Numerische Mathematik 4, 313—327 (1962)

On the Formulation of finite Difference analogues of the Dirichlet Problem for Poisson's Equation*

By

J. H. BRAMBLE and B. E. HUBBARD

1. Introduction

In the approximate solution of the Dirichlet problem for Poisson's equation many finite difference methods have been proposed. It is important for the formulation of the approximating problem to have some measure of the deviation of the finite difference solution from the exact solution. The question of convergence itself for a class of problems has been discussed by COURANT, FRIEDRICHS and LEWY [3].

In 1930 GERSCHGORIN [5] gave a method for obtaining an estimate of the order of convergence of the finite difference approximation to the solution of the Dirichlet problem for a class of elliptic equations. His method was based on a maximum principle for the tinite difference analogue. In a note in 1933 COLLATZ [1] proposed a certain boundary approximation and, using the techniques of GERSCHGORIN, showed that this approximation gives rise to an $O(h^2)$ estimate for the truncation error. The estimates of both GERSCHGORIN and COLLATZ assume the knowledge of bounds for certain higher derivatives of the solution of the Dirichlet problem.

From an analogy to probability theory COURANT, FRIEDRICHS, and LEWY [3] give a finite difference Green's function for the Dirichlet problem for Poisson's equation. Using this Green's function they give an analogue of Green's third identity. WASOW [14] studies the asymptotic behavior of the finite difference Green's function and LAASONEN [8] uses an explicit representation of the finite difference Green's function for the rectangle to obtain bounds in that case.

In this paper we obtain some further estimates of the type proposed by GERSCHGORIN. The approach taken here is to define an appropriate related finite difference Green's function for various finite difference analogues. In each case the analogue of Green's third identity is given and used to obtain estimates for the truncation error.

In section 2 the truncation error is studied for a finite difference approximation proposed by SHORTLEY and WELLER [10]. Although at points near the boundary the finite difference operator approximates the Laplace operator only to $O(h)$ it is seen that the resulting contribution to the truncation error is $O(h^3)$.

* This research was supported in part by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command under Contract No. AF 49(638)-228.

This has been shown to be the case for the homogeneous equation by FORSYTHE and WASOW [4].

In section 3 techniques similar to those of section 2 are used to obtain certain known results. For example the method of COLLATZ is treated from this point of view. From these considerations we are led to formulate finite difference approximations in which the matrix representing the linear system is not of "positive type", (MOTZKIN and WASOW [9]). In this connection we state a general theorem on error estimation for a class of finite difference analogues of the Dirichlet problem for Poisson's equation. This theorem is similar to a special case of a theorem of FORSYTHE and WASOW [4, p. 302] concerning approximations whose associated matrices are of "positive type".

In the final section we apply this theorem to a certain finite difference problem and show that the resulting truncation error is $O(h^4)$. Finally, combining the techniques of the previous two sections we formulate yet another $O(h^4)$ finite difference approximation.

For some other studies of this problem we refer the reader to [1], [4], [7], [11], [12], [13] and references therein.

2. Second Order Estimates

Throughout this paper we shall be concerned with finite difference approximations to the Dirichlet problem for Poisson's equation, i.e.

$$\begin{aligned} \Delta u(x, y) &= F(x, y), \qquad (x, y) \in R \\ u(x, y) &= f(x, y), \qquad (x, y) \in C. \end{aligned} \tag{2.1}$$

We assume that R is a bounded region in the (x, y) plane with boundary C.

We cover the (x, y) plane with a grid made up of two families of lines. Each family consists of lines, a distance h apart, parallel to one of the coordinate axes. The intersections will be called either "grid" or "mesh" points and a function defined at such points will be termed a "mesh" function. If we restrict ourselves to a finite portion of the plane such a function can be considered as a vector in a finite dimensional vector space. Such is the case with the point sets which we shall now define.

Let R_h be the set of those mesh points in R whose nearest neighbors in the x and y directions lie in R. Those grid points in R which do not belong to R_h will make up the set called C_h^*. The points of intersection of the grid with the boundary C form the set C_h.

For any point P belonging to $R_h + C_h^* + C_h$ we define its neighbors $N(P)$ to be those nearest points in $R_h + C_h^* + C_h$, lying along grid lines through P.

If $V(x, y)$ is an arbitrary mesh function defined on $R_h + C_h^* + C_h$ then for such vectors we define the finite difference operator Δ_h. If $(x, y) \in R_h$ then

$$\begin{aligned} \Delta_h V(x, y) \equiv h^{-2}\{&V(x+h, y) + V(x, y+h) + \\ &+ V(x-h, y) + V(x, y-h) - 4V(x,y\}. \end{aligned} \tag{2.2}$$

This is the usual $O(h^2)$ approximation of Δ for functions $v(x, y) \in C^4$ in R. In fact

$$|\Delta v(x,y) - \Delta_h v(x,y)| \leqq \frac{h^2}{6} M_4, \qquad (x, y) \in R_h, \tag{2.3}$$

where we have used the notation

$$M_j = \sup_{P\in R}\left\{\left|\frac{\partial^j U(P)}{\partial x^i\,\partial y^{j-i}}\right| \,\middle|\, i=0,1,\ldots,j\right\}. \tag{2.4}$$

At points of C_h^*, Δ_h is defined to be the five point divided difference approximation to Δ. For example if $(\bar{x},\bar{y})\in C_h^*$ is the point in Fig. 1 then

$$\begin{aligned}\Delta_h V(\bar{x},\bar{y}) \equiv 2h^{-2}\Big\{&\left(\frac{1}{\alpha+1}\right)V(\bar{x}+h,\bar{y}) + \frac{1}{\alpha(\alpha+1)}V(\bar{x}-\alpha h,\bar{y}) + \\ &+\left(\frac{1}{\beta+1}\right)V(\bar{x},\bar{y}+h) + \frac{1}{\beta(\beta+1)}V(\bar{x},\bar{y}-\beta h) - \left(\frac{1}{\alpha}+\frac{1}{\beta}\right)V(\bar{x},\bar{y})\Big\}.\end{aligned} \tag{2.5}$$

If $\alpha=\beta=1$ then Δ_h takes the same form as in (2.2). In fact, either α or β or both may equal 1. We note that Δ_h as defined in (2.5) approximates Δ to $O(h)$ for $v(x,y)\in C^3$ in R, i.e.

$$|\Delta v(\bar{x},\bar{y}) - \Delta_h v(\bar{x},\bar{y})| \leqq \frac{2M_3 h}{3}. \tag{2.6}$$

As in the formulation of Shortley and Weller [10] we allow any or all the neighbors of a point $P\in C_h^*$ to lie at a distance less than or equal to h from p. The appropriate analogy of (2.5) is assumed.

We consider the following finite difference analogue of (2.1),

$$\begin{aligned}\Delta_h U(x,y) &= F(x,y), \qquad (x,y)\in R_h + C_h^*,\\ U(x,y) &= f(x,y), \qquad (x,y)\in C_h.\end{aligned} \tag{2.7}$$

This is just a system of simultaneous linear equations for the determination of the mesh function $U(x,y)$. It is well-known that the associated determinant of such a system does not vanish and hence there exists a unique solution for (2.7); cf. Collatz [2, p. 46].

We shall show that the truncation error $\varepsilon(P)\equiv u(P)-U(P)$, $P\in R_h + C_h^* + C_h$. satisfies an inequality of the type

$$|\varepsilon|_M \leqq K h^2 \tag{2.8}$$

where K is a constant independent of P and h. In (2.8) we have used the notation

$$\psi_M = \sup_{P\in S\subset\bar{R}} \psi(P) \tag{2.9}$$

for any function ψ defined on a subset S of $\bar{R}$. Before proceeding we shall introduce the finite difference analogue of the Green's function, $G_h(P,Q)$ which is defined by

$$\begin{aligned}\Delta_{h,P} G_h(P,Q) &= -\delta(P,Q)h^{-2}, \qquad P\in R_h + C_h^*\\ G_h(P,Q) &= \delta(P,Q), \qquad P\in C_h\end{aligned} \tag{2.10}$$

for $Q\in R_h + C_h^* + C_h$.

Here

$$\delta(P,Q) = \begin{cases}1, & P=Q\\ 0, & P\neq Q.\end{cases} \tag{2.11}$$

We shall now prove the finite difference analogues of some well known theorems in potential theory.

Lemma 2.1 (Maximum Principle). For any mesh function $V(P)$ defined on $R_h + C_h^* + C_h$ if $\Delta_h V(P) \geqq 0$ for $P \in R_h + C_h^*$ then $V(P)$ takes on its maximum on C_h.

Such a result was first obtained by GERSCHGORIN [5] and is easily seen to be a special case of more general theorems given by COLLATZ [2].

Lemma 2.2 (Green's Third Identity). Let $V(P)$ be an arbitrary mesh function defined on $R_h + C_h^* + C_h$. Then for any $P \in R_h + C_h^* + C_h$

$$V(P) = h^2 \sum_{Q \in R_h + C_h^*} G_h(P, Q)\,[-\Delta_h V(Q)] + \sum_{Q \in C_h} G_h(P, Q)\, V(Q). \tag{2.12}$$

Proof. The relation (2.12) can be proved from the finite difference analogue of Green's second identity. It can be seen more simply from the following considerations. Let $W(P)$ be the right hand side of (2.12). By a direct calculation using the properties of the Green's function $G_h(P, Q)$ it follows that

$$\Delta_h W(P) = \Delta_h V(P), \qquad P \in R_h + C_h^* \tag{2.13}$$

$$W(P) = V(P), \qquad P \in C_h. \tag{2.14}$$

From the uniqueness of the solution of (2.7) we have $W(P) = V(P)$.

Lemma 2.3.

$$G_h(P, Q) \geqq 0, \qquad Q \in R_h + C_h^* + C_h. \tag{2.15}$$

Proof. Apply the maximum principle (lemma 2.1) to $-G_h(P, Q)$ for arbitrary but fixed $Q \in R_h + C_h^* + C_h$.

Lemma 2.4.

$$\sum_{Q \in C_h^*} G_h(P, Q) \leqq 1, \qquad P \in R_h + C_h^* + C_h. \tag{2.16}$$

Proof. Let the mesh function $W(P)$ be given by

$$W(Q) = \begin{cases} 1, & Q \in R_h + C_h^*, \\ 0, & Q \in C_h. \end{cases} \tag{2.17}$$

Then $\Delta_h W(Q) = 0$, $Q \in R_h$. It is easily seen from the definition of Δ_h on C_h^* that $-\Delta_h W(Q) \geqq h^{-2}$

Applying lemma 2.2 it follows that for $P \in R_h + C_h^*$

$$1 = h^2 \sum_{Q \in C_h^*} G_h(P, Q)\,[-\Delta_h W(Q)] \geqq \sum_{Q \in C_h^*} G_h(P, Q).$$

If $P \in C_h$ the inequality (2.16) is trivially satisfied.

Lemma 2.5. If d is the diameter of the smallest circumscribed circle containing R then

$$h^2 \sum_{Q \in R_h + C_h^*} G_h(P, Q) \leqq \frac{d^2}{16}, \qquad P \in R_h + C_h^* + C_h. \tag{2.18}$$

Proof. Let 0 be the center of the circumscribed circle about R of diameter d. Let $W(P)=\frac{r(P)^2}{4}$ for $P\in R_h+C_h^*+C_h$ where $r(P)$ is the Euclidean distance from 0 to P. Then

$$\Delta_h W(P)=1, \qquad P\in R_h+C_h^*.$$

Now define the mesh function

$$V(P)\equiv h^2 \sum_{Q\in R_h+C_h^*} G_h(P,Q).$$

We see from (2.10) that

$$\text{(2.19)} \qquad \begin{aligned} \Delta_h V(P)&=-1, \qquad & P\in R_h+C_h^* \\ V(P)&=0, & P\in C_h. \end{aligned}$$

Hence $\Delta_h[V(P)+W(P)]=0$ for $P\in R_h+C_h^*$ and $V(P)+W(P)\leqq\frac{d^2}{16}$ for $P\in C_h$. By the maximum principle, since $W\geqq 0$, it follows that

$$V(P)\leqq\frac{d^2}{16}, \qquad P\in R_h+C_h^*+C_h$$

which completes the proof of lemma 2.5.

We are now in a position to prove the principal result of section 2.

Theorem 1. Let $u(x, y)$ be the solution of (2.1) and $U(x, y)$ the solution of (2.7). Then the truncation error $\varepsilon(P)\equiv u(P)-U(P)$ satisfies the inequality

$$\text{(2.20)} \qquad |\varepsilon|_M\leqq\frac{M_4 d^2}{96}h^2+\frac{2M_3}{3}h^3.$$

Proof. Since $\varepsilon(P)=0$, $P\in C_h$ we see from lemma 2.2 that

$$\text{(2.21)} \qquad \varepsilon(P)=h^2\sum_{Q\in R_h+C_h^*} G_h(P,Q)[-\Delta_h\varepsilon(Q)].$$

It follows from (2.1) and (2.7) that

$$\text{(2.22)} \qquad |-\Delta_h\varepsilon(Q)|=|\Delta_h u(Q)-\Delta u(Q)|.$$

Applying (2.22) to (2.21) and using (2.3) and (2.6) we have

$$|\varepsilon(P)|\leqq\left[h^2\sum_{Q\in R_h} G_h(P,Q)\right]\frac{h^2 M_4}{6}+\left[\sum_{Q\in C_h^*} G_h(P,Q)\right]\frac{2M_3}{3}h^3.$$

Finally using lemmas 2.4 and 2.5 we arrive at (2.20), which completes the proof of theorem 1.

3. Other Boundary Approximations

In this section we derive some known inequalities using the results and techniques of the previous section. In addition we show that in certain cases the requirement of positivity can be removed near the boundary C. In this connection we give an example of a formulation of a finite difference analogue of (2.1) which fails to be of positive type at points of C_h^*. For this problem we derive an over-all $O(h^2)$ estimate for the truncation error and show that the contribution from points of C_h^* is $O(h^3)$ as in Theorem 1.

Finally we give a general theorem on error estimation for finite difference approximations to Poisson's equation. This result, in certain cases, is similar to a general theorem given in G. FORSYTHE and W. WASOW [p. 302, theorem 23.7, 4]. In the theorem stated here the condition of non-negativity is relaxed to admit more general approximations near the boundary.

Let $G_h^*(P, Q)$ be the finite difference Green's function for R_h with boundary C_h^*. This is given by

$$\begin{aligned} \Delta_{h,P}\, G_h^*(P, Q) &= -\delta(P, Q)\, h^{-2}, \qquad P \in R_h \\ G_h^*(P, Q) &= \delta(P, Q), \qquad P \in C_h^* \end{aligned} \tag{3.1}$$

for all $Q \in R_h + C_h^*$.

Just as in lemma 2.2 we have the identity

$$V(P) = h^2 \sum_{Q \in R_h} G_h^*(P, Q)\left[-\Delta_h V(Q)\right] + \sum_{Q \in C_h^*} G_h^*(P, Q)\, V(Q). \tag{3.2}$$

In addition all of the other lemmas of section 2 are valid if we make the substitutions

$$\begin{aligned} G_h &\to G_h^*, \\ R_h + C_h^* &\to R_h \\ C_h &\to C_h^*. \end{aligned} \tag{3.3}$$

We shall also need the following lemma.

Lemma 3.1. For $P \in R_h + C_h^*$

$$\sum_{Q \in C_h^*} G_h^*(P, Q) = 1. \tag{3.4}$$

Proof. Apply (3.2) to $V(P) \equiv 1$.

We first define a finite difference analogue of (2.1) and show that the truncation error is $O(h)$. FORSYTHE and WASOW [4] call this "interpolation of order zero". Let $V(P)$ satisfy

$$\begin{aligned} \Delta_h U(P) &= F(P), \qquad P \in R_h \\ U(P) &= f(P'), \qquad P \in C_h^* \end{aligned} \tag{3.5}$$

where P' is one of the neighbors of P in C_h. We have the following theorem.

Theorem 2. Let $u(x, y)$ be the solution of (2.1) and $U(x, y)$ the solution of (3.5). Then the truncation error $\varepsilon(P) = U(P) - U(P)$ satisfies the inequality

$$|\varepsilon|_M \leqq h M_1 + \frac{M_4 d^2}{96} h^2. \tag{3.6}$$

Proof. Applying (3.2) to $\varepsilon(P)$ we have

$$\varepsilon(P) = h^2 \sum_{Q \in R_h} G_h^*(P, Q)\left[-\Delta_h \varepsilon(Q)\right] + \sum_{Q \in C_h^*} G_h^*(P, Q)\, \varepsilon(Q). \tag{3.7}$$

We note that for $Q \in C_h^*$

$$|\varepsilon(Q)| = |u(Q) - U(Q)| = |u(Q) - U(Q')| \leqq h M_1. \tag{3.8}$$

Now from (2.3) and (3.5) we have also that

$$|\Delta_h \varepsilon(Q)| \leqq \frac{h^2}{6} M_4, \qquad Q \in R_h. \tag{3.9}$$

Taking absolute values of both sides of (3.7), using (3.8) and (3.9), and applying lemmas 2.5 and 3.1, the inequality (3.6) follows easily.

We next consider the finite difference analogue of (2.1) given by COLLATZ [1], [2]. He defines the following approximation to (2.1).

$$\begin{aligned} \Delta_h U(P) &= F(P), \qquad P \in R_h \\ U(P) &= f(P), \qquad P \in C_h. \end{aligned} \tag{3.10}$$

At a point P of C_h^* he prescribes that $U(P)$ lie on a straight line between the values of U at two neighbors of P, one of which is in R_h, the other in C_h. For example for the point $(\bar{x}, \bar{y})$ of Fig. 1 we have

$$U(\bar{x}, \bar{y}) = \frac{\alpha}{\alpha+1} U(\bar{x} + h, \bar{y}) + \frac{1}{\alpha+1} U(\bar{x} - \alpha h, \bar{y}). \tag{3.11}$$

Alternatively we could have interpolated in the y direction.

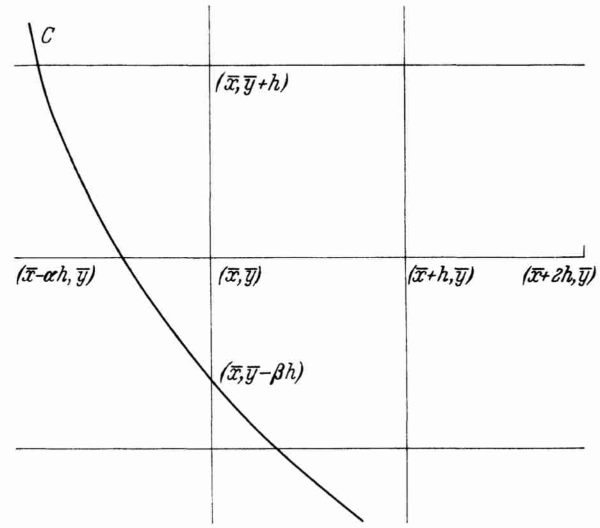

Fig. 1

As COLLATZ has shown [1] this method gives rise to an estimate of the truncation error which is $O(h^2)$. The contribution to the truncation error arising from the points of C_h^* is also $O(h^2)$. The following analysis again yields similar results.

Theorem 3 (COLLATZ). Let $u(x, y)$ be the solution of (2.1) and $U(x, y)$ the solution of (3.10) and (3.11). Then the truncation error $\varepsilon(P) = u(P) - U(P)$ satisfies

$$|\varepsilon|_M \leqq \left[M_2 + \frac{M_4 d^2}{48}\right] h^2. \tag{3.12}$$

Proof. As in the proof of theorem 2, $\varepsilon(P)$ satisfies (3.7). For any point $Q \in C_h^*$ we make use of (3.11) to bound $|\varepsilon(Q)|$ as follows.

$$\begin{aligned} |\varepsilon(\bar{x}, \bar{y})| &= |u(\bar{x}, \bar{y}) - U(\bar{x}, \bar{y})| \\ &= \left| u(\bar{x}, \bar{y}) - \frac{\alpha}{\alpha+1} U(\bar{x} + h, \bar{y}) - \frac{1}{\alpha+1} u(\bar{x} - \alpha h, \bar{y}) \right| \end{aligned} \tag{3.13}$$

where we have used the second relation in (3.10). Using the triangle inequality we have

$$|\varepsilon(\bar{x},\bar{y})| \leqq \left| u(\bar{x},\bar{y}) - \frac{\alpha}{\alpha+1} u(\bar{x}+h,\bar{y}) - \frac{1}{\alpha+1} u(\bar{x}-\alpha h,\bar{y}) \right| + \frac{\alpha}{\alpha+1} |\varepsilon|_M. \tag{3.14}$$

From Taylor's theorem and the fact that $0<\alpha\leqq 1$ it follows that

$$|\varepsilon(Q)| \leqq \frac{M_2}{2} h^2 + \frac{1}{2} |\varepsilon|_M. \tag{3.15}$$

Employing inequalities (3.15) and (3.9), as well as lemma 2.5 and (3.1) we see that

$$|\varepsilon(Q)| \leqq \frac{1}{2} |\varepsilon|_M + \left(\frac{M_2}{2} + \frac{M_4 d^2}{96}\right) h^2. \tag{3.16}$$

Since the right hand side is independent of Q the inequality (3.12) follows.

We next give an example of a finite difference analogue of (2.1) which fails to be of positive type at points of C_h^*. Let $U(P)$ satisfy the system

$$\begin{aligned} \Delta_h U(P) &= F(P), \qquad P \in R_h \\ U(P) &= f(P), \qquad P \in C_h. \end{aligned} \tag{3.17}$$

At a point P of C_h^* let $U(P)$ lie on a parabola through value of $U(P)$ at a neighboring point of C_h and two points of $R_h + C_h^*$. All four points involved must of course be colinear. In addition we require one of the points of $R_h + C_h^*$ to be a neighbor of P and the other to be taken at a distance $3h$ from P. For example, for the point $(\bar{x},\bar{y})$ in Fig. 1

$$\begin{aligned} U(\bar{x},\bar{y}) = \frac{3}{3+\alpha(\alpha+4)} \Big\{ & U(\bar{x}-\alpha h,\bar{y}) + \frac{\alpha}{2}(\alpha+3)\, U(\bar{x}+h,\bar{y}) - \\ & - \frac{\alpha}{6}(\alpha+1)\, U(\bar{x}+3h,\bar{y}) \Big\}. \end{aligned} \tag{3.18}$$

From Taylor's theorem it is easy to see that for a sufficiently smooth function $U(P)$ in R we have an inequality of the type

$$\begin{aligned} \Big| u(\bar{x},\bar{y}) - \frac{3}{3+\alpha(\alpha+4)} \Big\{ & u(\bar{x}-\alpha h,\bar{y}) + \frac{\alpha}{2}(\alpha+3)\, u(\bar{x}+h,\bar{y}) - \\ & - \frac{\alpha}{6}(\alpha+1)\, u(\bar{x}+3h,\bar{y}) \Big\} \Big| \leqq \frac{14 h^3 M_3}{3}, \end{aligned} \tag{3.19}$$

where $(\bar{x},\bar{y}) \in C_h^*$. In some cases the interpolation will be in the y direction. We thus have the following theorem.

Theorem 4. Let $u(x, y)$ be the solution of (2.1) and $U(x, y)$ the solution of (3.17) and (3.18). Then the truncation error $\varepsilon(P) = u(P) - U(P)$ satisfies

$$|\varepsilon|_M \leqq \frac{d^2 M_4}{12} h^2 + \frac{112}{3} M_3 h^3. \tag{3.20}$$

Proof. The proof follows in a manner analogous to that of theorem 3. Instead of (3.14) we have the inequality

$$\begin{aligned} |\varepsilon(\bar{x},\bar{y})| \leqq \Big| u(\bar{x},\bar{y}) - \frac{3}{3+\alpha(\alpha+4)} \Big\{ & u(\bar{x}-\alpha h,\bar{y}) + \frac{\alpha}{2}(\alpha+3)\, u(\bar{x}+h,\bar{y}) - \\ & - \frac{\alpha}{6}(\alpha+1)\, u(\bar{x}+3h,\bar{y}) \Big\} \Big| + \frac{7}{8} |\varepsilon|_M, \end{aligned} \tag{3.21}$$

for the point $(\bar{x}, \bar{y})$ of Fig. 1. From (3.21) and (3.19) it follows that

$$|\varepsilon(Q)| \leqq \frac{14 M_3}{3} h^3 + \frac{7}{8} |\varepsilon|_M, \tag{3.22}$$

where $Q \in C_h^*$. As before we have the analogue of (3.16)

$$|\varepsilon(Q)| \leqq \frac{d^2 M_4}{96} h^2 + \frac{14 M_3}{3} h^3 + \frac{7}{8} |\varepsilon|_M, \tag{3.23}$$

from which (3.20) easily follows.

We see from the previous theorem that the property of positivity of the matrix representing the linear system is not essential in the problem of obtaining error estimates. It is sufficient to replace this condition with a requirement of "interior positivity", provided a "strict" diagonal dominance is satisfied near the boundary. To be more precise we shall need some further definitions.

Let Δ_h now be defined as some finite difference analogue of Δ at mesh points in R. We define R_h to be the set of those mesh points in R at which the operator Δ_h is defined solely in terms of mesh points in R. Let C_h^* be the set of those mesh points in R which do not belong to R_h. At these points we define a finite difference operator Δ_h^* which involves points of $R_h + C_h^*$ and certain points on C. Those points of C which are involved in Δ_h^* at some point of C_h^* will be called C_h. We define $N(P)$ as those points, other than P involved in Δ_h at P, when $P \in R_h$. If $P \in C_h^*$ the definition is the same with Δ_h replaced by Δ_h^*. We define $\sigma(P, Q)$ to be the coefficient of $V(Q)$ in the expression for the operator applied to V at the point P. In terms of the operators Δ_h and Δ_h^* we define the following finite difference analogue of (2.1).

$$\begin{aligned} \Delta_h U(P) &= F(P) + \varphi[F(P)], \qquad && P \in R_h \\ \Delta_h^* U(P) &= H(P), && P \in C_h^* \\ U(P) &= f(P), && P \in C_h \end{aligned} \tag{3.24}$$

where $\varphi[F(P)]$ is an operator which is a linear combination of derivatives of F at P. $H(P)$ is, as yet, arbitrary. We give the following definitions.

Definition 3.1. The matrix representing the linear system (3.24) has the property of *interior positivity* if it is of positive type at each point of R_h, [9]. That is

$$\frac{\sigma(P, Q)}{\sigma(P, P)} < 0, \qquad P \in R_h, \qquad Q \in N(P). \tag{3.25}$$

Definition 3.2. The matrix representing the linear system (3.24) has the property of *strict diagonal dominance* if

$$\sum_{Q \in (R_h + C_h^*) \cap N(P)} |\sigma(P, Q)| \leqq |\sigma(P, P)|, \qquad P \in R_h \tag{3.26}$$

and

$$\sum_{Q \in (R_h + C_h^*) \cap N(P)} |\sigma(P, Q)| \leqq |\sigma(P, P)| \delta, \qquad P \in C_h^* \tag{3.27}$$

where δ is a constant less than 1 and independent of h.

Definition 3.3. Let $\overline{N(P)}=[N(P)\cup P]\cap R_h$. We say that R_h is "connected" if for every set S_h properly contained in R_h the decomposition

$$R_h=\Big[\bigcup_{P\in S_h}\overline{N(P)}\Big]\cup\Big[\bigcup_{Q\in R_h-S_h}\overline{N(Q)}\Big]$$

implies that

$$\Big[\bigcup_{P\in S_h}\overline{N(P)}\Big]\cap\Big[\bigcup_{Q\in R_h-S_h}\overline{N(P)}\Big]$$

is not empty.

We now state a theorem pertaining to systems of the type given by (3.24).

Theorem 5. Let $u(x, y)$ be the solution of (2.1) and $U(x, y)$ be the solution of (3.24). The matrix representing (3.24) is assumed to satisfy the properties

(a) Strict diagonal dominance with constant δ.

(b) Interior positivity.

(c) R_h is connected.

The operators Δ_h and Δ_h^* satisfy the inequalities

$$|\Delta_h v-[\Delta v+\varphi(\Delta v)]|\leqq C_1 M_{n+2}h^n,\qquad n\geqq 2 \tag{3.28}$$

and

$$\left|\frac{\Delta_h^* U(P)-H(P)}{\sigma(P,P)}\right|\leqq C_2 M_m h^m,\quad m\geqq 1,\quad P\in C_h^* \tag{3.29}$$

where C_1 and C_2 are constants, and H is a function given on C_h^*. In (3.28), v is any sufficiently smooth function in $\overline{R}$ and M_i is related to v. Then $\varepsilon(P)=u(P)-U(P)$ satisfies the inequality

$$|\varepsilon|_M\leqq\frac{1}{1-\delta}\left\{C_2 M_m h^m+C_1\frac{d^2}{16}M_{n+2}h^n\right\}. \tag{3.30}$$

Proof. We define the Green's function $G_h^*(P, Q)$ related to the operator Δ_h as defined on R_h with boundary C_h^*. That is

$$\Delta_{h,P}\, G_n^*(P, Q)=-h^{-2}\delta(P, Q),\qquad P\in R_h, \tag{3.31 a}$$

$$G_h^*(P, Q)=\delta(P, Q),\qquad P\in C_h^*. \tag{3.31 b}$$

for all $Q\in R_h+C_h^*$.

The existence of $G_h^*(P, Q)$ is assured by (a) and (c), cf. [2, p. 46]. Because of (a), (b) and (c) any function $V(P)$, $P\in R_h+C_h^*$ for which $\Delta_h V(P)\geqq 0$, $P\in R_h$ attains its maximum on C_h^*, cf. COLLATZ [p. 45, 2]. From this it follows easily that

$$G_h^*(P, Q)\geqq 0,\qquad Q\in R_h+C_h^*. \tag{3.32}$$

As before we have the identity

$$\varepsilon(P)=h^2\sum_{Q\in R_h}G_h^*(P, Q)\,[-\Delta_h\varepsilon(Q)]+\sum_{Q\in C_h^*}G_h^*(P, Q)\,\varepsilon(Q)\,. \tag{3.33}$$

This may be verified as in lemma 2.2. We observe that lemma 3.1 is valid for $G_h^*(P, Q)$ as defined in (3.31). As a consequence of (3.28) we see that

$$\Delta_h\left(\frac{r(P)^2}{4}\right)\equiv\Delta\left(\frac{r(P)^2}{4}\right)=1\,, \tag{3.34}$$

since we may take $M_{n+2}\equiv 0$. As in lemma 2.5 we thus have

$$h^2 \sum_{Q\in R_h} G_h^*(P, Q) \leqq \frac{d^2}{16}. \tag{3.35}$$

From the definition $\sigma(P, Q)$ and Δ_h^* we have

$$\varepsilon(P) = \frac{\Delta_h^* \varepsilon(P)}{\sigma(P, P)} - \frac{\sum\limits_{Q\in N(P)} \sigma(P, Q)\, \varepsilon(Q)}{\sigma(P, P}, \qquad P\in C_h^*. \tag{3.36}$$

Taking the absolute value of both sides of (3.36) it follows that

$$|\varepsilon(P)| \leqq \left| \frac{\Delta_h^*[u(P)-U(P)]}{\sigma(P, P)} \right| + \frac{\sum\limits_{Q\in(R_h+C_h^*)\cap N(P)} |\sigma(P, Q)|}{\sigma(P, P)} |\varepsilon|_M. \tag{3.37}$$

From (3.24), (3.27) and (3.29) we obtain

$$|\varepsilon(P)| \leqq C_2 M_m h^m + \delta |\varepsilon|_M. \tag{3.38}$$

If we now take the absolute value of both sides of (3.33) and apply (3.28), (3.35), (3.37) and lemma 3.1 the inequality (3.30) follows.

As a non-trivial application of theorem 5 an $O(h^4)$ analogue of (2.1) will be presented in the next section.

4. Fourth Order Estimates

Let Δ_h be the usual nine point approximation to Δ, defined by

$$\begin{aligned} \Delta_h v(x, y) = \frac{1}{6h^2} \{&4v(x+h, y) + 4v(x, y+h) + 4v(x-h, y) + \\ &+ 4v(x, y-h) + v(x+h, y+h) + v(x-h, y+h) + \\ &+ v(x-h, y-h) + v(x+h, y-h) - 20v(x, y)\}, \end{aligned} \tag{4.1}$$

[cf. KANTOROVICH and KRYLOV, p. 210, 6]. This definition of Δ_h fixes the sets R_h and C_h^* of theorem 5.

At points of C_h^*, the pure second derivatives of Δ are approximated to within terms of order h^2. Since Δ is invariant under rotations, these derivatives need not be restricted solely to u_{xx} and u_{yy}. That is, if (ξ, η) are new variables resulting from a rotation then

$$\Delta v \equiv v_{\xi\xi} + v_{\eta\eta}. \tag{4.2}$$

Consider the set of lines from $P\in C_h^*$ to the eight nearest grid points. From the definition of C_h^*, C must cut at least one of these lines. In the case where a horizontal line is cut by C we can approximate $v_{xx}(P)$ to $O(h^2)$ by an unbalanced four point formula which involves the value of v at the intersection of the line with C. For example, if P is the point (x, y) of Fig. 2, we have that

$$\begin{aligned} \Bigg| v_{xx}(x, y) - h^{-2} \Bigg\{ &\left(\frac{\alpha-1}{\alpha+2}\right) v(x+2h, y) + \frac{2(2-\alpha)}{\alpha+1} v(x+h, y) + \\ &+ \frac{6}{\alpha(\alpha+1)(\alpha+2)} v(x-\alpha h, y) - \left[\frac{6+\alpha(7-\alpha^2)}{\alpha(\alpha+1)(\alpha+2)}\right] v(x, y) \Bigg\} \Bigg| \\ \leqq \frac{1}{12} M_4 h^2 + \frac{1}{6} M_5 h^3 & \end{aligned} \tag{4.3}$$

and

$$(4.4) \qquad |v_{yy}(x, y) - h^{-2}\{v(x, y+h) + v(x, y-h) - 2v(x, y)\}| \leqq \frac{1}{12} M_4 h^2.$$

The analogous approximations are made if a vertical line is cut (or both are cut) by C. In this case we take Δ_h^* to be

$$(4.5) \qquad \begin{aligned} \Delta_h^* v(x, y) \equiv h^{-2} \Big\{ & v(x, y+h) + v(x, y-h) + \frac{\alpha-1}{\alpha+2} v(x+2h, y) + \\ & + \frac{2(2-\alpha)}{\alpha+1} v(x+h, y) + \frac{6}{\alpha(\alpha+1)(\alpha+2)} v(x-\alpha h, y) - \\ & - \left[2 + \frac{6+\alpha(7-\alpha^2)}{\alpha(\alpha+1)(\alpha+2)}\right] v(x, y) \Big\}. \end{aligned}$$

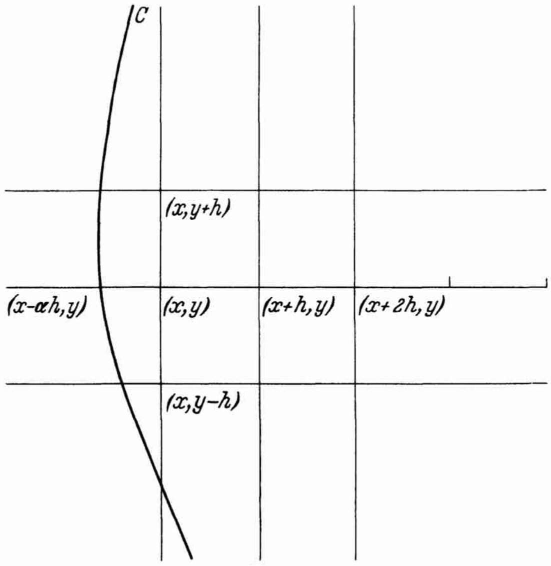

Fig. 2

In certain cases the boundary C may cut only a diagonal from P (see Fig. 3). In such a situation we consider the Laplace operator in the form (4.2) where

$$(4.6) \qquad \xi = \frac{x+y}{\sqrt{2}}, \qquad \eta = \frac{y-x}{\sqrt{2}},$$

which is just a rotation through 45 degrees. At least one of the terms $v_{\xi\xi}$, $v_{\eta\eta}$ will be approximated by a four point formula involving a point of C. For example if P is the point (x, y) of Fig. 3 we have

$$(4.7) \qquad \begin{aligned} \Big| v_{\xi\xi}(x, y) - \frac{1}{2h^2} \Big\{ & \left(\frac{\alpha-1}{\alpha+2}\right) v(x+2h, y+2h) + \\ & + \frac{2(2-\alpha)}{\alpha+1} v(x+h, y+h) + \frac{6}{\alpha(\alpha+1)(\alpha+2)} v(x-\alpha h, y-\alpha h) - \\ & - \left[\frac{6+\alpha(1-\alpha^2)}{\alpha(\alpha+1)(\alpha+2)}\right] v(x, y) \Big\} \Big| \leqq \frac{M_4}{6} h^2 + \frac{\sqrt{2} M_5}{3} h^3, \end{aligned}$$

and

$$\left| v_{\eta\eta}(x,y) - \frac{1}{2h^2}\{v(x-h,\,y+h)+v(x+h,\,y-h)-2v(x,y)\}\right| \leqq \frac{M_4}{6}h^2. \tag{4.8}$$

Then Δ_h^* is defined by

$$\begin{aligned}\Delta_h^* v(x,y) \equiv \frac{h^{-2}}{2}\Big\{ & v(x-h,\,y-h)+v(x+h,\,y-h)+\frac{\alpha-1}{\alpha+2}v(x+2h,\,y+2h) + \\ & + \frac{2(2-\alpha)}{\alpha+1}v(x+h,\,y+h)+\left[\frac{6}{\alpha(\alpha+1)(\alpha+2)}\right]v(x-\alpha h,\,y-\alpha h) - \\ & - \left[2+\frac{6+\alpha(7-\alpha^2)}{\alpha(\alpha+1)(\alpha+2)}\right]v(x,y)\Big\}.\end{aligned} \tag{4.9}$$

Thus having defined Δ_h^*, the set C_h of theorem 5 is fixed.

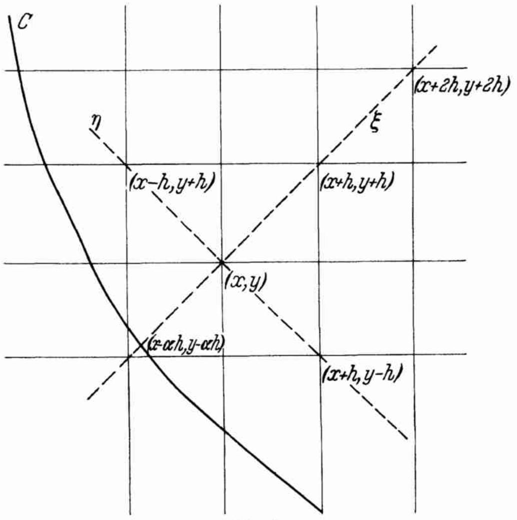

Fig. 3

In the finite difference problem (3.24) let

$$\varphi[F(P)] = \frac{h^2}{12}\Delta F(P), \tag{4.10}$$

and

$$H(P) = F(P). \tag{4.11}$$

For sufficiently small h, the hypotheses (a), (b) and (c) of theorem 5 are easily seen to be satisfied with $\delta = \frac{3}{4}$. The inequality (3.28) takes the form

$$\left|\Delta_h v - \left[\Delta v + \frac{h^2}{12}\Delta^2 v\right]\right| \leqq \frac{4}{5!}M_6 h^4. \tag{4.12}$$

Since for $P \in C_h^*$, $\sigma(P,P) \geqq \frac{2}{h^2}$ it follows that

$$\left|\frac{\Delta_h^* u(P) - F(P)}{\sigma(P,P)}\right| \leqq \frac{h^4}{6}M_4 + \frac{\sqrt{2}}{3}h^5 M_5. \tag{4.13}$$

applying theorem 5 we see that

$$|\varepsilon|_M \leqq \left[\frac{M_6 d^2}{120} + \frac{2M_4}{3} + \frac{4\sqrt{2}}{3}M_5 h\right]h^4. \tag{4.14}$$

In posing finite difference analogues of (2.11) one can combine the techniques of sections 2 and 3. To illustrate this we define an $O(h^4)$ approximation to (2.1).

Let R_h, C_h^* and C_h be defined as in section 2. The set of points of R_h whose eight nearest neighbors are not all in R will be denoted by C_h'. At points of $R_h - C_h'$ we define Δ_h by (4.1), and at points of C_h', Δ_h' is defined by (2.2). The operator Δ_h^* is taken to be an unbalanced $O(h^2)$ approximation to Δ, as in (4.5). We note that in this case only the pure second derivatives with respect to x and y are usesd.

We now pose the finite difference analogue of (2.1)

$$
\begin{aligned}
&\Delta_h U(P) = F(P) + \frac{h^2}{12}\Delta F(P), && P \in R_h - C_h' \\
&\Delta_h' U(P) = F(P), && P \in C_h' \\
&\Delta_h^* U(P) = F(P), && P \in C_h^* \\
&U(P) = f(P), && P \in C_h.
\end{aligned} \tag{4.15}
$$

This definition of Δ_h satisfies all hypotheses of theorem 5 except (3.28) with $n=4$. At points of C_h' (3.28) is valid for Δ_h' with $n=2$ and $\varphi(\Delta v) \equiv 0$. Hence theorem 5 is not directly applicable to (4.15). We can however modify the proof of theorem 5 to show that $|\varepsilon|_M = O(h^4)$.

Green's third identity (3.33) is valid for any vector $V(P)$, $P \in R_h + C_h^*$. If we apply (3.33) to the vector $V(P)=1$ for $P \in R_h$, $V(P)=0$ for $P \in C_h^*$, we obtain

$$
1 \geqq h^2 \sum_{Q \in C_h'} G_h^*(P,Q)[-\Delta_h' V(Q)] + h^2 \sum_{Q \in R_h - C_h'} G_h^*(P,Q)[-\Delta_h V(Q)]. \tag{4.16}
$$

Noting that the second term on the right hand side of (4.16) is non-negative and that $-\Delta_h V(Q) = 2h^{-2}$, for $Q \in C_h'$, we have that

$$
\sum_{Q \in C_h'} G_h^*(P,Q) \leqq \frac{1}{2}. \tag{4.17}
$$

Equation (3.33) can be written in the form

$$
\begin{aligned}
\varepsilon(P) = h^2 \sum_{Q \in C_h'} G_h^*(P,Q)[-\Delta_h' \varepsilon(Q)] + h^2 \sum_{Q \in R_h - C_h'} G_h^*(P,Q)[-\Delta_h \varepsilon(Q)] + \\
+ \sum_{Q \in C_h^*} G_h^*(P,Q)\,\varepsilon(Q).
\end{aligned} \tag{4.18}
$$

Thus we see that

$$
|\varepsilon(P)| \leqq \frac{h^2}{2}|\Delta_h' \varepsilon|_M + \frac{d^2}{16}|\Delta_h \varepsilon|_M + \left|\frac{\Delta_h^* \varepsilon(P)}{\sigma(P,P)}\right|_M + \frac{3}{4}|\varepsilon|_M. \tag{4.19}
$$

In this case it is easy to see that

$$
\left|\frac{\Delta_h^* \varepsilon(P)}{\sigma(P,P)}\right|_M \leqq \frac{M_4}{12} h^4 + \frac{M_5}{12} h^5. \tag{4.20}
$$

From (2.3), (4.12) and (4.20) we obtain the estimate

$$
|\varepsilon|_M \leqq \left[\frac{M_6 d^2}{120} + \frac{2M_4}{3} + \frac{M_5}{12} h\right] h^4. \tag{4.21}
$$

Note that the estimates (4.14) and (4.21) differ only in the higher order terms.

References

[1] Collatz, L.: Bemerkungen zur Fehlerabschätzung für das Differenzenverfahren bei partiellen Differentialgleichungen. Z. angew. Math. Mech. **13**, 56—57 (1933).

[2] — Numerical treatment of differential equations, 3rd ed. Berlin-Göttingen-Heidelberg: Springer 1960.

[3] Courant, R., K. Friedrichs and H. Lewy: Über die partiellen Differenzengleichungen der mathematischen Physik. Math. Ann. **100**, 32—74 (1928).

[4] Forsythe, G., and W. Wasow: Finite-difference methods for partial differential equations. New York: Wiley 1960.

[5] Gerschgorin, S.: Fehlerabschätzung für das Differenzenverfahren zur Lösung partieller Differentialgleichungen. Z. angew. Math. Mech. **10**, 373—382 (1930).

[6] Kantorovich, L., and V. Krylov: Approximate Methods of Higher Analysis. Netherlands: Noordhoff Ltd. 1958.

[7] Laasonen, P.: On the degree of convergence of discrete approximations for the solutions of the Dirichlet problem. Ann. Acad. Sci. Fenn. A. I. **246**, 1—19 (1957).

[8] — On the solution of Poisson's difference equation. J. Assoc. Comput. Mach. **5**, 370—382 (1958).

[9] Motzkin, T., and W. Wasow: On the approximation of linear elliptic differential equations by difference equations with positive coefficients. J. Math. Phys. **31**, 253—259 (1953).

[10] Shortley, G., and R. Weller: The numerical solution of Laplace's equation. J. Appl. Phys. **9**, 334—348 (1938).

[11] Walsh, J., and D. Young: On the accuracy of the numerical solution of the Dirichlet problem by finite differences. J. Res. Nat. Bur. Stand. **51**, 343—363 (1953).

[12] — — On the degree of convergence of solutions of difference equations to the solution of the Dirichlet problem. J. Math. Phys. **33**, 80—93 (1954).

[13] Wasow, W.: On the truncation error in the solution of Laplace's equation by finite differences. J. Res. Nat. Bur. Stand. **48**, 345—348 (1952).

[14] — The accuracy of difference approximations to plane Dirichlet problems with piecewise analytic boundary values. Quart. Appl. Math. **15**, 53—63 (1957).

Institute for Fluid Dynamics
and Applied Mathematics
University of Maryland
College Park, Maryland

(Received September 18, 1961)

Numerische Mathematik 4, 328—329 (1962)

d-Minimal Surfaces Spanning Polygons in E_n by Solving a Linear System

By

WALTER L. WILSON, JR.

1.0. Introduction

In [1], p. 368, §4.5, the author described a method involving solution of a linear system for computing approximations to conformal maps of regions bounded by polygons in a plane. Here, we extend the method to compute polyhedral approximations to minimal surfaces spanning a simple closed polygon Γ in Euclidean n-space.

Notation and operators used here are defined in the above cited paper.

Let Γ be a Jordan curve in the form of a polygon in E_n given by the vector function

$$\Gamma: g(t) \qquad 0 \leqq t \leqq l(\Gamma)$$

where $g(0) = g(l(\Gamma))$ is a vertex of Γ, t is arc length on Γ and the σ vertices of Γ are $\{g(t_k) \mid 0 = t_0 < t_1 < t_2 < \cdots < t_\sigma = l(\Gamma)\}$. Specifically, let

$$\text{(1.1)} \qquad \begin{aligned} g(t) &= g(t_k) + [g(t_{k+1}) - g(t_k)] \frac{t - t_k}{t_{k+1} - t_k} \\ &= g_k + \gamma_k t, \qquad t_k \leqq t \leqq t_{k+1} \end{aligned}$$

where $k = 0(1)\,\underline{\sigma - 1}$. These are σ line segments contained respectively in the lines

$$\text{(1.2)} \qquad L_k(t) = g_k + \gamma_k t, \qquad -\infty < t < +\infty.$$

Definition. $\Gamma_0: \{b_i(t) \mid i = 1(1)N\}$ is a discrete parametrization of Γ if (1) three elements of Γ_0, say b_{N-2}, b_{N-1} and b_N, are distinct vertices of Γ and are assumed *fixed*, (2) each b_j not one of the fixed points is *assigned* to a side of Γ and several (or no) elements of Γ_0 may be assigned to any side of Γ. The position of b_j on the side to which it is assigned is *not* prescribed. The orientation of elements of Γ_0 on Γ is prescribed*.

* The functional $A(\bar{\Gamma})$ in (2.1) below was derived assuming Γ is a topological image of a plane simple closed polygon with N vertices — the vertices corresponding to similarly oriented elements of any discrete parametrization of Γ. Therefore if m elements of Γ_0 are between two fixed points then all or any adjacent subset of these may be assigned to any side of Γ between the fixed points assuming the remainder of the set of m points are assigned to similarly oriented sides of Γ.

2.0. The Linear System

In [1] the author derived an analog

$$A(\bar{\Gamma}) = \sum_{i,j=1}^{N} F_{ij}(b_i - b_j)^2 \tag{2.1}$$

to the Douglas functional. (These are equations (4.5) and (1.1) respectively in [1].) $A(\bar{\Gamma})$ is defined on the discrete parametrizations of Γ. If $\underline{\Gamma_p}: \{b_i | i = 1(1)N\}$ is a discrete parametrization of Γ which locally minimizes $A(\bar{\Gamma})$ in the set of all discrete parametrizations containing the same three fixed points and the same number of points on each arc of Γ between fixed points then Γ_p is called minimizing. The unique d-harmonic surface Σ_p spanning the polygon Γ_p is a d-minimal surface. The vertices of Σ_p are defined by the Laplacian operator L in equation (3.10) in [1]. We now define the linear system which is used to determine a minimizing Γ_p. The particular Γ_p obtained depends on the choice of fixed points and assignments in Γ_0.

Let Γ_0 be a discrete parametrization of Γ as defined above. Then the positive definite quadratic form $A(\Gamma_0) = A(t)$ is minimized only if the non-fixed elements of Γ_0 are chosen so that $\frac{dA}{dt} = \frac{dA}{db_m} \cdot \frac{db_m}{dt} = 0$ That is, if

$$\sum_{\beta} F_{m\beta}(b_m - b_\beta) \cdot \dot{b}_m = -\sum_{\alpha=1}^{n} C_{m\alpha} \sum_{\beta=1}^{N} F_{m\beta} b_{\beta\alpha} = 0^{\star}, \qquad m = 1(1)\underline{N-3}, \tag{2.2}$$

where $C_{m\alpha}$ is the α-th component of the vector γ defined in (1.1) for the side of Γ to which b_m was assigned in Γ_0. These are $N-3$ equations in $n(N-3)$ unknowns — the components of elements of Γ_0.

If b_β is assigned to the side of Γ contained in the line $L_k(t)$ then we make the substitution $b_{\beta\alpha} = (g_k + \gamma_k a_\beta)_\alpha \equiv g_{\beta\alpha} + \gamma_{\beta\alpha} a_\beta$ where a_β is to be determined. Then, (2.2) may be written in the form

$$\sum_{\alpha=1}^{n} \gamma_{m\alpha} \sum_{\beta=1}^{N} F_{m\beta}(g_{\beta\alpha} + \gamma_{\beta\alpha} a_\beta) = 0, \qquad m = 1(1)\underline{N-3} \tag{2.3}$$

where a_{N-2}, a_{N-1} and a_N are fixed.

The system of $N-3$ equations (2.3) is solved for $\{a_\beta | \beta = 1(1)\underline{N-3}\}$ which determine a set $\bar{\Gamma}_p$ of points some of which may lie on the extensions of the sides to which they were assigned in Γ_0. If any element of $\bar{\Gamma}_p$ is on the extension of the assigned side then a reassignment of elements giving a new Γ_0 is necessary. When elements of $\bar{\Gamma}_p$ all lie on the assigned sides of Γ then $\bar{\Gamma}_p$ is Γ_p — a minimizing discrete parametrization of Γ. The d-minimal surface is then determined as indicated above.

References

[1] WILSON JR., W. L.: On Discrete Dirichlet and Plateau Problems. Numerische Mathematik 3, 359—373 (1961).

Department of Mathematics
University of Alabama
University, Alabama

(Received January 11, 1962)

$\star$ The first equality is obtained using the definition of scalar product in E_n and Corollary 4.1 in [1].