Difference Equations on a Mesh Arising from a General Triangulation

By R. B. Kellogg

1. Introduction. Consider the boundary value problem

$$(1) \qquad \begin{cases} Lu \equiv -(pu_x)_x - (pu_y)_y + qu = f \\ u = 0 \quad \text{on } \partial R \end{cases}$$

in a domain R with polygonal boundary ∂R. The coefficients p, q are assumed positive, bounded, and bounded away from 0. It may be shown [1, p. 20] that for f square integrable, (1) has a unique "generalized solution" $u \in H_0^1(R)$. (The notation is given in §2.) It may be conjectured that if p is smooth enough, u has generalized derivatives of the second order and $\| u \|_2 \leqq c \| f \|$. (In [3, p. 665] such a result is given if ∂R is sufficiently smooth.)

We consider a class of finite difference approximations of (1),

$$(2) \qquad L_1 v = f_1,$$

in which the mesh points of the approximation are the vertices of any triangulation of R by acute triangles. These difference approximations were first considered by MacNeal [2] and include as a special case the usual 5 point difference approximation [5, chapter VI] to (1). It will be shown that, if $u \in H_0^1(R) \cap H^2(R)$ is a solution of (1), a mean square norm of the error, $u - v$, is bounded by $c'h\| u \|_2$, where c' is an explicit constant and h is the maximum distance between neighboring mesh points.

This result contrasts with that of Nitsche and Nitsche [4] who obtain an $0(h^{2/5})$ error estimate of the maximum norm of $u^* - v$ for more general second order elliptic equations and more special difference approximations. (u^* is a certain average of u.)

In the theories of heat conduction and neutron diffusion it is important to let p, q be discontinuous. Let p, q be smooth in the closure of each of a finite number of subdomains R_i which make up the domain R. It is required that at each interface ∂R_i, the solution u satisfies

$$(3) \qquad u,\ p\partial u/\partial n \quad \text{continuous across } \partial R_i,$$

where n is the normal vector at ∂R_i. If u is twice differentiable in each R_i and satisfies (1), (3), then for any $\phi \in H_0^1(R)$,

$$(4) \qquad \iint_R \{p\phi_x u_x + p\phi_y u_y + q\phi u - f\phi\}\, dx\, dy = 0,$$

so u is the generalized solution whose existence is shown in [1]. The proofs in this paper are valid if $u \in H^2(T)$ where T is any triangle in the triangulation which gives rise to the finite difference approximation (2). One may conjecture that the unique generalized solution $u \in H_0^1(R)$ of (4) is in $H^2(R_i)$ for each subdomain R_i.

Received February 13, 1963, revised October 8, 1963.

If this is true and if the ∂R_i are polygons, then the results of this paper apply if the triangulation contains the polygons ∂R_i.

2. The Difference Equations. If u is a function on a domain U, let $\| u, U \| = \left\{ \int_U | u |^2 \, dx \, dy \right\}^{1/2}$. Define $\| u, U \|_1^2 = \| u, U \|^2 + \| u_x, U \|^2 + \| u_y, U \|^2$, $\| u, U \|_2^2 = \| u, U \|_1^2 + \| u_{xx}, U \|^2 + \| u_{xy}, U \|^2 + \| u_{yy}, U \|^2$. $H(U)$, $H^1(U)$, $H^2(U)$ will denote the closure under the corresponding norms of the set of functions infinitely differentiable in a neighborhood of $\bar{U}$. These are Hilbert spaces. $H_0(U)$, $H_0^1(U)$, $H_0^2(U)$ will denote the closed subspaces spanned by those infinitely differentiable functions which vanish outside a compact subset of U. The usual properties of these spaces will be assumed. In particular two simple inequalities should be noted. Namely

$$(5) \qquad \begin{cases} | u(P) | \leqq c_1 \| u, U \|_2, & U \in H^2(U) \\ \int | u | \, ds \leqq c_2 \| u, U \|_1, & U \in H^1(U). \end{cases}$$

In these inequalities U is a triangle and c_1, c_2 depend only on U. P is a vertex of U and, in the second inequality, the left side is a line integral taken along a line segment in U. From the first inequality it is seen that the $u(P)$ are meaningful quantities for our generalized solutions.

When $U = R$ we omit the U in the above norms and spaces.

Let $\mathfrak{T}$ be a triangulation of R such that the sides of the polygons ∂R, ∂R_i, all lie on the lines of $\mathfrak{T}$, and such that there are no obtuse triangles in $\mathfrak{T}$. Let $\mathcal{S}$ be the set of vertices of $\mathfrak{T}$, and let $\mathcal{S}_0$ denote the points of $\mathcal{S}$ lying inside R. Let there be N points of $\mathcal{S}_0$. We will say that two points of $\mathcal{S}$ are neighbors if they are both vertices of a triangle of $\mathfrak{T}$.

Let $\rho(P, Q)$ be the distance between points P and Q, and let $h = \max \rho(P, Q)$, the maximum being taken over all neighbors $P, Q \in \mathcal{S}$. Let $c_3 > 1$ be a constant such that

$$(6) \qquad c_3^{-1} \leqq \rho(A, B)/\rho(C, D) \leqq c_3$$

for each point $P \in \mathcal{S}$, where A, B, C, D range over the set consisting of P and its neighbors. The error bound will depend upon c_3, which may be thought of as a "local maximum mesh ratio". Let $h(P)$ be the maximum distance from P to any one of its neighbors.

Let $\mathcal{C}$ be the collection of all real valued mesh functions on $\mathcal{S}$, and let $\mathcal{C}_0 \subset \mathcal{C}$ consist of those functions vanishing outside $\mathcal{S}_0$. Then $\mathcal{C}_0$ is an N dimensional vector space and L_1 will be an N by N matrix acting on C_0. We introduce two inner products on $\mathcal{C}_0$. If $\alpha, \beta \in \mathcal{C}_0$, these are defined by

$$(\alpha, \beta) = \sum h(P)^2 \alpha(P) \beta(P),$$

$$(\alpha. \beta)_1 = (\alpha, \beta) + \sum\nolimits_1 (\alpha(P) - \alpha(Q))(\beta(P) - \beta(Q)).$$

The sum $\sum$ is taken over all $P \in \mathcal{S}$ and the sum $\sum_1$ is taken over all neighboring points $P, Q \in \mathcal{S}$. The corresponding norms are denoted by $\| \alpha \|$ and $\| \alpha \|_1$.

Now let $\mathfrak{T}(P)$ be the set of triangles in $\mathfrak{T}$ with $P \in \mathcal{S}_0$ as a vertex. Let $T \in \mathfrak{T}(P)$

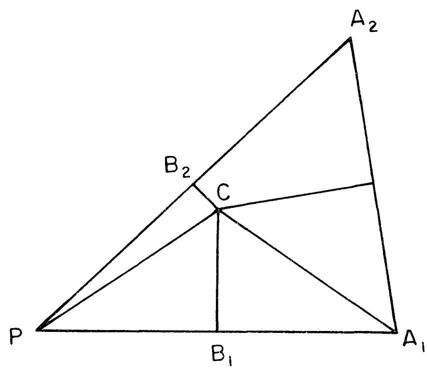

FIG. 1.

have vertices P, A_1, A_2, and let B_1C, B_2C be the perpendicular bisectors of PA_1, PA_2 (see figure 1). Since T is acute, C lies in T. Let U denote the quadrilateral defined by PB_1CB_2. We define functions $a_i(P, T)$, $b(P, T)$, $f_1(P, T)$ by the equations

$$a_i(P, T) = \frac{1}{\rho(P, A_i)} \int_{B_iC} p\, ds, \qquad i = 1, 2$$

$$b(P, T) = \iint_U q\, dx\, dy,$$

$$f_1(P, T) = \iint_U f\, dx\, dy.$$

Then the difference approximation (2) arising from the triangulation $\mathfrak{J}$ is defined by

$$L_1v(P) = \sum \{a_1(P, T)(v(P) - v(A_1)) + a_2(P, T)(v(P) - v(A_2)) + b(P, T)v(P)\}$$
$$= \sum f_1(P, T),$$

the sums being taken over $T \in \mathfrak{J}(P)$. Define functions $b(P)$, $a(P, Q)$ by

$$b(P) = \sum b(P, T), \quad T \in \mathfrak{J}(P)$$

$$a(P, Q) = \begin{cases} a_1(P, T) + a_2(P, T'), & Q \text{ a neighbor of } P \\ 0, & Q \text{ not a neighbor of } P, \end{cases}$$

where $T, T' \in \mathfrak{J}(P) \cap \mathfrak{J}(Q)$. Then (2) may be written

$$(7) \qquad L_1v(P) \equiv \sum a(P, Q)(v(P) - v(Q)) + b(P)v(P) = f_1(P), \qquad P \in \mathcal{S}_0 .$$

By requiring $v \in \mathcal{C}_0$ (7) is a system of N equations in N unknowns. L_1 is a symmetric, positive definite "Stieltjes" matrix. If $\mathfrak{J}(P)$ contains exactly 6 triangles for each P, L_1 is block 2-cyclic, and the system (7) may be solved numerically by one of the variety of methods discussed in [5].

One could introduce an area weight at each $P \in \mathcal{S}$ defined by $\sum_T | U |$, $T \in \mathfrak{J}(P)$, where $| U |$ is the area of the quadrilateral U, and use these weights to construct norms equivalent to $\| \alpha \|$, $\| \alpha \|_1$, but having more geometric meaning. The equivalence would be expressed with the constant c_3.

3. Some Remainder Terms. In this section we give two approximation formulae with the error bounded in a form suitable for our later use. Let

$$d = \max \{\rho(P, Q), P, Q \epsilon \bar{R}\}.$$

LEMMA 1. *There is a* $c_4 > 0$ *depending only on d such that, if U is the quadrilateral* PB_1CB_2, $u \in H^2(U)$, *and q is a nonnegative bounded function on U, then*

$$(8) \qquad \left| \iint_U qu\, dx\, dy - u(P) \iint_U q\, dx\, dy \right| \leqq c_4(\sup q)h(P)^2 \| u, U \|_2 .$$

Proof. It suffices to prove (8) for u having continuous second derivatives. Referring to Figure 1 we may take P to be the origin of coordinates and A_1 to lie on the positive x axis. Using polar coordinates let the line B_1CB_2 be given by $r = R(\theta)$, $0 \leqq \theta \leqq \alpha =$ the angle at P. For $(r, \theta) \in U$ one has

$$u(r, \theta) - u(P) = \int_{\rho=0}^{r} u_r(\rho, \theta)\, d\rho = ru_r(r, \theta) - \int_{\rho=0}^{r} \rho u_{rr}(\rho, \theta)\, d\rho.$$

Multiplying this by rq and integrating over U one finds that the left side of (8) is bounded by

$$\int_{\theta=0}^{\alpha} \int_{r=0}^{R(\theta)} rq \left\{ ru_r(r, \theta) - \int_{\rho=0}^{r} \rho u_{rr}(\rho, \theta)\, d\rho \right\} dr\, d\theta$$

$$\leqq (\sup q)h(P) \iint_U | u_r |\, dx\, dy + (\sup q)h(P)^2 \iint_U | u_{rr} |\, dx\, dy$$

$$\leqq (\sup q)h(P)[1 + h(P)]\, | U |^{1/2} \cdot \| u, U \|_2$$

which proves (8) since $| U | \leqq h(P)^2$ and $1 + h(P) \leqq 1 + d$.

LEMMA 2. *There is a* $c_5 > 0$ *depending only on* c_3 *such that, if V is the triangle* PCA_1, *L is the line segment* B_1C, η *is a unit vector pointing from P to* A_1, $u \in H^2(V)$, *and p is a nonnegative bounded function on L, then*

$$(9) \qquad \left| \int_L p(\eta \cdot \nabla)u\, ds - \frac{u(A_1) - u(P)}{\rho(P, A_1)} \int_L p\, ds \right| \leqq c_5(\sup p)h(P)\, \| u, V \|_2 .$$

Proof. It suffices to prove (9) for u having continuous second derivatives. Referring to Figure 1 we may take B_1 to be the origin of coordinates and A_1 to lie on the positive x axis. Let $\rho(P, B_1) = \rho(A_1, B_1) = a$, $\rho(C, B_1) = b$. If $\xi(y) = a(b - y)/b$ the line CA_1 contains the points $(\xi(y), y)$ and the line CP contains the points $(-\xi(y), y)$, $0 \leqq y \leqq b$. The inequality (9) follows from the two inequalities

$$(10) \qquad \left| \int_{y=0}^{b} p(0, y) \left[u_x(0, y) - \left[\frac{u(\xi(y), y) - u(-\xi(y), y)}{2\xi(y)} \right] dy \right| \right.$$
$$\leqq c_6(\sup p)h(P)\, \| u, V \|_2$$

$$(11) \qquad \left| \int_{y=0}^{b} p(0, y) \left[\frac{u(\xi(y), y) - u(-\xi(y), y)}{2\xi(y)} - \frac{u(a, 0) - u(-a, 0)}{2a} \right] dy \right|$$
$$\leqq c_7(\sup p)h(P)\, \| u, V \|_2$$

where c_6 and c_7 are positive constants depending only on c_3. To prove (10) one may use Taylor's theorem with integral remainder term to bound the left side of (10) by

$$\frac{1}{2}(\sup p)\int_{y=0}^{b}\int_{t=-\xi(y)}^{\xi(y)} |u_{xx}(t,y)|\,dt \leqq \frac{1}{2}(\sup p)\,|V|^{1/2}\,\|u,V\|_2$$

and note that $|V| \leqq h(P)^2$.

The inequality (11) will now be proved. Define $|D^2u| = [u_{xx}^2 + u_{xy}^2 + u_{yy}^2]^{1/2}$. Then

$$\pm\,(u_x(\theta\xi(y),y) - u_x(\theta b,0)) = \pm\int_{t=0}^{y}\frac{d}{dt}u_x(\theta\xi(t),t)\,dt$$

$$\leqq a^{-1}(a^2+b^2)^{1/2}\int_{t=0}^{y}|D^2u|\,(\theta\xi(t),t)\,dt.$$

If this is integrated with respect to θ over $(-1, 1)$ one obtains

$$\pm\left(\frac{1}{\xi(y)}\int_{-\xi(y)}^{\xi(y)} u_x(x,y)\,dx - \frac{1}{b}\int_{-b}^{b} u_x(x,0)\,dx\right)$$

(12)

$$\leqq a^{-1}(a^2+b^2)^{1/2}\int_{t=0}^{y}\int_{s=-\xi(t)}^{\xi(t)}\frac{1}{\xi(t)}\,|D^2u|\,(s,t)\,ds\,dt.$$

After multiplying both sides of (12) by $p(0, y)$, integrating with respect to y over $(0, b)$, and interchanging the order of the y and t integrations, one finds that the left side of (11) is bounded by

$$(\sup p)ba^{-2}(a^2+b^2)^{1/2}\iint_V |D^2u|\,dx\,dy \leqq (\sup p)c_5\,|V|\,\|u,V\|_2\,.$$

This proves (11) because $|y| \leqq h(P)^2$.

4. The Discretization Error. For our error bounds we assume there exists a $c_6 > 1$ such that in the closure of R,

$$(13)\qquad 1/c_6 \leqq p(x,y), \qquad q(x,y) \leqq c_6\,.$$

We also define a constant c_7 by the condition that no $P \in \mathcal{S}$ has more than c_7 neighbors.

LEMMA 3. *There is a c_8 depending on c_3, c_6, and c_7 such that*

$$(14)\qquad c_8^{-1}\|\alpha\|_1 \leqq \{\textstyle\sum \alpha(P)L_1\alpha(P)\}^{1/2} \leqq c_8\|\alpha\|_1$$

for any $\alpha \in \mathcal{C}_0$, the sum being taken over $P \in \mathcal{S}$.

Proof. One has

$$(15)\qquad \textstyle\sum \alpha(P)L_1\alpha(P) = \frac{1}{2}\sum_1 a(P,Q)(\alpha(P)-\alpha(Q))^2 + \sum b(P)\alpha(P)^2.$$

The proof follows easily from (15).

L_1 is symmetric and (15) shows that it is positive definite. Hence we define an inner product on $\mathcal{C}_0$ by $(\alpha, \beta)' = \sum \alpha(P)L_1\beta(P)$, and denote the corresponding norm by $\|\alpha\|'$.

THEOREM 1. *Let $u \in H_0^1 \cap H^2$ be a solution of (1), and let $v \in \mathcal{C}_0$ be the corresponding solution of (2). Then there is a constant c_9 depending only on c_3, c_6, c_7, and d,* *such that, if* $e \in \mathcal{C}_0$ *is defined by* $e(P) = u(P) - v(P)$, *then*

$$\| e \|_1 \leqq hc_9 \| u \|_2 .$$

Proof. Using (14), one has

$$\| e \|_1 \leqq c_8^2 \| e \|'.$$

Hence the theorem follows from the inequality

$$(16) \qquad | (e, e)' | \leqq hc_{10} \| e \|_1 \| u \|_2 ,$$

where c_{10} depends only on c_3, c_6, c_7, and d. One has $L_1 e = L_1 u - f_1$. Because $u \in H^2$, one has, referring to Figure 1,

$$f_1(P) = \sum \left\{ \int p \frac{du}{dn} ds + \iint qu \, dx \, dy \right\},$$

the sum being taken over all triangles $T \in \mathfrak{J}(P)$; the line integral is taken over the line segments B_1CB_2, and the area integral is taken over the quadrilateral PB_1CB_2. Analogous to (15), a calculation gives

$$(17) \qquad (e, e)' = \tfrac{1}{2}\sum\nolimits_1 [e(P) - e(Q)]E(P, Q) + \sum e(P)F(P),$$

where

$$E(P, A_1) = \frac{u(P) - u(A_1)}{\rho(P, A_1)} \int p \, ds - \int p \frac{du}{dn} ds,$$

the line integral being taken over B_1C and the corresponding perpendicular bisector on the other side of PA_1 (see Figure 1), and

$$F(P) = u(P) \iint q \, dx \, dy - \iint up \, dx \, dy,$$

the area integrals being taken over all the quadrilaterals PB_1CB_2 of triangles $T \in \mathfrak{J}(P)$. Using Lemmas 1 and 2 we obtain

$$\begin{aligned} | (e, e)' | &\leqq \tfrac{1}{2} c_5 c_6 \sum\nolimits_1 h(P) \| u, T \|_2 \, | e(P) - e(Q) | + c_4 c_6 \sum h(P)^2 \| u, T \|_2 \, | e(P) \\ &\leqq c_{11} h \{ \sum | e(P) - e(Q) |^2 \}^{1/2} \| u \|_2 + c_{11} h \{ \sum h(P)^2 e(P)^2 \}^{1/2} \| u \|_2 \\ &\leqq 2 c_{11} h \| e \|_1 \| u \|_2 , \end{aligned}$$

which proves the theorem.

It is easily seen that the proof remains valid if $u \in H^2(T)$ for each triangle T of $\mathfrak{J}$.

To extend this result to the case $q \geqq 0$ it seems necessary to make further restrictions on the triangulation. The first requirement is

(A) There is a $c_{12} > 1$ such that whenever $A, B, C, D \in \mathcal{S}$ and A and B are neighbors and C and D are neighbors, one has

$$(c_{12})^{-1} \leqq \rho(A, B)/\rho(C, D) \leqq c_{12} .$$

To state the second condition, let a line λ of $\mathfrak{J}$ be a sequence $\{P_1, P_2, \cdots, P_n\}$ of points of $\mathcal{S}$ such that P_i is a neighbor of P_{i+1}, $1 \leqq i < n$, define the ends of λ to be the points P_1, P_n, and define the length of λ to be $\sum \rho(P_i, P_{i+1})$, $1 \leqq i < n$.

The second condition is

(B) $\mathcal{S}$ may be written as a union of a set of lines λ such that no two lines have a point in common and each line has a least one endpoint on ∂R. Given such a decomposition of $\mathcal{S}$, let c_{13} denote the maximum length of the lines λ in the decomposition.

We also assume that there exists a $c_{14} > 1$ such that in the closure of R,

$$\begin{cases} p(x, y), q(x, y) \leqq c_{14} \\ q(x, y) \geqq 0 \\ p(x, y) \geqq 1/c_{14} \end{cases} \tag{18}$$

Then Lemma 3 is easily extended as follows.

LEMMA 4. *Suppose* $\mathfrak{I}$ *satisfies* (A) *and* (B) *and suppose* (18) *holds. Then there is a* c_{15} *depending on* c_7, c_{12}, c_{13}, *and* c_{14}, *such that*

$$(c_{15})^{-1} \| \alpha \|_1 \leqq \{\sum \alpha(P)L_1\alpha(P)\}^{1/2} \leqq c_{15} \| \alpha \|_1$$

for any $\alpha \in \mathcal{C}_0$, *the sum being taken over* $P \in \mathcal{S}$.

Proof. Let $\lambda = \{P_1, \cdots, P_n\}$ be one of the lines of (B). Then

$$| \alpha(P_j)| \leqq \sum | \alpha(P_{i+1}) - \alpha(P_i)| \leqq [(n - 1)\sum(\alpha(P_{i+1}) - \alpha(P_i))^2]^{1/2}, \qquad 1 \leqq i < n.$$

Hence

$$\sum(n - 1)^{-2}\alpha(P_i)^2 \leqq \sum(\alpha(P_{i+1}) - \alpha(P_i))^2, \qquad 1 \leqq i < n.$$

Now for any j, $1 \leqq j \leqq n$,

$$c_{13} \geqq \sum\rho(P_i, P_{i+1}) \geqq (n - 1)h(P_j)(c_{12})^{-1}.$$

Hence

$$\sum h(P_i)^2\alpha(P_i)^2 \leqq (c_{12}c_{13})^2\sum(\alpha(P_{i+1} - \alpha(P_i))^2, \qquad 1 \leqq i < n.$$

The left sum may be extended over $1 \leqq i \leqq n$. This is obvious if $\alpha(P_n) = 0$, and if $\alpha(P_1) = 0$ the same argument may be applied to the lines λ ordered in the other direction. Summing this over all lines λ of the decomposition and using (15),

$$(\alpha, \alpha) \leqq 4\, c_{12}^3 c_{13}^2 \sum \alpha(P)L_1\alpha(P).$$

The rest of the proof follows that of Lemma 3.

Using this lemma the following theorem may be proved in the same manner as Theorem 1.

THEOREM 2. *Assume* (A), (B), *and* (18). *Then there is a constant* c_{16} *depending only on* c_7, c_{12}, c_{13}, c_{14}, *and* d, *such that if* $u \in H_0^1 \cap H^2$ *is a solution of* (1) *and* $v \in \mathcal{C}_0$ *is the corresponding solution of* (2), *and* $e(P) = u(P) - v(P)$, *then*

$$\| e \|_1 \leqq hc_{16} \| u \|_2 .$$

Bettis Atomic Power Laboratory
Pittsburgh, Pennsylvania

1. J. L. Lions, *Equations Differentielles Operationelles*, Springer, Berlin, 1961.
2. R. H. MacNeal, "An asymetrical finite difference network," *Quart. Appl. Math.*, v. 12, 1953, p. 295–310.
3. L. Nirenberg, "Remarks on strongly elliptic partial differential equations," *Comm. Pure Appl. Math.*, v. 8, 1955, p. 649–674.
4. J. Nitsche & J. C. C. Nitsche, "Error estimates for the numerical solution of elliptic differential equations," *Arch. Rational Mech. Anal.*, v. 5, 1960, p. 293–306.
5. R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, New York, 1962.