = \begin{pmatrix} Z^{-1} & 0 \\ B & E - BYB^t \end{pmatrix} \begin{pmatrix} I & ZB^t \\ 0 & I \end{pmatrix},$$

where Y and Z are two approximations to X^{-1}. As the efficiency of the preconditioner mainly depends on the choice for Z, we only varied this; we have chosen $Y = [X^{-1}]^{(1)}$, i.e., the inner tridiagonal part of X^{-1}, throughout. The preconditioners now correspond to various choices for Z:

– C_r is the recursive preconditioner obtained by letting $Z = X^{-1}$. This is realized by storing a factorization of X. Note that solution of systems with C involves solving systems with X, which is not a vectorizable operation.

– C_e is an explicit (whence vectorizable) preconditioner, obtained by letting $Z = Y$; as is shown in [AP] this preconditioner is positive definite for the Poisson problem, but not necessarily so in general, so if we were to use it for other than Poisson problems it may lead to a divergent method. Two ways of mending this (while retaining easy vectorizability) are given in the last two preconditioners:

– C_x is a so-called 'product-expansion' preconditioner; if X is factored $X=(I-L)D^{-1}(I-L^t)$, we let $Z=(I+L^t)D(I+L)$. This is the smallest instance of the general product expansion approximation (see [AE])

$$(I-L)^{-1} \approx \sum_{k=0}^{2^p-1} L^k = \prod_{k=0}^{p} (I+L^{2^k}).$$

It is easy to see that this preconditioning matrix is positive definite.

– C_{rm}, finally, is the 'modified' recursive factorization, i.e., it is a recursive factorization like C_r, but modified by adding a positive diagonal to Y, in order to satisfy a generalized rowsum criterion $Cv = Av$, where v is a positive vector (in the numerical tests we have taken $v^t = (1, \ldots, 1)$). It is readily proved that this preconditioner is positive definite (see [AP], [AE]). For modified point factorizations as asymptotically faster rate of convergence can be proved (see [AB]), and for block factorizations this seems to be true under certain conditions too, see [B] and for model type problems [AE]. The tests presented in the next section indicate that, at least for the Poisson problem, an order reduction occurs.

We will now analyse the spectrum of the peconditioned spline system with a C_r preconditioner (from now on denoted just C) for the Laplace problem. In this analysis we consider the eigenvalue spectrum of $C^{-1}K$ as the product of $C^{-1}(-\Delta_5)$ and $(-\Delta_5)^{-1}K$, where we use the limit form of the peconditioner, i.e., we consider the preconditioner to be built solely from the limit blocks appearing in the factorization of a constant-coefficient matrix. Also we neglect the boundary of the domain in the construction of the spline matrix.

Symmetric block Toeplitz matrices with Toeplitz blocks, for instance deriving from constant coefficient problems, all have eigenvectors

$$u_{nm}(x, y) = \sin n\pi x \ \sin m\pi y.$$

As Δ_5, the spline matrix, and the limit form of the preconditioner all of this type, we have, denoting

$$\lambda_n = 4 \sin^2 \pi n h/2$$
$$\beta_n = 4\lambda_n + 2\lambda_{2n}$$
$$\alpha_n = 120 - 26\lambda_n - \lambda_{2n}$$

the following eigenvalues corresponding to the eigenvectors u_{nm}:

$$\text{for } -\Delta_5 : \delta_{nm} = \lambda_n + \lambda_m$$

$$\text{for the spline system: } \sigma_{nm} = \beta_n\alpha_m + \beta_m\alpha_n$$

$$\text{for the preconditioner } C = C_r : \gamma_{nm} = \xi_n + \xi_n^{-1} + \lambda_m - 2,$$

where

$$\xi_n = \mathrm{d}(\mu\lambda_n + (1-\mu)^2)$$

is an eigenvalue of the limit pivot block; the quantities d and μ (the limit pivot and decay rate of the pivot block) can be determined numerically for any given (constant coefficient) differential equation.

We need a measure of the 'smoothness' of the eigenvectors; for this we use a scaled L_2-norm of the gradient of the vector. Thus eigenvector u_{nm} has smoothness n^2+m^2.

After these preparations we can analyse the spectra of $C^{-1}(-\Delta_5)$, $(-\Delta_5)^{-1}K$, and $C^{-1}K$. In particular we will be interested in how smooth and fast oscillating eigenvectors are treated. The easiest way to investigate this is graphically*. Thus we will plot (for $n, m = 1, \ldots, h^{-1}-1$) eigenvalues against n and m, and we will plot the smoothness n^2+m^2 against the eigenvalues.

We start by plotting the eigenvalues $\delta_{nm}^{-1}\sigma_{nm}$ of $(-\Delta_5)^{-1}K$ for $h=1/40$ in figure 4.1. It is seen that such a preconditioned system has fairly uniform eigenvalues for fast to medium fast oscillating eigenvectors, but that it has large eigenvalues for the smoothest eigenvectors. This behaviour is born out by figure 4.2 where we plot the smoothness against the eigenvalue. Note that we know the large eigenvalues to be bounded uniformly in h from the spectral equivalence shown above.

Consider next the system $C^{-1}(-\Delta_5)$. It is known that the eigenvalues satisfy

$$\gamma_{nm}^{-1}\delta_{nm} \in (\varepsilon_1, 1+\varepsilon_2),$$

where ε_2 is a small number, and $\varepsilon_1 = O(h^2)$, with a clustering of eigenvalues around 1, see [AL]; from figure 4.3 we see that the eigenvalues tending to zero correspond to the smooth eigenvectors.

Now combining these two preconditioned systems, we find essentially the same behaviour as for the preconditioned spline system for any but the smooth eigenvectors. For these we see the product of big, but bounded, eigenvalues from $(-\Delta_5)^{-1}K$, and eigenvalues tending to zero from $C^{-1}(-\Delta_5)$. The resulting plot is figure 4.4. If we now plot the smoothness against the eigenvalue (in figure 4.5 for $h=1/40$ and in figure 4.6 for $h=1/80$), we obtain the same shape as in figure 4.2, but the smooth eigenvectors do not have increasing eigenvalues this time. Instead they can be seen to 'travel to the left' through the spectrum for subsequently smaller meshsizes.

The resulting behaviour of these multiplied preconditioned systems, then, is that of a spectrum that stays essentially bounded for all mesh sizes up till a certain limit; for smaller values of h the smoothest eigenvectors will make the lower bound on the spectrum decrease. From well-known bounds for the convergence of preconditioned conjugate gradient methods (see [AB]) we now expect for these preconditioned spline systems a number of iterations that increases slowly for meshsizes coarse enough, until the eigenvalues corresponding to smooth eigenvectors start protruding at the left end of the spectrum (for the Poisson equation this can be calculated to happen at $h^{-1}=58$). For finer mesh sizes the number of iterations will increase in the same way as for linear element systems.

* Note that we have full knowledge of eigensystems, so any qualitative result can be backed up computationally. However, such calculations are tedious, and probably less insightful than graphs.

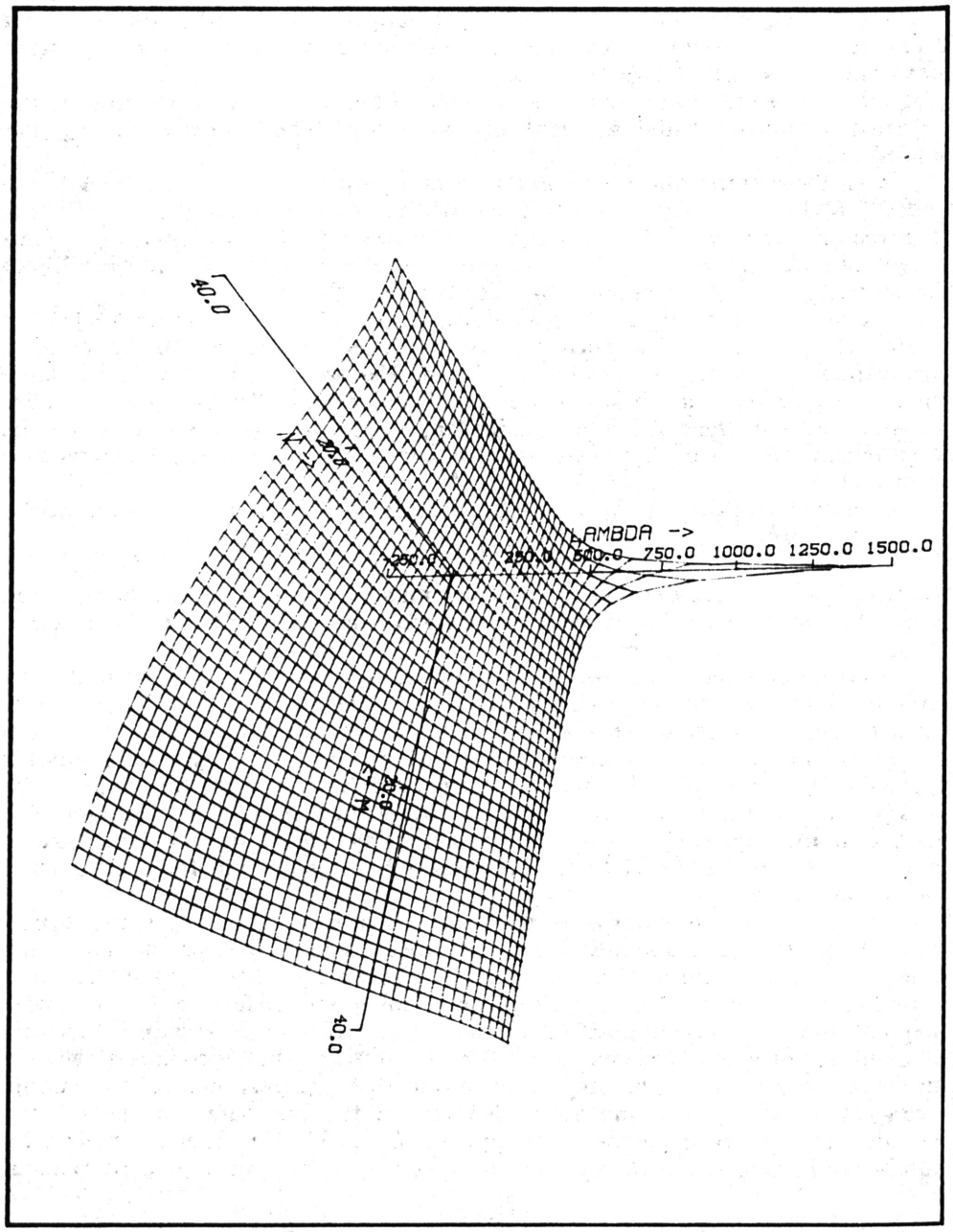

Fig. 4.1. $\lambda_{nm}((-\Delta_5)^{-1}K)$ eigenvalue plot, 40 points per line

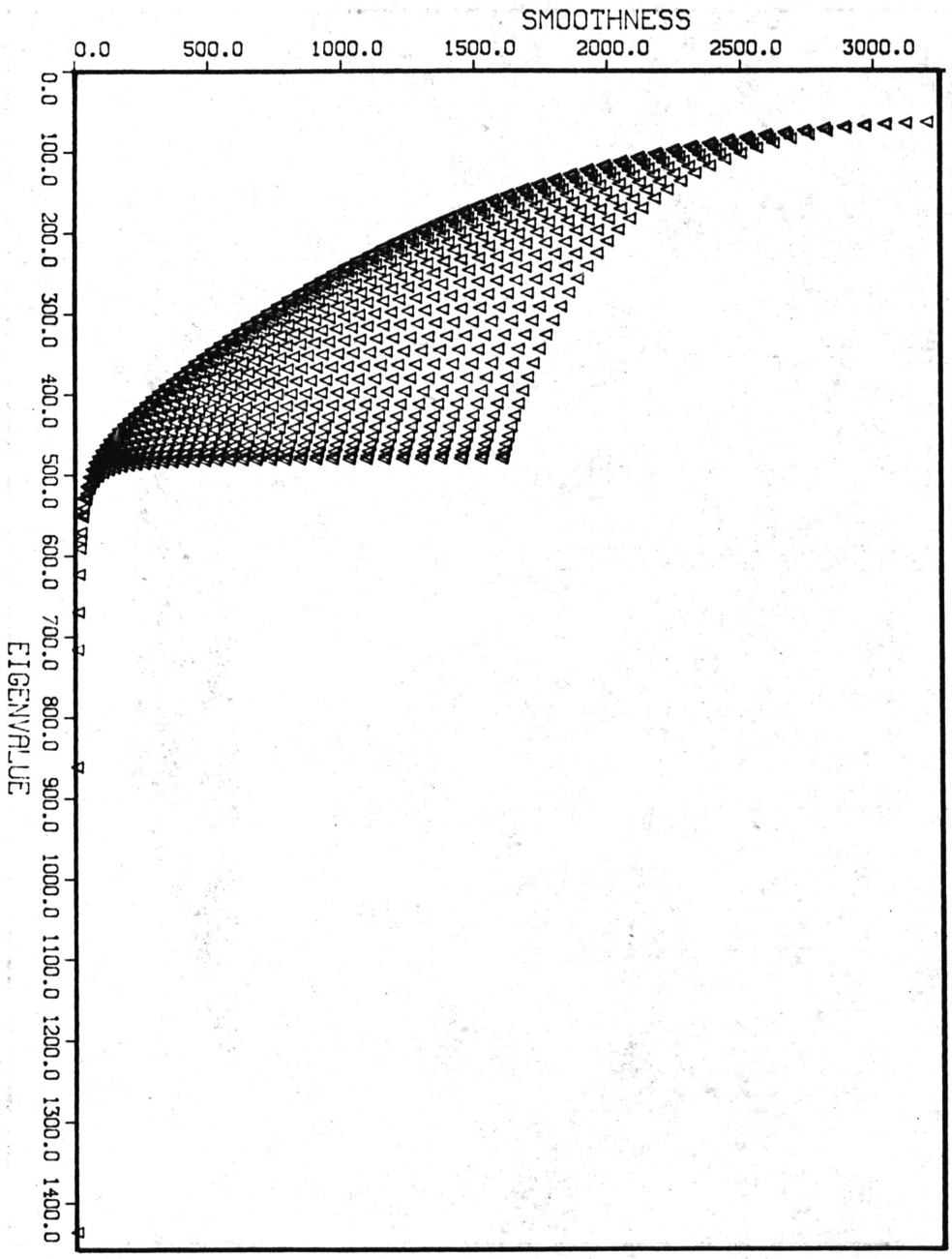

Fig. 4.2. L_1-norm of eigenvectors of $(-\Delta_5)^{-1}K$ versus eigenvalues, 40 points per line

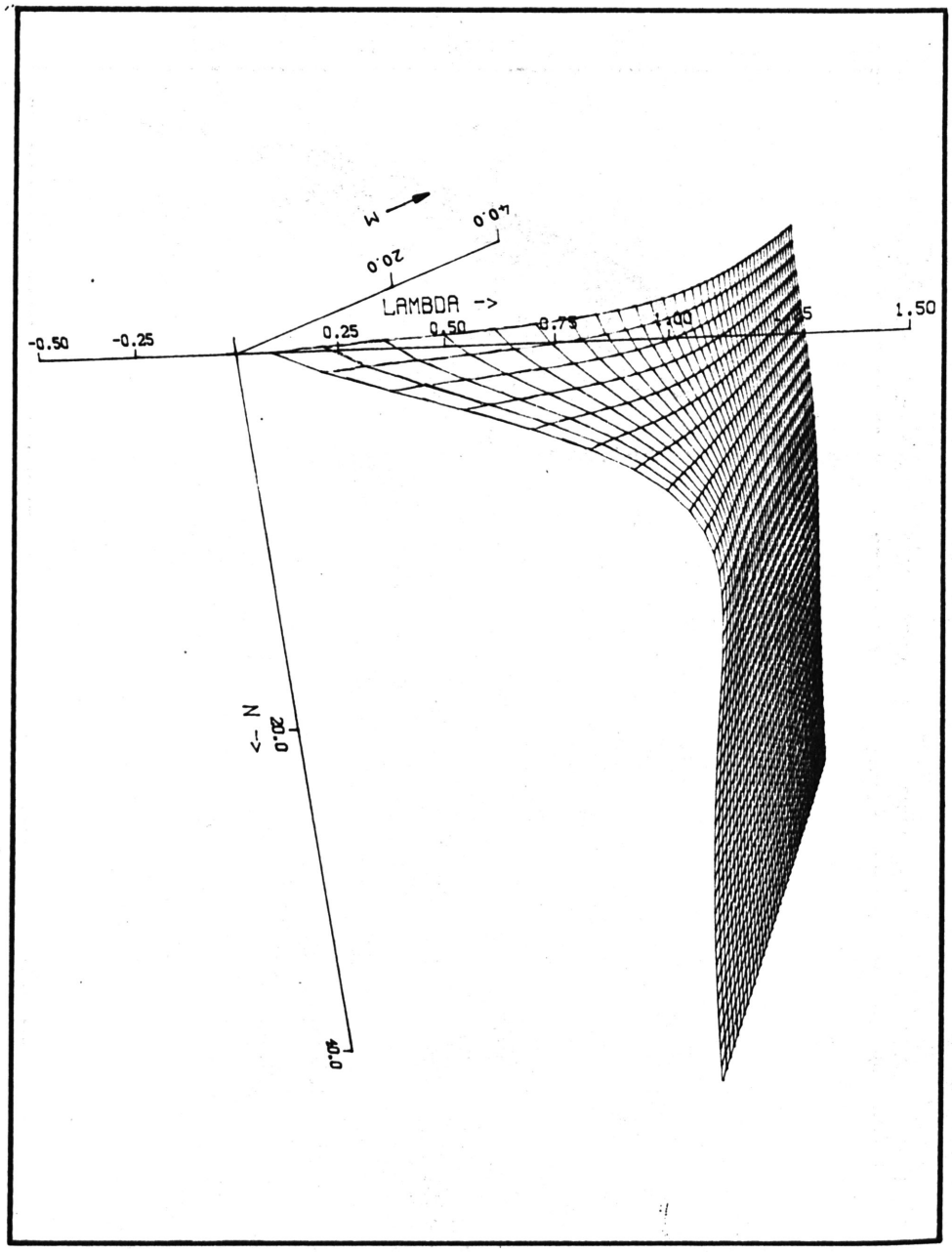

Fig. 4.3. $\lambda_{nm}(C_r^{-1}A)$ eigenvalue plot, 40 points per line

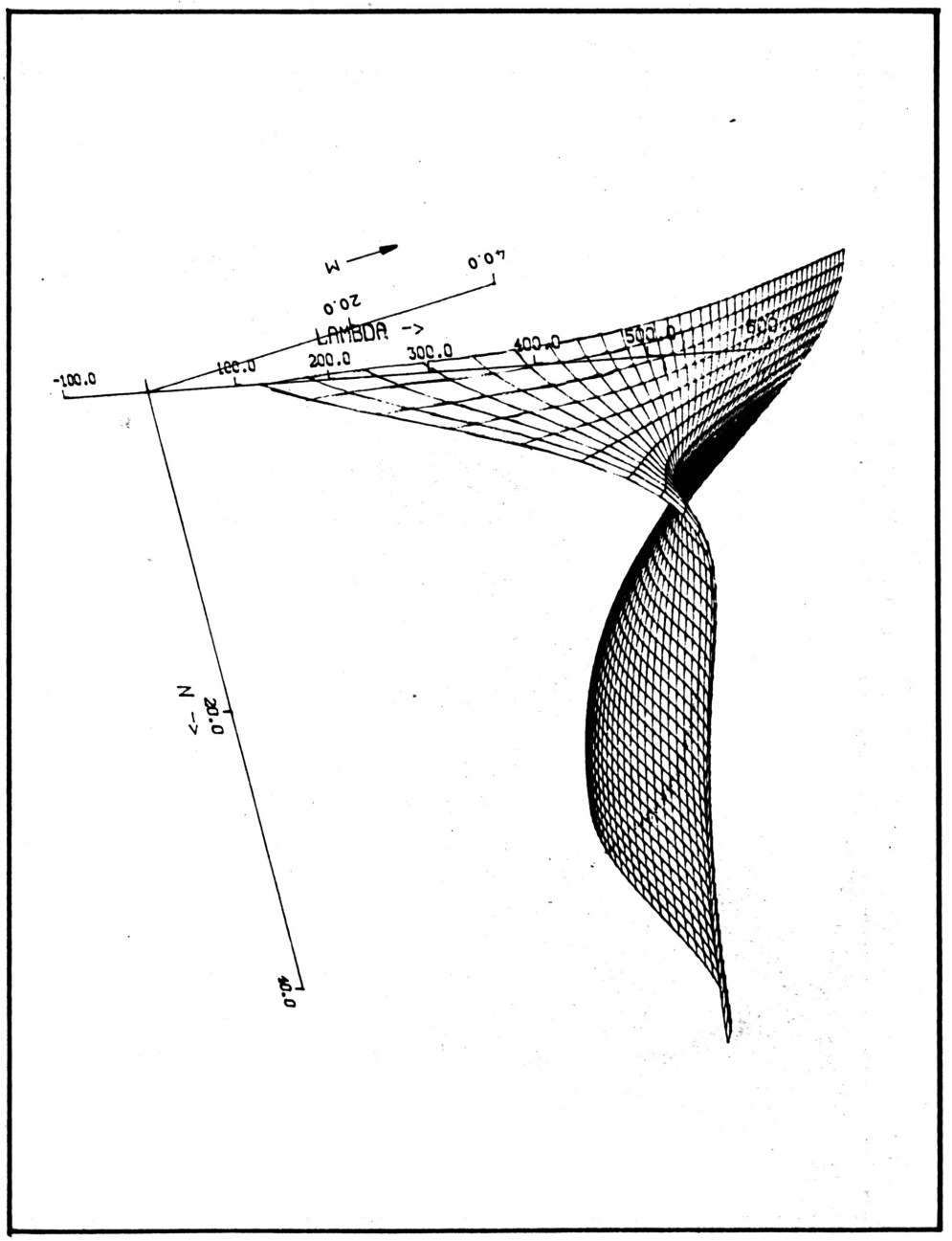

Fig. 4.4. $\lambda_{nm}(C_r^{-1}K)$ eigenvalue plot, 40 points per line

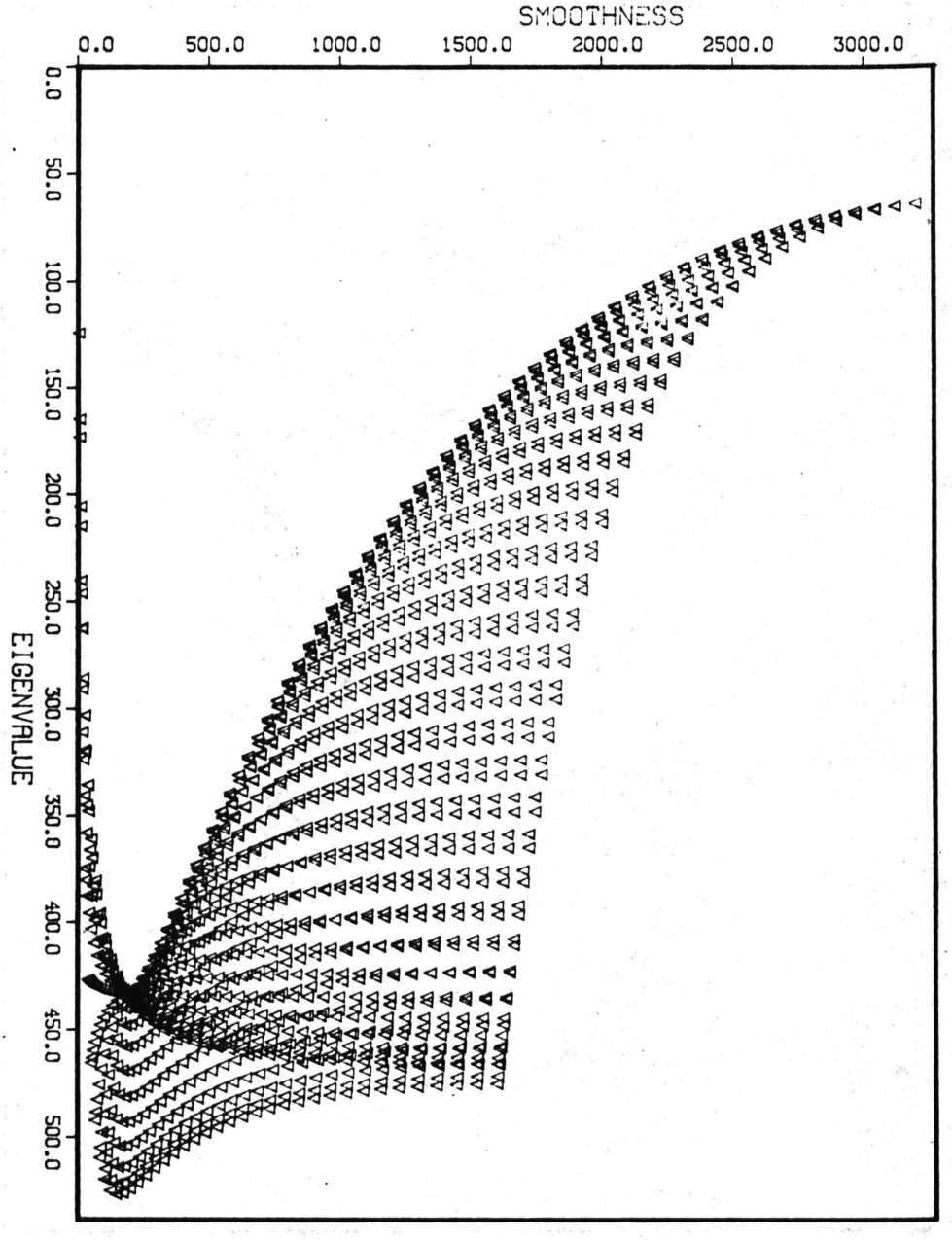

Fig. 4.5. L_1-norm of eigenvectors of $C_r^{-1}K$ versus eigenvalues, 40 points per line

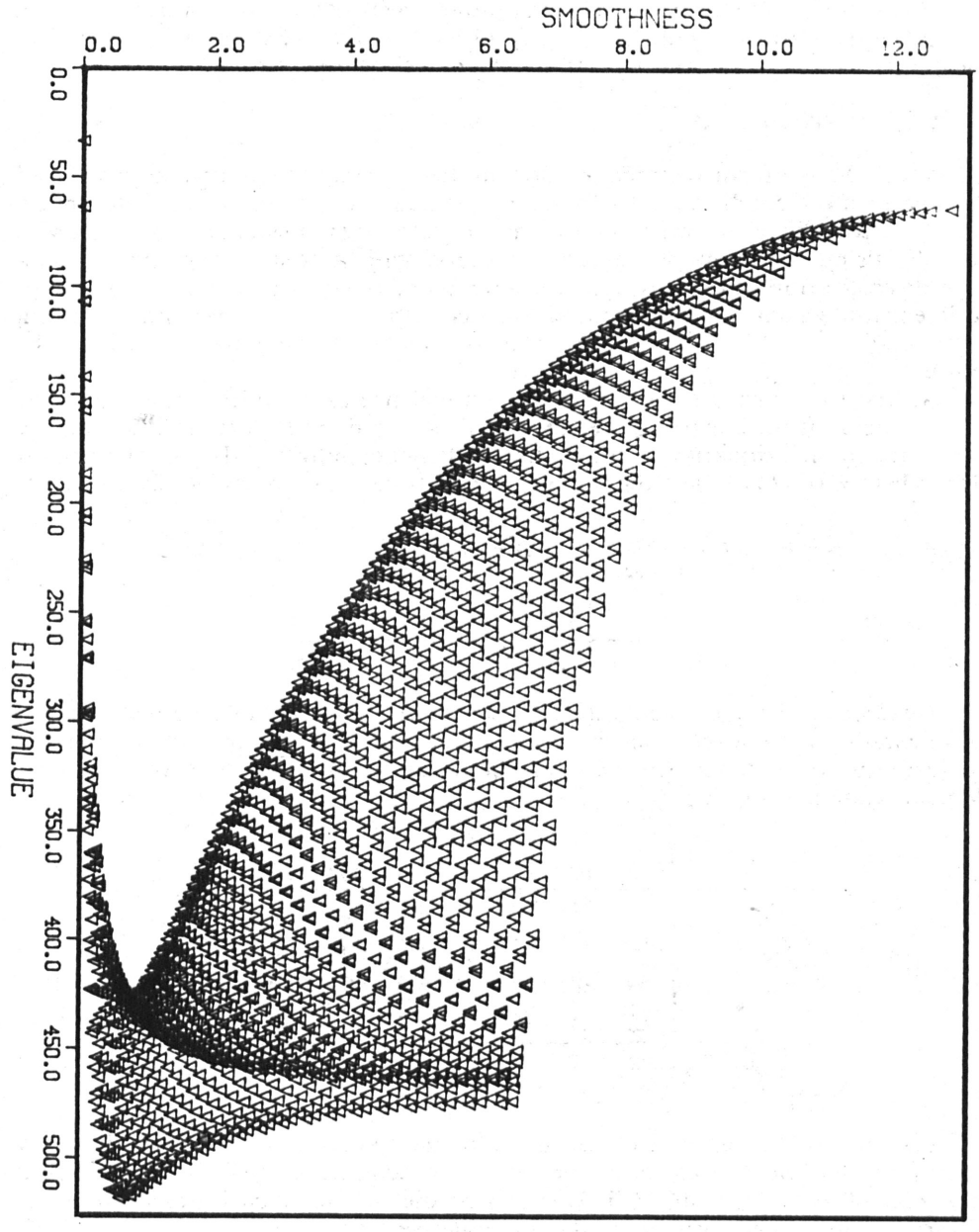

Fig. 4.6. L_1-norm of eigenvectors of $C_r^{-1}K$ versus eigenvalues, 80 points per line

Finally we note that for C_x and other vectorizable preconditioners the clustering of eigenvalues of $C^{-1}(-\Delta_5)$ is less pronounced, so we expect any levelling effects in the numbers of iterations to be less strong.

5. Numerical tests

We have performed some numerical tests using the above approach of employing a preconditioner for a linear element system on the regular mesh $\Delta_{n+1,n+1}$ (see above) in order to solve the quadratic splines equations. This was done by using the conjugate gradient method with a coefficient matrix arising from discretization of (3.1) by spline elements and various preconditioners based on linear finite elements. The method used as initial vector $C^{-1}rhs$ and as halting criterion $g^tC^{-1}g < 10^{-10}|rhs|^2$, where C is the preconditioning matrix, and g is the residual.

At first we tested the actual spectral equivalence of the spline system and the matrix arising from linear finite elements. This was done by preconditioning the CG by the full factorization of the Δ_5 linear elements matrix. Table 4.1 shows that for mesh sizes of 1/20 or less the number of iterations remains essentially constant.

Table 4.1

Numbers of iterations for a CG method on (3.1) preconditioned with Δ_5

h^{-1}	5	10	20	40	56	80
$\#it$	6	15	28	30	31	30

Next we used some preconditioners based on the linear finite elements matrix to precondition the solution of the spline equations. For purposes of comparison we furthermore included the numbers of iterations needed to solve the linear element system using C_r.

Table 4.2

h^{-1}	5	10	20	40	56	80	113	160
C_r	15	21	30	31	31	35	41	52
C_e	14	26	34	36	42	46	55	75
C_x	19	28	33	37	40	49	59	80
C_{rm}	15	24	36	43	48	55	64	75
C_r on Δ_5	2	4	7	12	16	22	27	35

Looking at the numbers of iterations for the splines system, first of all we see a leveling off (most noticeable for the C_r preconditioner) for mesh sizes in the order of $h = 1/20$, 1/40, 1/56. This was predicted on the basis of the spectrum of the peconditioned system in the previous section.

For finer meshes the number of iterations increases proportional to that of the preconditioned solution of Δ_5 systems. It is important to note that the proportionality constant involved is rather lower than what would have been expected from the theoretical estimates in the previous section. In fact, we see that

the iterative solution of the linear system arising from a quadratic splines approximation takes about twice the number of iterations it takes to solve a linear finite elements system, which is much less than what would have been obtained if we would have preconditioned by an iterative solution of $(-\Delta_5)$ itself. Taking into account the higher accuracy of the splines approximation this makes it a very attractive scheme for the solution of elliptic equations.

6. References

[AB] O. Axelsson, V. A. Barker. Finite element solution of boundary value problems, Theory and computation. Academic Press, Orlando, Fl., 1984.

[AE] O. Axelsson, V. Eijkhout. Vectorizable Preconditioners for Elliptic Difference Equations in Three Space Dimensions (to appear in Journal for Computational and Applied Mathematics).

[AL] O. Axelsson, G. Lindskog. On the eigenvalue distribution of a class of preconditioning matrices. *Numer. Math.*, **48**, 1986, 479-498.

[AP] O. Axelsson, B. Polman. On approximate factorization methods for block matrices suitable for vector and parallel processors. *Lin. Alg. Appl.*, **77**, 1986, 3-26.

[AV] O. Axelsson, P. Vassilevski. Algebraic multilevel preconditioning methods, I, (to appear in Numer. Math.).

[BW] P. E. Bjørstad, O. B. Widlund. Iterative methods for the solution of elliptic problems on regions partitioned into substructures. *SIAM J. Numer. Anal.*, **23**, 1986, 1097-1120.

[EV] V. Eijkhout, P. Vassilevski. Positive definiteness aspects of vectorizable preconditioners. *Parallel Computing*, **10**, 1989, 93-100.

[M] S. D. Margenov. Application of spline-macroelements in solving the biharmonic equation. – In: Numerical methods '84, proceedings of the International Conference on Numerical Methods, Sofia, Bulgaria, 1985 (in Russian).

[MK] Marchuk, Yu. A. Kuznetsov. A generalized conjugate gradient method in a subspace and its applications. – In: Problems in Comp. Math. and Math. Modelling (ed. Marchuk), Nauka, Moscow, 1985, 10-32.

Department of Mathematics, University of Nijmegen
Toernooiveld 5, 6525 ED Nijmegen
the Netherlands
Center for Informatics and Computer Technology
Bulgarian Academy of Sciences
Acad. G. Bonchev str bl. 25A
1113 Sofia
BULGARIA

Received 21. 12. 1989