

## CAPACITANCE MATRIX METHOD FOR PETROV-GALERKIN PROCEDURE

SEIYOUNG CHUNG

### 1. Introduction

In this paper a capacitance matrix method is developed for the Poisson equation on a rectangle

$$(1-1) \quad \mathbf{L}u \equiv -(u_{xx} + u_{yy}) = f, \quad (x, y) \in \Omega \equiv (0, 1) \times (0, 1)$$

with the homogeneous Dirichlet boundary condition

$$(1-2) \quad u = 0, \quad (x, y) \in \partial\Omega$$

where  $\partial\Omega$  is the boundary of the region  $\Omega$ .

This method considered here is an application of the cubic spline method of I. Sloan in [1] to the boundary value problem in two space variables and is basically the discretization of the  $H^1$ -Galerkin method on an  $n \times n$  uniform partition, which is a Petrov-Galerkin procedure using a trial space of bicubic splines, a test space of piecewise-bilinear functions, and the composite Simpson's 9-point rule or the composite 4-point Gaussian quadrature for the integrals in two dimension. Then the resulting linear system is slightly different from the one which can be solved at the cost of  $O(n^2 \log n)$  by a matrix decomposition method using fast Fourier transforms(FFT). The capacitance matrix method utilizes this special structure to solve the linear system at the cost of  $O(n^3)$ , which is not optimal but quite efficient considering that the number of the unknowns is  $n^2$ , a quarter of the number of the unknowns in the orthogonal spline collocation. Notice that the existence and the

---

Received May 17, 1994.

1991 AMS Subject Classification: .

Key words: capacitance matrix method, fast direct method Petrov-Galerkin procedure.

uniqueness of the approximation for this method follow directly from [1]. The order of convergence will be discussed in section 2.

The method may be viewed as a "quasolocation" approximation (i.e. a quadrature-based modification of the collocation approximation) in [2,3] for the boundary-integral equations on smooth curves. The  $H^1$ -Galerkin method with Dirichlet boundary conditions was proposed by de Boor[4] and was also analyzed by G. Fairweather[5]. But only the  $O(h^3)$  convergence was obtained in [4]. The modified cubic spline collocation method, an  $O(h^4)$  cubic spline method developed independently by Archer[6] and Daniel and Swartz[7], employs nodal cubic spline collocation for the solution of a higher order perturbation of the differential equation. Another  $O(h^4)$  cubic spline method can be found in Chawla and Subramanian[8].

The practical difficulty in the  $H^1$ -Galerkin method lies in evaluating the inner products i.e. the integrals. But it can be overcome with ease in this discrete  $H^1$ -Galerkin procedure, and the matrix in the resulting linear system can be determined once for all for any number  $n$ . Furthermore it has a special structure as discussed in section 2. An algorithm involving fast direct method and capacitance matrix technique will be constructed in section 3.

## 2. Petrov-Galerkin procedure

Given  $n \geq 1$ , let  $\Omega_h$  be a uniform partition of the region  $\Omega$

$$\Omega_h = \{(x_i, y_j) \mid x_i = ih, y_j = jh, i, j = 0, 1, \dots, n, h = 1/n\}.$$

Let  $S(\Omega_h) \subset C^2(\Omega)$  be the set of bicubic splines defined on the uniform partition  $\Omega_h$  which satisfy the homogeneous boundary condition (1-2). Let  $T(\Omega_h)$  be the set of piecewise-bilinear functions on  $\Omega_h$  which do not satisfy the boundary conditions. Then the discrete  $H^1$ -Galerkin method of [1] is: find  $u_h \in S(\Omega_h)$  such that

$$(2-1) \quad \langle \mathbf{L}u_h, z_h \rangle = \langle f, z_h \rangle, \quad z_h \in T(\Omega_h),$$

where

$$\langle u, v \rangle = Q_h(uv)$$

and  $Q_h$  is a quadrature rule. In this paper the composite Simpson's 9-point rule in two dimension, a rule with fourth-order error, for the partition  $\Omega_h$  will be employed, see[14, chapter 5]:

$$Q_h(u) = \frac{h^2}{36} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} [u_{ij} + u_{i+1,j} + u_{i,j+1} + u_{i+1,j+1} + 4(u_{i+1/2,j} + u_{i,j+1/2} + u_{i+1,j+1/2} + u_{i+1/2,j+1}) + 16u_{i+1/2,j+1/2}],$$

where  $u_{i,j} = u(x_i, y_j)$  and  $x_{i+1/2} = (x_i + x_{i+1})/2$ . The composite 4-point Gauss rule may be employed and the method here can be applied to the resulting system.

To discuss the order of convergence, we notice that it is well known that there exists a positive constant  $\gamma$  such that

$$(2-2) \quad \|Lu\|_{L_2(\Omega)} \geq \gamma \|u\|_{H^2(\Omega)}$$

if  $u$  satisfies the boundary conditions (1-2), where  $H^m(\Omega)$  is the Sobolev space equipped with the norm

$$\|u\|_{H^m(\Omega)} = \left( \sum_{0 \leq |\alpha| \leq m} \|D^\alpha u\|_{L_2(\Omega)}^2 \right)^{1/2}$$

for any positive integer  $m$ . This inequality (2-2),  $u \in H^6(\Omega)$  and a quadrature rule with fourth-order error in (2-1) satisfies the assumptions of theorem 4.1 in [1]. It therefore follows from theorem 4.1 of [1] that (2-1) has a unique solution  $u_h \in S(\Omega_h)$  satisfying

$$\|u - u_h\|_{H^i(\Omega)} \leq \gamma h^{4-i} \|u\|_{H^6(\Omega)}, \quad i = 0, 1, 2.$$

Let  $\hat{B}_i(x)$ , for  $0 \leq i \leq n$ , be the cubic spline with the support  $[x_{i-2}, x_{i+2}]$  in one space variable over the extended partition

$$x_{-3} < x_{-2} < x_{-1} < x_0 < \dots < x_n < x_{n+1} < x_{n+2} < x_{n+3},$$

$$x_k = kh, \quad k = -3, -2, -1, 0, \dots, n, n+1, n+2, n+3, \quad h = 1/n,$$

such that

$$\hat{B}_i(x_j) = \begin{cases} 2/3 & \text{if } i = j \\ 1/6 & \text{if } i = j \pm 1 \\ 0 & \text{if otherwise.} \end{cases}$$

We define the cubic B-spline  $B_i(x)$ , for  $0 \leq i \leq n$ , satisfying the homogeneous boundary conditions (1-2):

$$\begin{aligned} B_0(x) &= \hat{B}_0(x) - 4\hat{B}_{-1}(x), & B_1(x) &= \hat{B}_1(x) - \hat{B}_{-1}(x), \\ B_i(x) &= \hat{B}_i(x), & i &= 2, 3, \dots, n-2, \\ B_{n-1}(x) &= \hat{B}_{n-1}(x) - \hat{B}_{n+1}(x), & B_n(x) &= \hat{B}_n(x) - 4\hat{B}_{n+1}(x). \end{aligned}$$

Then we may represent an approximation  $u_h \in S(\Omega_h)$  by

$$u_h = \sum_{k=0}^n \sum_{l=0}^n w_{kl} B_{kl}(x, y),$$

where  $B_{kl}(x, y) = B_k(x)B_l(y)$ , for  $k, l = 0, 1, \dots, n$ .

We mean  $z_j(x)$ , for  $0 \leq j \leq n$ , by piecewise-linear "hat" function in one space variable with support  $[x_{j-1}, x_{j+1}]$ :

$$z_j(x_k) = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if otherwise.} \end{cases}$$

The equation (2-1) is equivalent to

$$(2-3) \quad \sum_{k=0}^n \sum_{l=0}^n w_{kl} \langle \mathbf{L}B_{kl}, z_{ij} \rangle = \langle f, z_{ij} \rangle, \quad i, j = 0, 1, \dots, n,$$

where  $z_{ij}(x, y) = z_i(x)z_j(y)$ . Let's define the following:

$$\mathbf{f} = (f_{00}, \dots, f_{0n}, \dots, f_{n0}, \dots, f_{nn})^T, \quad f_{ij} = 864 \langle f, z_{ij} \rangle,$$

$$\mathbf{w} = (w_{00}, \dots, w_{0n}, \dots, w_{n0}, \dots, w_{nn})^T,$$

$$A = (a_{ij}), \quad a_{ij} = -6h \int_0^1 B''(x)z_i(x)dx, \quad i, j = 0, 1, \dots, n,$$

$$B = (b_{ij}), \quad b_{ij} = 144h^{-1} \int_0^1 z_i(x)B_j(x)dx, \quad i, j = 0, 1, \dots, n.$$





It follows immediately from the above Proposition and the properties of matrix tensor product that the equation (3-1) is equivalent to

$$(3-2) \quad (D_1 \otimes D_2 + D_2 \otimes D_1)(S \otimes S)\mathbf{w} = (S \otimes S)\mathbf{f},$$

which in turn shows that the following algorithm holds.

### Fast direct algorithm

1. Compute  $\hat{\mathbf{f}} = (S \otimes S)\mathbf{f}$ .
2. Solve  $(D_1 \otimes D_2 + D_2 \otimes D_1)\hat{\mathbf{w}} = \hat{\mathbf{f}}$ .
3. Compute  $\mathbf{w} = (S \otimes S)\hat{\mathbf{w}}$ .

It can be easily verified that the step 1 and the step 3 of the algorithm can be performed by FFT at the cost of  $O(m^2 \log m)$ . Therefore the total amount of arithmetic operations of the algorithm is  $O(m^2 \log m)$  since the coefficient matrix of the system in step 2 is the diagonal matrix of order  $m^2$ . Notice that all steps of the algorithm can be implemented for parallel computations, see[11].

Let  $Q$  and  $R$  be nonsingular matrices of order  $m^2$  whose  $k$ -th rows are different from each other only for  $k \in \mathcal{K}$ , where  $\mathcal{K}$  is any subset of the index set  $\{1, 2, \dots, m^2\}$ . Define an  $m^2$ -dimensional vector  $\hat{\mathbf{b}} = (\hat{b}_i)$  for a given  $m^2$ -dimensional vector  $\mathbf{b} = (b_i)$

$$\hat{b}_i = \begin{cases} 0, & \text{if } i \in \mathcal{K} \\ b_i, & \text{if otherwise.} \end{cases}$$

We seek the solution of  $R\mathbf{x} = \mathbf{b}$  in the form

$$\mathbf{x} = \mathbf{y} + \sum_{i \in \mathcal{K}} v_i \mathbf{g}_i, \quad v_i \in \mathbf{R},$$

where  $\mathbf{y}$  is the solution of the equation  $Q\mathbf{y} = \hat{\mathbf{b}}$  and  $\mathbf{g}_i$ 's are  $m^2$ -dimensional vectors to be determined. Let  $\mathbf{e}_i$  denote the  $i$ -th standard unit vector and  $(\mathbf{x})_i$  the  $i$ -th component of a vector  $\mathbf{x}$  for any  $i$ . We now describe the capacitance matrix technique, see [12,13].

### Capacitance matrix algorithm

1. Solve  $Q\mathbf{g}_j = \mathbf{e}_j, \quad j \in \mathcal{K}$ .

2. Compute the capacitance matrix  $C = (c_{ij})$ :

$$c_{ij} = (R\mathbf{g}_j)_i, \quad i, j \in \mathcal{K}$$

3. Solve  $Q\mathbf{y} = \hat{\mathbf{b}}$ .

4. Compute  $\mathbf{h} = (h_i)$ ,  $h_i = b_i - (R\mathbf{y})_i$ ,  $i \in \mathcal{K}$ .

5. Solve  $C\mathbf{v} = \mathbf{h}$ .

6. Solve  $Q\mathbf{x} = \hat{\mathbf{b}} + \sum_{i \in \mathcal{K}} v_i \mathbf{e}_i$ ,  $\mathbf{v} = (v_i)$ ,  $i \in \mathcal{K}$ .

It is easy to verify that

$$\begin{aligned} (R\mathbf{x})_j &= (R\mathbf{y})_j + \sum_{i \in \mathcal{K}} v_i (R\mathbf{g}_i)_j \\ &= \begin{cases} \hat{b}_j + \sum_{i \in \mathcal{K}} v_i (\mathbf{e}_i)_j = b_j + 0 = b_j, & \text{if } j \in \mathcal{K} \\ (R\mathbf{y})_j + \sum_{i \in \mathcal{K}} v_i c_{ji} = (R\mathbf{y})_j + h_j = b_j, & \text{if } j \notin \mathcal{K}, \end{cases} \end{aligned}$$

which justifies the algorithm. Note that the capacitance matrix  $C$  is usually dense and hence that the algorithm is useful only when  $Q\mathbf{x} = \mathbf{b}$  can be solved easily and the number of different rows is quite small comparing to the order of the coefficient matrix  $R$ .

Introducing

$$\begin{aligned} T_1 &= -T(-6, 2), & T_2 &= T(78, 32) \\ Q &= T_1 \otimes T_2 + T_2 \otimes T_1, & R &= A \otimes B + B \otimes A, \end{aligned}$$

we see that the capacitance matrix method with the fast direct algorithm may be applied to solve the equation (2-3) and the index set  $\mathcal{K}$  is

$$\mathcal{K} = \left\{ k = (i-1)m + j \mid \begin{array}{l} (i = 1, 2, m-1, m, j = 1, 2, \dots, m) \\ \text{or } (i = 1, 2, \dots, m, j = 1, 2, m-1, m) \end{array} \right\}.$$

The number of elements of the index set  $\mathcal{K}$  is  $8m - 16$ . Therefore the step 1 requires the arithmetic operations of  $O(m^3 \log m)$  if the fast direct algorithm is applied without any modification. To reduce the order of the cost by one, we may take advantage of the structure of



*R.* Since matrices  $A$  and  $B$  are pentadiagonal, it can be easily verified that for  $i = (k - 1)m + l$ ,

$$(R\mathbf{g}_j)_i = \sum_{r=k-2}^{k+2} \sum_{s=l-2}^{l+2} (b_{ls}a_{kr} + a_{ls}b_{kr})(\mathbf{g}_j)_{sr}$$

and hence that to calculate  $c_{ij} = (R\mathbf{g}_j)_i$  for all  $i \in \mathcal{K}$  in step 2, we need only  $16(m - 4)$  elements of the vector  $\mathbf{g}_j$ , not all of them:

$$\{(\mathbf{g}_j)_i \mid i \in \mathcal{H}\},$$

$$\mathcal{H} = \{i = (r - 1)m + s \mid$$

$$(r = 1, 2, 3, 4, m - 3, m - 2, m - 1, m, s = 1, \dots, m)$$

$$\text{or } (r = 5, 6, \dots, m - 4, s = 1, 2, 3, 4, m - 3, m - 2, m - 1, m)\}.$$

We have shown that the capacitance matrix  $C$  can be formulated at the cost of  $O(m^2)$ . Now for each  $j \in \mathcal{K}$  in step 1, we solve  $(\mathbf{g}_j)_i$ 's only for  $i \in \mathcal{H}$  not for all  $i$  at the cost of  $O(m \log m)$  by means of FFT. Therefore the step 1 requires the arithmetic operations of  $O(m^2 \log m)$ . It is easy to see that the cost for step 3, step 4 and step 6 are  $O(m^2 \log m)$ ,  $O(m^2)$  and  $O(m^2 \log m)$  respectively. The matrix  $C$  in step 5 is a dense matrix of order  $(8m - 16)$ , which means that our algorithm requires the arithmetic operations of  $O(m^3)$  or  $O(n^3)$  in total.

ACKNOWLEDGEMENTS. This work was initiated while the author was visiting the University of Kentucky as a post-doctoral fellow, which was supported by KOSEF. The author would like to appreciate the support and also would like to thank Professor Graeme Fairweather for his kindly help and suggestions in preparing this paper.

### References

1. I.H. Sloan, Dat Tran and G. Fairweather, *A fourth-order cubic spline method for linear second-order two point boundary value problems* (to appear).
2. I.H. Sloan, *A quadrature-based approach to improving the collocation method*, Numer. Math. **54** (1988), 41-56.
3. G.A. Chandler and I.H. Sloan, *Spline quallocation methods for boundary integral equations*, Numer. Math. **58** (1990), 537-567.
4. C. de Boor, *The method of projections as applied to the numerical solution of two point boundary value problems*, Ph.D Thesis (1966), Univ. of Michigan.

5. G. Fairweather, *Finite Element Galerkin Methods for Differential Equations*, Lecture Notes in Pure and Applied Mathematics, Vol. 34, Marcel Dekker, 1978.
6. D. Archer, *An  $O(h^4)$  cubic spline collocation methods for quasilinear parabolic equations*, SIAM J. Numer. Anal. **14** (1977), 620-637.
7. J.W. Daniel and B.K. Swartz, *Extrapolated collocation for two point boundary value problems using cubic splines*, J. Inst. Maths. Applics. **16** (1975), 161-174.
8. M.M. Chawla and R. Subramanian, *A fourth-order cubic spline method for second-order nonlinear two-point boundary-value problems*, J. Comp. Appl. Math. **23** (1988), 1-10.
9. D.E. Rutherford, *Some Continuant Determinants arising in Physics and Chemistry*, Proceedings of the Royal Society of Edinburgh, Section A **LXII-PartIII** (1947), 229-236.
10. ———, *Some Continuant Determinants arising in Physics and Chemistry II*, Proceedings of the Royal Society of Edinburgh, Section A **LXIII-PartIII** (1950), 232-241.
11. B. Bialecki, G. Fairweather and K.R. Bennett, *Fast direct solvers for piecewise Hermite bicubic orthogonal spline collocation equations*, SIAM. J. Numer. Anal. **29** (1992), 156-193.
12. M. Dryja, *A capacitance matrix method for Dirichlet problem on polygon region*, Numer. Math. **39** (1982), 51-64.
13. B.L. Buzbee, F.W. Dorr, J.A. George and G.H. Golub, *The direct solution of the discrete Poisson equation on irregular regions*, SIAM. J. Numer. Anal. **8** (1971), 722-736.
14. P.F. Davis and P. Rabinowitz, *Methods of Numerical Integration*, Academic Press, 1984.

Department of Mathematics  
Chungnam National University  
Taejon 305-764, Korea  
e-mail: sychung@math.chungnam.ac.kr