Int. J. Solids Structures, 1973, Vol. 9, pp. 1013 to 1034. Pergamon Press. Printed in Great Britain

BOUNDS ON THE SPECTRAL AND MAXIMUM NORMS OF THE FINITE ELEMENT STIFFNESS, FLEXIBILITY AND MASS MATRICES

ISAAC FRIED†

Boston University, Boston, Massachusetts 02215, U.S.A.

Abstract—Upper and lower bounds are established on the spectral and maximum norms (and hence on the corresponding condition numbers) of the stiffness, flexibility (the inverse of the stiffness) and mass matrices generated from regular and irregular meshes of finite elements. Explicit expressions for these bounds are derived, in terms of the intrinsic and discretization parameters, for second and fourth order problems in one, two and three dimensions discretized with linear, triangular and tetrahedronal elements.

INTRODUCTION

A PROCEDURE has been devised in Ref. [1] (also briefly described in Ref. [2]) for bounding the extremal eigenvalues of the global stiffness and mass matrices, generated by the finite element method, in terms of the extremal eigenvalues of the element stiffness and mass matrices, the maximal number of elements meeting at a nodal point and the exact fundamental frequency of the structure. The attractiveness of this procedure lies both in its generality and computability; it can be applied to regular as well as to irregular meshes, and once an estimate for the fundamental frequency of the structure is known numerical bounds are readily computed.

To be more precise: the bounds on the eigenvalues of the global mass matrix involve the eigenvalues of only the element mass matrices whereas the bounds on the eigenvalues of the global stiffness matrix involve the eigenvalues of both the element mass and stiffness matrices in addition to the fundamental frequency of the structure. Thus, for the stiffness matrix, the mass matrix needed for bounding its eigenvalues may be derived with any fictitious density distribution. This mass matrix has therefore nothing to do with the actual matrix that appears in the dynamic analysis of the structure. But once a fictitious density distribution has been decided upon the fundamental frequency of the structure appearing in the bounding expressions must refer to this particular density distribution. The simplest distribution is certainly the uniform one. Since the density distribution in this case is independent of the mesh so will the frequency be. Also, the fundamental frequency for this distribution has been computed and recorded [3] for a fairly large number of regularly shaped structures. These can be used for estimating [4] the fundamental frequencies of less regular structures.

Indeed, by using the technique of Ref. [1] with uniform density distribution we were able to obtain a substantial amount of information about the behavior of the condition

† Assistant Professor, Department of Mathematics.

numbers of the stiffness matrix of membranes and plates [5], nearly incompressible elastic solids [6] and arches and shells either thin [7, 8] or thick [9, 10].

It was observed, however, in these analyses that with a uniform density distribution the sharpness of the bounds decreased as the mesh became less and less regular (i.e. with large variations in the size of the elements). To sharpen the bounds in irregular meshes we introduce in this paper nonuniform, mesh dependent, density distributions. The fundamental frequency is thus no more mesh independent and we will derive some bounds on it.

In the case of regularly shaped domains, uniform meshes and symmetric (periodic) boundary conditions the stiffness matrix becomes regular enough for its eigenvalues to be computed algebraically. This has been used by Kelsey, Lee and Mak [11] to study the influence of the basic discretization and intrinsic parameters (such as mesh size and order of differential operator) on the spectral condition number of the stiffness matrix associated with such regular problems.

Bounds on the maximum (l_∞) norm of the finite element stiffness, flexibility and mass matrices were established by Descloux [12, 13] and in Refs. [6 and 14]. The difference between these two analyses being that whereas the bounds established by Descloux on the norms of the stiffness and flexibility matrices involve bounds on the l_∞ norms of the mass matrix and its inverse, the analysis of Refs. [6 and 14] does not involve the mass matrix at all. The role of the fundamental frequency in the spectral (l_2) norms is taken over, in Refs. [6 and 14], by the maximum of Green's (the influence) function of the structure. This restricts the applicability range of these l_∞ bounds since not all structures (membranes, for instance) can carry point loads and the maximum of Green's function becomes infinite. Here we replace the point loads by a load distributed over small circles (spheres in space) and study with this the dependence of the l_∞ norms of the stiffness and flexibility matrices, for two and three dimensional second order problems discretized with triangular and tetrahedronal elements, on the mesh parameters.

SPECTRAL (l_2) NORM AND CONDITION NUMBER

The method of finite elements reduces, like any other discretization method, the continuous boundary value problem and eigenvalue problem

$$Lu = f \quad \text{and} \quad Lu = \lambda \rho u \quad \text{in } D \tag{1}$$

where L is a linear (elliptic) differential operator, where $\rho \geq 0$ is the density distribution and where the function u has to satisfy some conditions on the boundary ∂D of D, into the algebraic problems

$$KU = F \quad \text{and} \quad KU = \mu MU \tag{2}$$

where K and M are the global stiffness and mass (Gram) matrices, respectively.

Since K and M are symmetric and at least positive semi-definite the spectral norms of the stiffness matrix K and flexibility matrix K^{-1} can be written as

$$\begin{aligned} \|K\|_2 &= \text{maximum eigenvalue of } K \\ \|K^{-1}\|_2 &= 1/\text{minimum eigenvalue of } K \end{aligned} \tag{3}$$

and the same is true for the mass matrix M. We denote by N the number of rows (columns) in K and M, and by λ_1^K, λ_N^K, λ_1^M and λ_N^M the lowest (1st) and highest (Nth) eigenvalues of

K and M. In this notation the spectral condition numbers $C_2(K) = \|K\|_2\|K^{-1}\|_2$ of K and $C_2(M) = \|M\|_2\|M^{-1}\|_2$ of M become

$$C_2(K) = \lambda_N^K/\lambda_1^K \quad \text{and} \quad C_2(M) = \lambda_N^M/\lambda_1^M. \tag{4}$$

We denote by k_e and m_e the element stiffness and mass matrices associated with the eth element. By u_e we denote the portion of U associated with the eth element. The quadratics U^TKU and U^TMU can be written then as

$$U^TKU = \sum_{e=1}^{Ne} u_e^T k_e u_e \quad \text{and} \quad U^TMU = \sum_{e=1}^{Ne} u_e^T m_e u_e \tag{5}$$

where summation is carried over all the Ne finite elements in the mesh.

If we denote by $\phi_1\phi_2, \ldots, \phi_n$ the n shape function in the element, the element quadratic $u_e^T m_e u_e$ is given by

$$u_e^T m_e u_e = \int_V \rho(u_1\phi_1 + u_2\phi_2 + \ldots + u_n\phi_n)^2 \mathrm{d}v \tag{6}$$

and hence for $\rho \geq 0$ (equality holding in one dimension only at distinct points, in two dimensions only on distinct lines and in three dimensions only on distinct surfaces), if the shape functions are linearly independent then m_e is positive definite for all $e = 1, 2, \ldots, Ne$.

By λ_1^k, λ_n^k, λ_1^m and λ_n^m we denote the lowest (1st) and highest (nth) eigenvalues of the element stiffness and mass matrices k and m. We have then that

$$k_1^e u_e^T u_e \leq u_e^T k_e u_e \leq \lambda_n^{k_e} u_e^T u_e \quad e = 1, 2, \ldots, Ne \tag{7}$$

and

$$\lambda_1^{m_e} u_e^T u_e \leq u_e^T m_e u_e \leq \lambda_n^{m_e} u_e^T u_e \quad e = 1, 2, \ldots, Ne. \tag{8}$$

For a normalized ($U^TU = 1$) vector U it can be shown that

$$1 \leq \sum_{e=1}^{Ne} u_e^T u_e \leq p_{\max} \tag{9}$$

where $p_{\max}$ denotes the maximum number of elements meeting at a nodal point.

To derive an upper bound on λ_N^K and λ_N^M we assume U in equation (5) to be the normalized eigenvector corresponding to λ_N^K and λ_N^M. Equations (7, 8 and 9) lead then to

$$\lambda_N^K \leq p_{\max} \max_e(\lambda_n^{k_e}) \quad \text{and} \quad \lambda_N^M \leq p_{\max} \max_e(\lambda_n^{m_e}). \tag{10}$$

In order to derive a lower bound on λ_N^K and λ_N^M we select U in equation (5) to be such that for the element where $\max_e(\lambda_n^{k_e})$ occurs, $u_e^T u_e = 1$. From this selection of U we get

$$\lambda_N^K \geq \max_e(\lambda_n^{k_e}) \quad \text{and} \quad \lambda_N^M \geq \max_e(\lambda_n^{m_e}). \tag{11}$$

Or

$$\max_e(\lambda_n^{k_e}) \leq \lambda_N^K \leq p_{\max} \max_e(\lambda_n^{k_e}) \tag{12}$$

and

$$\max_e(\lambda_n^{m_e}) \leq \lambda_N^M \leq p_{\max} \max_e(\lambda_n^{m_e}) \tag{13}$$

where e ranges over all the Ne finite elements in the mesh.

In an entirely analogous manner we obtain for the lowest eigenvalues of K and M

$$\min_e(\lambda_1^{k_e}) \leq \lambda_1^K \leq \lambda_N^K \tag{14}$$

and

$$\min_e(\lambda_1^{m_e}) \leq \lambda_1^M \leq \lambda_N^M. \tag{15}$$

It is appropriate to recall here a related theorem by Irons and Teharne [15] asserting that if

$$\omega \leq \frac{u_e^T k_e u_e^T}{u_e^T m_e u_e} \leq \Omega \tag{16}$$

for all $e = 1, 2, \ldots, Ne$, then also

$$\omega \leq \frac{U^T K U}{U^T M U} \leq \Omega \tag{17}$$

holds. But since the element stiffness matrix k is usually only positive semidefinite, both λ_1^k and ω are zero and the lower bound in both equations (14 and 17) is reduced to zero, expressing the mere trivial fact that $0 \leq \lambda_1^K \leq \lambda_N^K$.

To obtain non trivial bounds on λ_1^K we make use of the variational nature of the finite element method and Rayleigh's principle. It asserts that if λ_1 is the exact eigenvalue of the structure then for any $U \neq 0$

$$U^T K U / U^T M U \geq \lambda_1. \tag{18}$$

Choosing U in equation (18) to be the eigenvector corresponding to λ_1^K we obtain from equations (15 and 18) that

$$\lambda_1^K \geq \lambda_1 \min_e(\lambda_1^{m_e}). \tag{19}$$

For deriving an upper bound on λ_1^K we choose U to be the eigenvector corresponding to the lowest approximate finite element eigenvalue μ_1 such that

$$U^T K U / U^T M U = \mu_1 \tag{20}$$

which results, due to equation (13), in

$$\lambda_1^K \leq \mu_1 p_{\max} \max_e(\lambda_n^{m_e}). \tag{21}$$

Therefore

$$\lambda_1 \min_e(\lambda_1^{m_e}) \leq \lambda_1^K \leq \mu_1 p_{\max} \max_e(\lambda_n^{m_e}) \tag{22}$$

and a sufficient condition for $\|K^{-1}\|_2 = 1/\lambda_1^K$ to exist is that $\lambda_1^{m_e} > 0$ for $e = 1, 2, \ldots, Ne$. This is assured, however, by the linear independence of the shape functions.

With all that, the bounds on $C_2(K)$ and $C_2(M)$ become

$$\frac{\max(\lambda_n^k)}{\mu_1 \max(\lambda_n^m) p_{\max}} \leq C_2(K) \leq \frac{\max(\lambda_n^k) p_{\max}}{\lambda_1 \min(\lambda_1^m)} \tag{23}$$

and

$$1 \leq C_2(M) \leq \frac{\max(\lambda_n^m)p_{\max}}{\min(\lambda_1^m)}. \tag{24}$$

In equation (23) μ_1 is finite element approximation to λ_1. For a sufficiently fine mesh they will be close enough [16] for λ_1 to be substituted for μ_1.

The upper bound on $C_2(K)$ given in equation (23) was derived under the assumption that the finite elements strictly satisfy the continuity requirements posed by the variational principle. However, the upper bound on $C_2(K)$ will hold asymptotically for a sufficiently fine mesh, even if the continuity requirements are not strictly enforced (that is, the elements do not conform), provided that the lowest finite element eigenvalue converges. Then the exact eigenvalue λ_1 in equation (16) can be replaced by the corresponding approximate eigenvalue μ_1.

It should be noticed that if the density distribution over the structure is being chosen independently of the mesh then λ_1 (the fundamental frequency squared) will depend only on the shape of the structure (domain), the order of the differential equation involved, the boundary conditions, the density distribution and the (elastic) coefficients. The usefulness of a nonuniform density distribution is in drawing $\max(\lambda_n^m)/\min(\lambda_1^m)$ as close as possible to 1. This may tie λ_1 to the mesh parameters and the next section is devoted to this problem.

ESTIMATION OF THE FUNDAMENTAL FREQUENCY

The eigenvalues of the element mass matrix m depend on the geometry of the element, the degree of the polynomial in the shape functions and the density inside it. For a uniform mesh the most appropriate density distribution is uniform since then $\max(\lambda_n^{me})/\min(\lambda_1^{me})$ is uniform over all the Ne finite elements. In the case of a nonuniform mesh we choose the density so as to make λ_1^{me} equal over all the Ne elements. This will tie the fundamental frequency of the structure to the mesh layout. It may well happen that due to a regular variation of the mesh size the density distribution can also be chosen to vary regularly enough for the fundamental frequency to be directly computed. In the general case the mesh will be too irregular for a direct computation of the fundamental frequency and we need some universal bounds on it which we intend to develop now.

Let u be the eigenfunction corresponding to the minimum eigenvalue λ_1 such that

$$\lambda_1 = E(u, u)/(u, u) \tag{25}$$

in which $E(u, u)$ denotes the quadratic internal (say elastic) energy and where (u, u) denotes the kinetic energy explicitly defined by

$$(u, u) = \int_D \rho u^2 \, dv. \tag{26}$$

Let the total mass of the structure be one

$$\int_D \rho \, dv = 1. \tag{27}$$

We obtain, then, from equation (26) that

$$(u, u) \leq \max_D |u^2| \tag{28}$$

which prompts us to introduce the characteristic number Φ

$$\Phi = \min_u (E(u, u)/\max_D |u^2|) \tag{29}$$

and

$$\Phi \leq \lambda_1 \tag{30}$$

for any density distribution over the structure maintaining a unit mass. This Φ number is, in fact, the fundamental eigenvalue (frequency squared) of the structure with all the mass concentrated at the crest of the corresponding eigenfunction.

We denote by $G(x, \xi)$ the influence (Green's) function of the structure (i.e. the deflection at point ξ due to a unit point load at point x, or vice versa). The continuous eigenproblem can be written with this function in the form

$$u(x) = \lambda \int_D G(x, \xi)\rho(\xi)u(\xi)\, d\xi. \tag{31}$$

Suppose that x appearing in equation (31) is just the point at which $u(x)$ attains its maximum. If all the mass is concentrated at that point we get from equation (31) that

$$\Phi = 1/\Gamma, \Gamma = \max_D G(x, x). \tag{32}$$

For a string of unit tension and length L: $\Gamma = L/4$ and hence $\phi = 4/L$ whereas a uniform density yields $\lambda_1 = \pi^2/L$. This Φ is the fundamental eigenvalue of the string with all its mass concentrated at the center. No other density distribution with unit mass will result in a lower frequency. For a simply supported beam of uniform density, unit mass and unit bending flexibility: $\lambda_1 = \pi^4/L^3$, $\Gamma = L^3/48$ and consequently $\Phi = 48/L^3$. This Φ is, again, the fundamental eigenvalue of the beam with all its mass at the center and no other density distribution produces a lower eigenvalue. For a clamped circular plate of radius R, unit bending flexibility and unit mass: $\lambda_1 = 104\pi/R^2$ and $\phi = 16\pi/R^2$. By a well known theorem of Rayleigh this Φ is lower than the Φ number of any other clamped plate completely enclosed inside the circle (the proof to this theorem hinges on the twin facts that the displacements of the clamped plate can be extended up to the clamped circle and that the fundamental frequency minimizes the quotient in equation (25)).

The membrane and the three dimensional elastic solid can carry no point loads, $G(x, x) = \infty$ in these cases, and consequently the Φ number as defined in equation (29) vanishes. To obtain a nontrivial bound on λ_1 we must avoid in these cases the infinite density of the point mass. We therefore pose an addition restriction on ρ and define Φ as

$$\Phi = \min_{u,\rho} [E(u, u)/(u, u)] \tag{33}$$

with

$$0 \leq \rho \leq \rho_{\max} \quad \text{and} \quad \int_D \rho\, dv = 1. \tag{34}$$

We will compute now the Φ number for the problem

$$\begin{aligned} u_{xx}+u_{yy}+u_{zz}+\lambda u &= 0 \quad \text{in } D \\ u &= 0 \quad \text{on } \partial D \end{aligned} \tag{35}$$

where D is a spherical (circular) domain. This Φ number is inferior to the Φ number of any other domain with u vanishing on its surface completely enclosed inside the sphere. Multiplying equation (31) by $u(x)\,dx$ and integrating results in

$$\int_D u^2(x)\,dx = \lambda \int_D \int_D G(x,\xi)\rho(\xi)u(x)u(\xi)\,dx\,d\xi. \tag{36}$$

Green's function $G(x,\xi)$ is positive for the cases we consider and hence according to equation (36), $u(x)$ corresponding to the lowest eigenvalue λ_1 cannot be negative. Had it been negative at some area we could have replaced it by $-u(x)$ leaving the left hand integral in equation (36) unchanged but increasing the right hand one and consequently decreasing (the assumed lowest) eigenvalue λ_1. If we assume again that the point x in equation (31) is that at which $u(x)$ attains its highest value, then since $u(\xi) \geq 0$ we obtain from this equation that

$$1 \leq \lambda_1 \int_D G(x,\xi)\rho(\xi)\,d\xi. \tag{37}$$

The fundamental solutions of equation (35) with $\rho = 0$ are

$$\psi(r) = \frac{1}{2\pi}\log(1/r) \quad \text{and} \quad \psi(r) = \frac{1}{4\pi r} \tag{38}$$

for the circle and sphere, respectively. Green's function for these domains is given in terms of the fundamental solutions by

$$G(x,\xi) = \psi(r) - \psi\left(\frac{s}{R}p\right) \tag{39}$$

where in n dimensions

$$s^2 = \sum_{j=1}^{n} \xi_j^2, \qquad r^2 = \sum_{j=1}^{n} (x_j - \xi_j)^2, \qquad p^2 = \sum_{j=1}^{n} \left(x_j - \frac{R^2}{s^2}\xi_j\right)^2. \tag{40}$$

Since $G(r,\xi)$ is monotonic the density $\rho(\xi)$ maximizing the integral in equation (37) is constant inside one of the circular contours of $G(x,\xi)$. Let $u(x)$ be the value of the integral in equation (37), C the contour value, x_0 the center of the circular contour and a its radius, then with $z = e^{4\pi C}$ we have that

$$\begin{aligned} u(x) &= \frac{1}{2\pi}\log\left(\frac{1-xx_0}{a}\right) + \frac{1}{4\pi}\left[1 - \left(\frac{x-x_0}{a}\right)^2\right] \\ x_0 &= x(1-z)/(x^2-z), \qquad a = \sqrt{(z)(x^2-1)}/(x^2-z). \end{aligned} \tag{41}$$

The maximum of $u(x)$ occurs at the center $x = 0$ such that $G(0,r) = (1/2\pi)\log(R/r)$ and consequently

$$\Phi \geq \frac{4\pi}{1+2\log(R/a)}, \qquad \pi a^2 \rho_{\max} = 1 \tag{42}$$

where equality is attained for a sufficiently small a. For a sphere of radius R we have

$$G(0,r)=\frac{1}{4\pi}\left(\frac{1}{r}-\frac{1}{R}\right) \tag{43}$$

and therefore

$$\Phi \geq \frac{4\pi R}{3R/2a-1}, \qquad 4\pi a^3\rho_{\max}/3=1 \tag{44}$$

where equality is approached as a decreases.

If the density is distributed over the mesh in such a way that $\lambda_1^{m_e}$—the lowest eigenvalue of the eth element—is equal over all the finite elements in the mesh, then the lower bound on λ_1^K and the upper bound on $C_2(K)$ become

$$\lambda_1^K \geq \Phi\lambda_1^m \tag{45}$$

and

$$C_2(K) \leq \frac{\max(\lambda_n^{k_e})p_{\max}}{\Phi\lambda_1^m}. \tag{46}$$

SECOND ORDER PROBLEMS

We start analyzing this class of problems by carefully considering the simplest of them all—the fixed string discretized by linear finite elements. The relevant element stiffness and mass matrices are given by

$$k_e=\frac{1}{h_e}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix} \quad \text{and} \quad m_e=\frac{\rho_e h_e}{6}\begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix} \tag{47}$$

where h_e and ρ_e are element size and the density inside it, respectively. It is appropriate to point out here that λ_1^m and λ_n^k appearing in the upper bound expression on $C_2(K)$ in equation (23) are computed from k_e and m_e with no regard to the boundary conditions since any constraint on these matrices can only lower λ_n^k and raise λ_1^m. But λ_n^k in the lower bound expression on $C_2(K)$ in equation (23) must be computed from a constrained element matrix.

For a uniform density distribution $\rho_e=1$ we have that

$$\lambda_2^{k_e}=2/h_e, \qquad \lambda_1^{m_e}=h_e/6 \quad \text{and} \quad \lambda_2^{m_e}=h_e/2. \tag{48}$$

Hence with $\lambda_1=\pi^2$ and $\mu_1\cong\lambda_1=\pi^2$ we get from equations (12, 22 and 23) that

$$1/h_{\min}\leq\lambda_N^K\leq 4/h_{\min}, \qquad \pi^2 h_{\min}/6\leq\lambda_1^K\leq\pi^2 h_{\max} \tag{49}$$

and therefore

$$\frac{1}{\pi^2 h_{\min}h_{\max}}\leq C_2(K)\leq\frac{24}{\pi^2 h_{\min}^2} \tag{50}$$

where $h_{\min}$ and $h_{\max}$ are the extremal mesh sizes. For the mass matrix M we get

$$1\leq C_2(M)\leq 6h_{\max}/h_{\min}. \tag{51}$$

Varying the density according to

$$\rho_e h_e = 1/Ne \tag{52}$$

where Ne denotes the total number of finite elements in the mesh, we get from equation (45), recalling that for the fixed string of unit length $\Phi = 4$, that

$$C_2(K) \leq Ne/h_{\min}. \tag{53}$$

Next we consider the two dimensional membrane discretized with linear (first order) triangular elements. Here

$$k_{ij} = \frac{1}{2A} l_i l_j n_i . n_j \quad \text{and} \quad m = \frac{2A}{12}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \tag{54}$$

where l_i, A and n_i denote the ith side of the triangle, its area and a unit inward vector normal to the ith side. Assuming a uniform density $\rho = 1$ we get

$$\lambda_1^m = A/6, \qquad \lambda_3^m = 2A/3 \quad \text{and} \quad 1/\sin\theta \leq \lambda_3^k \leq 6/\operatorname{tg}\theta \tag{55}$$

where θ is the smallest angle in the triangle. For a sufficiently fine mesh ($\lambda_1 \cong \mu_1$) we thus obtain that

$$1/\sin\theta \leq \lambda_N^K \leq 6p_{\max}/\operatorname{tg}\theta, \qquad \lambda_1 A_{\min}/6 \leq \lambda_1^K \leq 2\lambda_1 p_{\max} A_{\max}/3 \tag{56}$$

and consequently

$$\frac{3}{2\lambda_1 \sin\theta p_{\max} A_{\max}} \leq C_2(K) \leq \frac{36 p_{\max}}{\lambda_1 \operatorname{tg}\theta A_{\min}} \tag{57}$$

in which $A_{\max}$ and $A_{\min}$ are the extremal element areas. With the density distribution

$$\rho_e A_e = 1/Ne \qquad e = 1, 2, \ldots, Ne \tag{58}$$

the first eigenvalue of m becomes $\lambda_1^m = 1/6Ne$ and hence according to equations (45 and 46)

$$\lambda_1^K \geq \frac{2\pi}{3Ne} \frac{1}{1+2\log(R/a)} \tag{59}$$

where

$$\pi a^2 = Ne A_{\min}. \tag{60}$$

Therefore

$$C_2(K) \leq \frac{9 p_{\max} Ne(1+2\log(R/a))}{\pi \operatorname{tg}\theta}. \tag{61}$$

The function $\log(R/a)$ grows rather slowly with R/a and the main source of ill conditioning with triangular elements is seen to be very acute angles rather than large mesh ratios. This is of prime importance for problems including singularities where steep variations in the mesh size are required [17] for assuring the full accuracy provided by the shape functions. These large variations in the mesh size will not introduce the feared decline in the condition of the stiffness matrix unless the passage from the small elements to the

larger one will necessitate exceptionally thin elements. Without the ill-conditioning argument one can hardly see why not use polynomial elements, properly spaced around the singular point instead of including singular terms into the shape functions, a procedure that certainly encumbers the finite element method.

The *a priori* estimates for λ_1^K, λ_N^K and $C_2(K)$ as given in equations (56, 57, 59 and 61) will be compared now with numerically computed values. The problem chosen for these tests is that of a circular (in fact polygonal) membrane fixed on its circumference and discretized with an irregular mesh of triangular element as shown in Fig. 1. The fundamental eigenvalue of a circular membrane of unit radius and with a uniform unit density

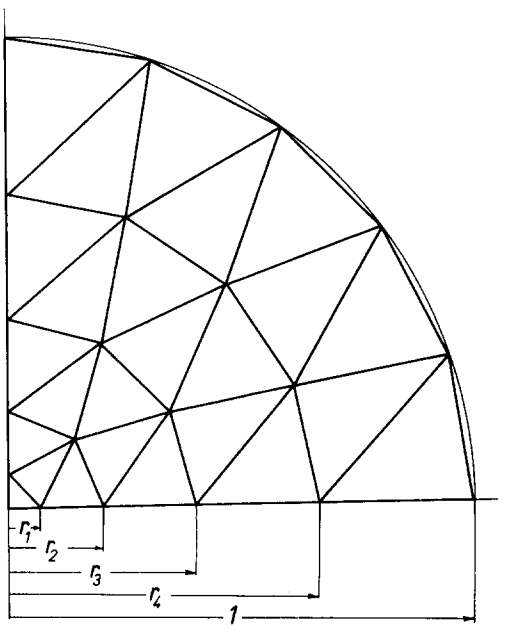

FIG. 1. Circular membrane discretized with triangular elements. Figure shows one quarter of the circle.

is approximately $\lambda_1 = 5{\cdot}7$ (the replacement of the circular boundary by a polygonal one results in only negligible [18] changes in the fundamental eigenvalue). First we consider linear elements. For these, distributed uniformly along the radius (that is, $r_i = i/5$ in Fig. 1) and with a uniform density we have *a priori* that

$$\min_e(\lambda_1^{m^e}) = 1/300, \qquad \lambda_1^K \geq 0{\cdot}02. \tag{62}$$

The bound on λ_N^K can be computed directly by Gerschgorin's theorem, yielding $\lambda_N^K \leq 15$ and consequently *a priori*

$$C_2(K) \leq 750, \qquad \log C_2(K) \leq 2{\cdot}88. \tag{63}$$

Computationally we obtain

$$\lambda_N^K = 11, \qquad \lambda_1^K = 0{\cdot}37, \qquad C_2(K) = 30, \qquad \log C_2(K) = 1{\cdot}5. \tag{64}$$

The number of decimals lost in the numerical solution of the algebraic system is roughly given [19] by $\log C_2(K)$ (for more on the relation between the condition number and the round-off errors see Refs. [20, 21 and 22]). Relative to (14 or 24) decimals available in the modern computers the *a priori* estimate given by equation (63) is, therefore, excellent.

Next we apply our estimate for λ_1^K, λ_N^K and $C_2(K)$ to the same membrane with the same elements distributed this time cubically along the radius (that is, $r_i = i^3/125$). Here equation (57) with a uniform mesh distribution yields $\lambda_1^K \geq 1{\cdot}8\,10^{-4}$. From Gerschgorin's theorem we get $\lambda_N^K \leq 20$ and hence *a priori*

$$C_2(K) \leq 10^5, \qquad \log C_2(K) \leq 5{\cdot}0. \tag{65}$$

Under the assumption that all the mass is concentrated near the center we get $\lambda_1^K \geq 1/300$, $\lambda_N^K \leq 20$ and consequently we have that *a priori*

$$C_2(K) \leq 6000 \quad \text{or} \quad \log C_2(K) \leq 3{\cdot}8. \tag{66}$$

Computationally we get

$$\lambda_1^K = 0{\cdot}14, \qquad \lambda_N^K = 15{\cdot}5, \qquad C_2(K) = 110, \qquad \log C_2(K) = 2{\cdot}01. \tag{67}$$

Again, both estimate (57) with a uniform density distribution and estimate (61) with a clustered mass give reasonably good and realistic bounds.

We have repeated the above calculations for a membrane discretized with triangular quadratic elements (with six nodal points, three at the vertices and three at the mid sides). For a uniform mesh of four elements along the radius we obtain computationally that

$$\lambda_N^K = 8{\cdot}1, \qquad \lambda_1^K = 0{\cdot}09, \qquad C_2(K) = 92, \qquad \log C_2(K) = 1{\cdot}96 \tag{68}$$

while for a mesh varying cubically along the radius we get

$$\lambda_N^K = 16{\cdot}4, \qquad \lambda_1^K = 1/30, \qquad C_2(K) = 480, \qquad \log C_2(K) = 2{\cdot}68 \tag{69}$$

and indeed the stiffness matrix remains well conditioned even with a strongly irregular mesh.

The most versatile three dimensional element is the tetrahedronal element. The first order element has four nodal points at the vertices and its element stiffness and mass matrices are given by

$$k_{ij} = \frac{1}{6V} s_i s_j n_i \cdot n_j \qquad i, j = 1, 2, 3, 4 \quad \text{and} \quad m = \frac{V}{10}\begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix} \tag{70}$$

where V denotes the volume of the element, s_i $i = 1, 2, 3, 4$ its four faces and n_i $i = 1, 2, 3, 4$ four unit inward vectors normal to the faces. The extremal eigenvalues of k and m are here (with $\rho = 1$)

$$\lambda_1^m = V/10, \qquad \lambda_4^m = V/2 \quad \text{and} \quad s_{\max}^2/6V \leq \lambda_4^k \leq 2s_{\max}^2/3V \tag{71}$$

resulting with $s_{\max}^2/V = H$ in

$$H_{\min}/6 \leq \lambda_N^K \leq 2p_{\max}H_{\max}/3 \tag{72}$$

$$\lambda_1 V_{\min}/10 \leq \lambda_1^K \leq \mu_1 p_{\max} V_{\max}/2 \tag{73}$$

or

$$\frac{H_{\max}}{3\mu_1 p_{\max} V_{\max}} \leq C_2(K) \leq \frac{20 p_{\max} H_{\max}}{3\lambda_1 V_{\min}} \tag{74}$$

and the stiffness matrix becomes invariably ill-conditioned as $H_{\max}$ increases and the element becomes thinner. With a nonuniform density distribution we get from equations (45, 46 and 42) that

$$C_2(K) \leq \frac{5 p_{\max} H_{\max} N e}{3\pi R}(3R/2a - 1) \tag{75}$$

where R is the radius of the sphere completely enclosing the domain and where

$$4\pi a^3 = 3 N e V_{\min}. \tag{76}$$

If $R \gg a$ equations (75 and 76) lead to the simpler bound

$$C_2(K) \leq \tfrac{4}{3} p_{\max} N e \tfrac{3}{2} \frac{H_{\max}}{V_{\min}^{\frac{1}{3}}}. \tag{77}$$

In Ref. [6] it has been shown that in the case of a nearly incompressible solid ($v \uparrow \frac{1}{2}$), the condition number of the stiffness matrix is bounded by

$$c_2 \frac{E}{\mu_1} \frac{h^{-2}}{1-2v} \leq C_2(K) \leq c_1 \frac{E}{\lambda_1} \frac{h^{-2}}{1-2v} \tag{78}$$

where h is the dia. of the element, E is the elastic modulus, v is Poisson's ratio, and c_1 and c_2 are two coefficients independent of h and v. According to equation (78) the condition of the stiffness matrix deteriorates (provided that $\mu_1 \cong \lambda_1$ and that λ_1 remains bounded as $v \uparrow \frac{1}{2}$). However, even if v is very near to $\frac{1}{2}$ the discretization error may not warrant the introduction of the exact v, it might any way be much higher than the error due to a sufficiently small change in v. Only a finer mesh with a lower discretization error will require a Poisson ratio nearer to $\frac{1}{2}$.

FOURTH ORDER PROBLEMS

The analysis of two dimensional plate problems discretized with rectangular or triangular elements is a bit too bulky for the present paper. We will be able nevertheless to get significant insight into the behavior of $C_2(K)$ in these problems by considering in detail a more accessible problem—that of a one dimensional beam discretized with cubic elements. The element stiffness and mass matrices for an element of size h are given by

$$k = \frac{1}{h^3}\begin{pmatrix} 12 & 6h & -12 & 6h \\ 6h & 4h^2 & -6h & 2h^2 \\ -12 & -6h & 12 & -6h \\ 6h & 2h^2 & -6h & 4h^2 \end{pmatrix} \text{ and } m = \frac{h}{420}\begin{pmatrix} 156 & 22h & 54 & -13h \\ 22h & 4h^2 & 13h & -3h^2 \\ 54 & 13h & 156 & -22h \\ -13h & -3h^2 & -22h & 4h^2 \end{pmatrix} \tag{79}$$

referring to u_1, u_{x1}, u_2 and u_{x2}. Assume first a uniform mesh and a uniform density $\rho = 1$. Suppose also that h is scaled out from the interior of k and m by replacing u_x by hu_x. With

this the extremal eigenvalues of k and m become

$$\lambda_4^k = \sqrt{(5)}12/h^3, \qquad \lambda_1^m = h/180 \quad \text{and} \quad \lambda_4^m = h/6 \tag{80}$$

yielding for the simply supported beam of length 1 (where $\lambda_1 = \pi^4$)

$$h^{-4} \leq C_2(K) \leq 86h^{-4}. \tag{81}$$

In the case of a nonuniform mesh we scale the matrix by changing the u_x nodal values into $h_a u_x$ where $h_a = (h_{\max} h_{\min})^{\frac{1}{2}}$. We thus obtain from equation (79), assuming that $\rho_e h_e = 1/Ne$, that

$$\lambda_1^m = z^2/180Neh_a^2, \qquad Ne/z^2 = \sum_{e=1}^{Ne} 1/h_e^2. \tag{82}$$

Consequently, since in this case $\lambda_N^K \leq 36/h_{\min}^3$, we get for the simply supported beam ($\Phi = 48$) that

$$C_2(K) \leq 135 \frac{Neh_{\max}}{h_{\min}^2 z^2}. \tag{83}$$

We could have obtained a sharper upper bound on $C_2(K)$ than that given in equation (83) by either a more careful scaling or by the addition of rotary inertia to go with the u_x nodal values. In the coming sections we will also derive a better bound on $C_2(K)$ via the maximum norms.

For the circular ring we obtained elsewhere [7] that

$$C_2(K) = c\left(\frac{r}{t}\right)^2 h^{-2} \tag{84}$$

where r/t denotes the radius of curvature to thickness ratio of the ring, h the size of the element and where c is independent of r, t and h. According to equation (84) the stiffness matrix K becomes ill-conditioned as the ring becomes thinner. However, in this case the extensional contribution to the solution which is proportional to t^2 diminishes too. On the other hand the total discretization error might be much higher than the extensional portion of the energy and there will be no need to introduce into the extensional energy the exact t. The thickness should be made to diminish simultaneously [8] with the mesh so as to balance the extensional energy and the discretization error.

In the same manner, by balancing the shear energy portion and the discretization error we were successful in constructing [9, 10] well conditioned C^0 plate bending elements.

UNIFORM MESHES; GENERAL RESULTS

Using the present bounding technique we are able to obtain for the case of a uniform mesh results of considerable generality.

We consider in this section problems associated with the operator

$$\nabla^{2m} = (\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2)^m.$$

Here for one element

$$(u, u) = \int_V u^2 \, dx \, dy \, dz \tag{85}$$

where $0 \leq x \leq h$, $0 \leq y \leq h$ and $0 \leq z \leq h$. With the mapping $x = h\xi$, $y = h\xi$ and $z = h\xi$, equation (75) is changed into

$$(u, u) = h^v \int_{V'} u^2 \, d\xi \, d\eta \, d\zeta \tag{86}$$

in which v is the dimension and $0 \leq \xi \leq 1$, $0 \leq \eta \leq 1$ and $0 \leq \zeta \leq 1$. If the nodal values are scaled in such a way that h is eliminated from the interior of the element matrices (h appears inside these matrices when the nodal values include derivatives of various degree) then

$$\lambda_1^m = c_1 h^v \quad \text{and} \quad \lambda_n^m = c_2 h^v \tag{87}$$

c_1 and c_2 being independent of h.

The element stiffness matrices are obtained from

$$E(u, u) = h^{v-2m} \int_{V'} \tilde{\nabla}^m u \, . \, \tilde{\nabla}^m u \, d\xi \, d\eta \, d\zeta \tag{88}$$

where

$$\tilde{\nabla}^m = (\partial^m/\partial\xi^m, \partial^m/\partial\eta^m, \partial^m/\partial\zeta^m). \tag{89}$$

Thus if h is scaled out from the interior of k, equation (78) yields

$$\lambda_n^k = c_3 h^{v-2m} \tag{90}$$

and consequently

$$c_5 h^{-2m} \leq C_2(K) \leq c_4 h^{-2m} \tag{91}$$

where c_4 and c_5 are independent of h, and where h was assumed small enough for μ_1 to be replaced by λ_1.

Equation (91) is of far reaching consequences. It indicates that whatsoever the degree of the shape polynomial and whatsoever the type of nodal values, the rate at which the condition number grows depends, at least in the case of a regular mesh, solely on the order of the problem. Fourth order problems are thus liable to be more ill-conditioned than second order problems. Also, since $C_2(K)$ grows with h, for the same number of elements one dimensional problems are likely to be more ill-conditioned than two and three dimensional problems. More important, since the rate at which $C_2(K)$ increases with the mesh refinement does not depend on the degree of shape polynomials, one is able to obtain a better total (discretization + round-off) accuracy by the use of higher order elements. This has indeed been confirmed numerically in Refs. [1, 7 and 22].

MAXIMUM NORMS

The maximum (l_∞) norm $\|K\|_\infty$ of the matrix $K(N \times N)$ is defined by

$$\|K\|_\infty = \max_i \sum_{j=1}^{N} |K_{ij}|. \tag{92}$$

Correspondingly the maximum condition number is defined by

$$C_\infty(K) = \|K\|_\infty \|K^{-1}\|_\infty \tag{93}$$

and since $\|K\|_\infty \geq \|K\|_2$, $C_\infty(K) \geq C_2(K)$ and a bound on $C_\infty(K)$ constitute a bound also on $C_2(K)$.

It is simple to bound $\|K\|_\infty$. In fact, if the maximal number of elements meeting at a point is $p_{\max}$ then each row of K is assembled from no more than $p_{\max}$ elements and we have that

$$\|K\|_\infty \leq p_{\max} \max_e \|k_e\|_\infty \quad e = 1, 2, \ldots, Ne. \tag{94}$$

In order to bound $\|K^{-1}\|_\infty$ we again make use of the variational nature of the finite element method. The finite element solution is obtained by minimizing the total potential energy $\pi(u)$, that can be written in the general form

$$\pi(u) = E(u, u) - (f, u) \tag{95}$$

where $E(u, u)$ is the internal (say elastic) energy and where (f, u) is the work of the forces f. Let $\hat{u}$ denote the finite element solution; then

$$\pi(\hat{u}) - \pi(u) = E(u - \hat{u}, u - \hat{u}) \geq 0. \tag{96}$$

Since the first variation of $\pi(u)$ vanishes at the solution $\hat{u}$ (or u) we have that

$$E(\hat{u}, \hat{u}) = \tfrac{1}{2}(f, \hat{u}) \quad \text{and} \quad \pi(\hat{u}) = -\tfrac{1}{2}(f, \hat{u}) \tag{97}$$

and since $E(u, u)$ is positive definite, a positive point force results in a positive displacement at the point of application. It also results from equations (96 and 97) that

$$(f, u) \geq (f, \hat{u}). \tag{98}$$

Choosing f to be a point unit force (impulse) at P we get that

$$u(P) \geq \hat{u}(P) \tag{99}$$

which means that at the point of application of the load the variational (finite element, Rayleigh–Ritz) solution never exceeds the exact solution.

Let $G(x, \xi)$ be the influence (Green's) function of the structure. The finite element response at nodal point j to an impulse at nodal point i is K_{ij}^{-1} while the response at the point of application itself is K_{ii}^{-1}. We conclude therefore from equation (99) that (the existence of K^{-1} is assured in the l_2 norm and hence also in the l_∞ norm)

$$\max_{1 \leq i \leq N} (K_{ii}^{-1}) \leq \Gamma, \qquad \Gamma = \max_x G(x, x). \tag{100}$$

It is also readily shown that for a positive definite matrix K of dimension N

$$\|K\|_\infty < N \max_{1 \leq i \leq N} (K_{ii}), K_{ii} > 0. \tag{101}$$

Combining equations (92, 100 and 101) we get that

$$C_\infty(K) < N\Gamma p_{\max} \max_e \|k_e\|_\infty \tag{102}$$

which was also obtained in Ref. [14]. Notice that $\Gamma = \max G(x, x)$ is an intrinsic property of the structure. It plays the role of the fundamental eigenvalue in the l_2 norms. However, a large Γ means a flexible structure whereas a large λ_1 means a stiff structure.

A remarkable feature of the bound in equation (102) is that it does not include estimates for the mass matrix and we therefore expect from it better bounds than from the l_2

estimates. On the other hand Γ might well become infinite for some problems or for some choice of nodal variables making the l_∞ estimates less general. A possibility to overcome this is by spreading the forces in equation (98) over small areas rather than concentrating them at a point and we will soon employ this device.

The advantages and limitations of equation (102) will become apparent when applied to some specific problems. We start with the fixed string whose element matrices are given in equation (47). Here

$$\max\|k\|_\infty = 2/h_{\min}, \qquad \Gamma = \tfrac{1}{4} \quad \text{and} \quad p_{\max} = 2 \tag{103}$$

and therefore according to equation (102)

$$C_\infty(K) < N/h_{\min} \tag{104}$$

similarly to the l_2 bound in equation (40). This is a rather simple case since for an impulse the finite element solution coincides here with the true solution. The pattern of entries in K^{-1} can be readily deduced from this: (a) all entries in K^{-1} are positive (maximum principle), (b) the largest diagonal entry is at the center and is bounded below by $\frac{1}{4}$ for any number of elements and any mesh ratio, (c) K^{-1} is dense (this denseness expresses the fact that a disturbance is an elliptic system as caused by the application of a point force propagates to all points) and (d) the off-diagonal entries in K^{-1} decrease linearly as one moves away from the diagonal.

With regard to K^{-1} there can be a fundamental difference between elements having only u as nodal values and those that include also the derivatives of u. For the latter the influence function $G(x, \xi)$ in equation (100) should be replaced by its derivatives i.e. the response not only to unit forces but also to unit torques. But not all structures, even if able to carry point load, are able to carry point torques and the bound in equation (102) fails. This suggests that in this case the diagonal entries in K^{-1} corresponding to the higher derivatives do not remain bounded (unless scaled) as the mesh is being refined.

Also for the Laplace equation in two and three dimensions equation (102) is not directly applicable since the fundamental solution for these problems is $\log(1/r)$ and $1/r$ respectively, and the origin $r = 0$ is a problem. We can, however, gain insight into the behavior of $\|K\|_\infty$ and $\|K^{-1}\|_\infty$ in the case of a two or three dimensional Laplace equation by computing these norms directly from the stiffness and flexibility matrices. To this end we consider a circular and spherical domains with $u = 0$ on the boundary. For the membrane we have

$$k_e = \frac{2e-1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \tag{105}$$

and hence

$$K = \frac{1}{2}\begin{pmatrix} 1 & -1 & & & \\ -1 & 4 & -3 & & \\ & -3 & 8 & -5 & \\ & & \ddots & \ddots & \ddots \end{pmatrix}. \tag{106}$$

The maximum deflection (in fact, $\max K_{ii}^{-1}$) due to a point force is given here by

$$u_{\max} = 2\left(1 + \frac{1}{3} + \frac{1}{5} + \ldots + \frac{1}{2N-1}\right) \tag{107}$$

such that for a large N, $u_{max} = 2 \log N$. The sum of the (positive) displacements along the radius (sum of K_{ij}^{-1} over j) renders

$$2\left(1+\frac{2}{3}+\frac{3}{5}+\ldots+\frac{N}{2N-1}\right). \tag{108}$$

Or

$$\|K^{-1}\|_\infty = N, \|K\|_\infty = 2N \quad \text{and} \quad C_\infty(K) = 2N^2. \tag{109}$$

In the same manner we obtain for the sphere

$$u_{max} = N(1+\tfrac{1}{7}+\tfrac{1}{19}+\tfrac{1}{37}+\ldots) \tag{110}$$

or $u_{max} = N$. The sum of the displacements along the radius is given by

$$N(1+\tfrac{2}{7}+\tfrac{3}{19}+\tfrac{4}{37}+\ldots).$$

From which

$$\|K^{-1}\|_\infty = N \log N, \|K\|_\infty = 9N \quad \text{and} \quad C_\infty(K) = 9N^2 \log N \tag{111}$$

is obtained.

It is interesting to notice the difference between the behavior of K_{ii}^{-1} in the case of a string (with bounded Green's function) and in the case of an elastic solid, or membrane (with unbounded Green's function). In the first case K_{ii}^{-1} remains bounded as the mesh is being refined while in the latter two cases it grows like $\log N$ in two dimensional problems and like N in three dimensional problems, in analogy with the corresponding fundamental solutions $\log(1/r)$ and $1/r$.

From the element matrices in equations (54 and 70) it is obvious that the (positive definite) stiffness matrix for the Laplace equation discretized with first order (linear) triangular and tetrahedral elements will be such that $K_{ij} \leq 0$ for all $i \neq j$, provided that the elements have acute angles. It is a well known theorem that in this case $K_{ij}^{-1} > 0$ for all i, j. We will make use of this fact to derive bounds on $C_\infty(K)$ for matrices generated from such elements. Consider then the Laplace equation with $u = 0$ on the boundary discretized by first order acute triangular or tetrahedronal elements. Inside each element we draw a circle (sphere) of radius a centered around the element's center of area and assume a uniform load $1/\omega$ over it, where ω denotes the area (volume) of the inscribed circle. Let u and $\hat{u}$ be the exact and finite element solutions, respectively, due to this load. Equation (98) yields in this case

$$\frac{1}{\omega}\int_\omega u \, dv \geq \frac{1}{\omega}\int_\omega \hat{u} \, dv \tag{112}$$

or since $\hat{u}$ is linear inside each element equation (112) is reduced to

$$\frac{1}{\omega}\int_\omega u \, dv \geq \hat{u}_c \tag{113}$$

where the subscript c stands for center of area (gravity). Also, for the triangular element with the three nodal points 1, 2 and 3

$$\hat{u}_c = \tfrac{1}{3}(\hat{u}_1+\hat{u}_2+\hat{u}_3). \tag{114}$$

The element load vector b for each triangular element is

$$b = \frac{1}{3}\begin{pmatrix}1\\1\\1\end{pmatrix} \tag{115}$$

and consequently

$$\hat{u}_c = \tfrac{1}{9}(K_{11}^{-1} + K_{22}^{-1} + K_{33}^{-1} + 2K_{12}^{-1} + 2K_{13}^{-1} + 2K_{23}^{-1}). \tag{116}$$

Since $K_{ij}^{-1} > 0$, equation (116) leads to

$$\hat{u}_c \geq \tfrac{1}{9} \max_i (K_{ii}^{-1}). \tag{117}$$

To estimate the integral in equation (113) we construct a circle completely enclosing the original domain. Since u can be extended from the boundary of the original domain to the circle by $u = 0$, the total potential energy of the circle will be lower than that of the original domain. If the radius of the encircling disc is R and the loaded circle inside the element is a, we obtain that

$$\hat{u}_c \leq \frac{1}{8\pi}[1 + 4\log(R/a)] \tag{118}$$

and consequently

$$\|K^{-1}\|_\infty < \frac{9N}{8\pi}[1 + 4\log(R/a)]. \tag{119}$$

Since here

$$\|K\|_\infty \leq 6p_{\max}/\operatorname{tg}\theta \tag{120}$$

we finally get that

$$C_\infty(K) < 2p_{\max} N \operatorname{ctg}\theta[1 + 4\log(R/a)] \tag{121}$$

θ being the smallest angle and a the radius of the smallest circle inside the elements, which is centered around their center of gravity.

For the case of three dimensional problems discretized with first order acute tetrahedron we obtain in the same manner that

$$\hat{u}_c \leq \frac{1}{4\pi R}\left(\frac{6R}{5a} - 1\right) \tag{122}$$

or

$$\max(K_{ii}^{-1}) \leq \frac{4}{\pi R}\left(\frac{6R}{5a} - 1\right). \tag{123}$$

Thus

$$\|K^{-1}\|_\infty < \frac{4N}{\pi R}\left(\frac{6R}{5a} - 1\right). \tag{124}$$

Also, according to equation (71)

$$\|K\|_\infty \leq \tfrac{2}{3} p_{\max} H_{\max} \tag{125}$$

and therefore

$$C_\infty(K) < \frac{N}{R} p_{\max} H_{\max} \left(\frac{6R}{5a} - 1 \right) \tag{126}$$

a being again the radius of the smallest sphere inscribed inside the element and centered around its center of gravity.

Equation (126) suggest that in the l_∞ norm the stiffness matrix for the three dimensional problem be more illconditioned than that for the two dimensional case. This conclusion is also implied in equations (109 and 111) obtained by direct calculations. Nevertheless, even if the bound on K_{ii}^{-1} is not too pessimistic the bound on $\|K^{-1}\|_\infty$ in equation (124) obtained from equation (101) could not be that good since in three dimensional problems the off-diagonal entries in the stiffness matrix vary faster (somewhat like $1/r$) than those in the corresponding two dimensional problems (which vary somewhat like $\log(1/r)$).

The beam element whose matrices are given in equation (79) is associated with the a nodal displacement and a nodal rotation. To bound K_{ii}^{-1} corresponding to these values we need the exact response of the beam to both a unit force and a unit torque. In both cases Γ is bounded and independent of h and we get from equation (102)

$$C_\infty(K) < 48 \Gamma N h_{\min}^{-3} \tag{127}$$

even without any scaling.

Knowing a bound on $C_2(K)$ we can bound $C_\infty(K)$ through the inequality

$$\|K^{-1}\|_\infty \leq N^{\frac{1}{2}} \|K^{-1}\|_2 \tag{128}$$

leading readily with equation (91) to

$$C_\infty(K) \leq c h^{-2m-\nu/2} \tag{129}$$

where ν denotes the dimension. For the string ($m = 1$, $\nu = 1$) equation (129) yields $C_\infty(K) \leq c h^{-2.5}$ which is rather pessimistic.

MIXED VARIATIONAL PRINCIPLES

To avoid the high continuity requirements in fourth (or higher) order problems they can be decomposed into a system of lower order equations with extra unknowns. As an example to this consider the beam equation

$$\frac{d^4 u}{dx^4} = f(x) \qquad 0 \leq x \leq 1. \tag{130}$$

It can be separated into the two equations

$$\frac{d^2 u}{dx^2} = M \quad \text{and} \quad \frac{d^2 M}{dx^2} = f \tag{131}$$

where M is proportional to the bending moment. Typical boundary conditions for the beam are:

$$\begin{aligned} &(1)\quad u = 0, M = 0 && \text{simply supported end,}\\ &(2)\quad u = 0, \mathrm{d}u/\mathrm{d}x = 0 && \text{clamped end,}\\ &(3)\quad M = 0, \mathrm{d}M/\mathrm{d}x = 0 && \text{free end.} \end{aligned} \tag{132}$$

We associate with equations (131) and (132) the quadratic functional $\pi_R(u, M)$

$$\pi_R(u, M) = \int_0^1 \left(\frac{1}{2}M^2 + \frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}M}{\mathrm{d}x} + uf\right)\mathrm{d}x. \tag{133}$$

Setting the first variation of $\pi_R(u, M)$ equal to zero we obtain as Euler's equations the two equations in (131) plus the following natural boundary conditions: (a) Where M is not prescribed $\mathrm{d}u/\mathrm{d}x = 0$. (b) Where u is not prescribed $\mathrm{d}M/\mathrm{d}x = 0$. Obviously, it is sufficient for both u and M in equation (133) to be only continuous and for a finite element scheme in which u and M vary linearly inside the element we get the global system

$$\begin{pmatrix} 0 & D_2 \\ D_2^T & G \end{pmatrix}\begin{pmatrix} u \\ M \end{pmatrix} = \begin{pmatrix} F \\ 0 \end{pmatrix} \tag{134}$$

where

$$D_2 = h^{-2}\begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & \ddots & \ddots & \ddots \end{pmatrix} \quad \text{and} \quad G = \frac{1}{6}\begin{pmatrix} 2 & 1 & & & \\ 1 & 4 & 1 & & \\ & 1 & 4 & 1 & \\ & & \ddots & \ddots & \ddots \end{pmatrix}. \tag{135}$$

In practice the entries in equation (134) are ordered in a band form but for analytic reasons we find it convenient to keep the ordering of equation (134). We notice also that D_2 is the second order difference operator or the stiffness matrix of the string whereas G is its mass matrix.

To find the condition of the global matrix in equation (134) we have to solve

$$\begin{pmatrix} 0 & D_2 \\ D_2^T & G \end{pmatrix}\begin{pmatrix} u \\ M \end{pmatrix} = \lambda\begin{pmatrix} u \\ M \end{pmatrix} \tag{136}$$

which after the elimination of u from it is reduced to

$$D_2^T D_2 M = (\lambda^2 I - \lambda G)M. \tag{137}$$

Since D_2 is a second order difference operator, $D_2^T D_2$ is a fourth order operator with eigenvalues ranging from 0(1) to $0(h^{-4})$. Hence, since G is almost an identity matrix (in the sense that its eigenvalues are bounded), the range of λ is computed from the equations

$$1 = \lambda^2 - \lambda \quad \text{and} \quad h^{-4} = \lambda^2 - \lambda \tag{138}$$

yielding that the condition number of the global matrix in equation (135) varies like $0(h^{-2})$.

Thus, by decomposing the biharmonic equation into a coupled system of second order equations we were able to reduce the condition number of the global matrix from $0(h^{-4})$ to $0(h^{-2})$. For the same number of elements the global matrix of the mixed principle is better conditioned than the corresponding global stiffness matrix. But that is not yet the whole story. It remains to be seen how the discretization error behaves in both cases and how many elements are required in each case for a comparable discretization accuracy. Indeed, some numerical experiments [1] suggest that even though the condition of the mixed matrix grows slower than the condition of the stiffness matrix, the higher rate of convergence of the latter makes it possible to obtain with it the same discretization accuracy as with the mixed principles with less elements. Therefore, at final counts the total (discretization + round-off) error in the displacement method might not be less than that of the mixed method.

The indefiniteness of the global matrix with mixed principles is also a drawback, but an obvious case where separation is useful is where the systems completely decouple as with simply supported beams or polygonal plates.

REFERENCES

[1] I. FRIED, Discretization and Round-off Errors in the Finite Element Analysis of Elliptic Boundary Value Problems and Eigenvalue Problems. Ph.D. Thesis, Massachusetts Institute of Technology (June, 1971).

[2] I. FRIED, Condition of finite element stiffness matrices generated from non-uniform meshes. *AIAA Journal* **10** (2), 219–221 (1972).

[3] A. W. LEISSA, Vibration of Plates. NASA *report* SP–160 (1969).

[4] G. POLYA and G. SZEGÖ, Isoperimetric inequalities in mathematical physics. *Ann. Maths Studies* **27**, Princeton University Press (1951).

[5] I. FRIED, Bounds on the extremal eigenvalues of the finite element stiffness and mass matrices and their spectral condition number. *J. Sound Vibration* **22** (4), 407–418 (1972).

[6] I. FRIED, Influence of Poisson's ratio on the condition of the finite element stiffness matrix. *Int. J. Solids Struct.* **9**, 323–329 (1973).

[7] I. FRIED, Basic computational problems in the finite element analysis of shells. *Int. J. Solids Struct.* **7** (12), 1705–1715 (1971).

[8] I. FRIED, Shape functions and the accuracy of arch finite elements. *AIAA J.* **11** (3), 287–297 (1973).

[9] I. FRIED, Shear in C^0 and C^1 bending finite elements. *Int. J. Solids Struct.* **9** (4), 449–460 (1973).

[10] I. FRIED and SHOK KENG YANG, Triangular, nine degrees of freedom C^0 plate bending element of quadratic accuracy. *Q. Appl. Math.* (To appear).

[11] S. KELSEY, K. N. LEE and C. K. K. MAK, The condition of some finite element coefficient matrices, computer-aided engineering. *Proceedings of the Symposium held at the University of Waterloo*, Canada. (Edited by G. M. L. GLADWELL), pp. 267–283 (1971).

[12] J. DESCLOUX, On some finite element matrices. *Siam J. Numerical Anal.* **9** (2), 260–265 (1972).

[13] J. DESCLOUX, Finite elements and numerical stability. EQUADIFF III Brno, 28 August–2 September 1972.

[14] I. FRIED, The l_2 and l_∞ Condition numbers of the finite element stiffness and mass matrices and the pointwise convergence of the method. *Conference on the Mathematics of Finite Elements and Applications*, Brunel University (18–20 April, 1972).

[15] B. M. IRONS and G. TREHARNE, A bound theorem in eigenvalues and its practical applications. *3rd Conference on Matrix Methods in Structural Mechanics*, Wright-Patterson Air Force Base, Ohio (1971).

[16] I. FRIED, Accuracy of finite element eigenproblems. *J. Sound Vibration* **18** (2), 289–295 (1971).

[17] I. FRIED and SHOK KENG YANG, Best finite element distribution around a singularity. *AIAA J.* **10** (9), 1244–1246 (1972).

[18] I. FRIED, Boundary and interior approximation errors in the finite element method. (To appear).

[19] F. L. BAUER, Genauigkeitsfragen bei der Losung Linearer Gleichungssysteme. *Zamm* **46** (7), 409–421 (1966).

[20] J. R. ROY, Numerical error in structural solutions. *J. Struct. Div. Proc. ASCE* 1039–1054 (1971).

[21] Y. YAMAMOTO, Some considerations on round-off errors of the finite element method. In *Advances in Computational Methods in Structural Mechanics and Design* (Edited by J. T. ODER, R. W. CLOUGH and Y. YAMAMOTO), University of Alabama Press (1972).

[22] I. FRIED, Discretization and computational errors in high-order finite elements. *AIAA J.* **9**, 2071–2073 (1971).

(*Received* 25 *September* 1972; *revised* 17 *January* 1973)

Абстракт—Учреждаются верхние и нижние пределы спектральных и максимальных норм [и следовательно соответсвующих чисел состояния] для матриц жесткости, гибкости [обратная величина жесткости] и массы, образованных из регулярных и нерегулярных сеток конечных элементов. Выводятся выражения в явной форме для этих пределов, в виде собственных параметров и параметров раздела. Эти выражения касаются задач второго и четвертого порядков, в одном, двух или трех размерах, для дискретных элементов линейных, трехугольных и четырехгранных.