JOURNAL OF APPROXIMATION THEORY 29, 305–322 (1980)

On the Approximation of a Bivariate Function by the Sum of Univariate Functions

W. A. LIGHT

University of Lancaster, Lancaster, England

AND

E. W. CHENEY*

University of Texas, Austin, Texas 78712

Communicated by the Editors

Received March 12, 1979

DEDICATED TO THE MEMORY OF P. TURÁN

INTRODUCTION

The approximation problem considered here is to represent or approximate a prescribed continuous and real-valued function of two variables by the sum of two continuous functions of one variable:

$$f(x, y) \approx g(x) + h(y).$$

To make the problem more precise, let X and Y be compact topological spaces, and let $f \in C(X \times Y)$. Denote by M the set of functions ϕ which have the form

$$\phi(x, y) = g(x) + h(y) \qquad g \in C(X), h \in C(Y).$$

The *distance* from f to M is defined by

$$\operatorname{dist}(f, M) = \inf_{\phi \in M} \|f - \phi\| = \inf_{\phi \in M} \sup_{(x,y)} |f(x, y) - \phi(x, y)|.$$

An element $\phi \in M$ is sought such that $\|f - \phi\| = \operatorname{dist}(f, M)$. Such a ϕ is termed a *best approximation* to f. Alternatively, one asks for an $f^* \in C(X \times Y)$ such that $f^* - f \in M$ and $\|f^*\|$ is a minimum, i.e., $\|f^*\| = \operatorname{dist}(f, M)$.

* This author was supported by the Science Research Council of the U.K. and by the Office of Research, U. S. Army at The University of Texas at Austin.

0021-9045/80/080305-18$02.00/0

640/29/4-4

In 1951, Diliberto and Straus published in [3] an algorithm for solving this problem, and in [4], Golomb has shown that a generalized form of this algorithm is applicable in any normed linear space. The procedure of [3] is most easily explained in terms of two non-linear averaging operators defined by the equations

$$(\mathscr{M}_x f)(y) = \tfrac{1}{2} \sup_x f(x, y) + \tfrac{1}{2} \inf_x f(x, y),$$

$$(\mathscr{M}_y f)(x) = \tfrac{1}{2} \sup_y f(x, y) + \tfrac{1}{2} \inf_y f(x, y).$$

The algorithm is then simply

$$f_1 = f, \qquad f_{2n} = f_{2n-1} - g_n, \qquad f_{2n+1} = f_{2n} - h_n,$$
$$g_n = \mathscr{M}_y f_{2n-1}, \qquad h_n = \mathscr{M}_x f_{2n}.$$

One of the principal results of [3] can be summarized thus:

THEOREM. *The sequence $\{f_n\}$ is equicontinuous and possesses cluster points. Each cluster point f^* is a solution to the problem: $f^* - f \in M$ and $\| f^* \| =$* dist(f, M). *Furthermore, $\mathscr{M}_x f^* = \mathscr{M}_y f^* = 0$, $\| f_n \| \downarrow$* dist(f, M), *and* $\| f_{n-1} - f_n \| \downarrow 0$.

1. Convergence of Functions in the Algorithm

One of the questions left unanswered in [3] is whether the sequence $\{f_n\}$ converges. This was answered affirmatively by Aumann in [2]. We give a new proof of this result in Theorem 1.1 below.

LEMMA 1.1. $\| \mathscr{M}_x f - \mathscr{M}_x F \| \leqslant \| f - F \|$. *Similarly for* $\mathscr{M}_y$.

Proof. Let $\delta = \| f - F \|$. Then we have the pointwise inequality

$$-\delta + F \leqslant f \leqslant \delta + F.$$

Since $\mathscr{M}_x$ is order-preserving, and constant-preserving,

$$-\delta + \mathscr{M}_x F \leqslant \mathscr{M}_x f \leqslant \delta + \mathscr{M}_x F.$$

This is equivalent to the inequality in the lemma. ∎

In the following three lemmas we give some of Aumann's results [2]. Proofs are included for completeness.

The following lemma is elementary.

LEMMA 1.2. *For two functions f and F,*

$$\max f(x) - \max F(x) \leqslant \max\{f(x) - F(x)\}$$

with equality holding only if there exists a point at which all three maxima are attained.

LEMMA 1.3. [Aumann, [2]]. *Let τ_n denote the maximum of $|f_n(p)|$ as p ranges over points satisfying $|f_{n+1}(p) - f_n(p)| = \| f_{n+1} - f_n \|$. If $\| f_{n+1} - f_n \| = \| f_n - f_{n-1} \|$ then $\tau_{n-1} \geqslant \tau_n + \| f_n - f_{n-1} \|$.*

Proof. We suppose $n = 2k$; the case when n is odd is similar. Using Lemma 1.1 and the fact that $\mathscr{M}_x f_{2k-1} = 0$, we have

$$\| h_k \| = \| \mathscr{M}_x f_{2k} \| = \| \mathscr{M}_x f_{2k} - \mathscr{M}_x f_{2k-1} \| \leqslant \| f_{2k} - f_{2k-1} \| = \| g_k \|. \quad (1)$$

By the definition of τ_{2k} there exists y_0 such that

$$| h_k(y_0)| = \| h_k \| \qquad \text{and} \qquad \tau_{2k} = \max_x | f_{2k}(x, y_0)|.$$

Since $\| h_k \| = \| g_k \|$ by hypothesis, one of the following holds:

(i) $(\mathscr{M}_x f_{2k} - \mathscr{M}_x f_{2k-1})(y_0) = \| g_k \|$,

(ii) $(\mathscr{M}_x f_{2k} - \mathscr{M}_x f_{2k-1}(y_0) = -\| g_k \|$.

Suppose (i) holds. Then rewrite (i) in the forms

$$\tfrac{1}{2}[\max_x f_{2k}(x, y_0) - \max_x f_{2k-1}(x, y_0)] + \tfrac{1}{2}[\min_x f_{2k}(x, y_0) - \min_x f_{2k-1}(x, y_0)] = \| g_k \|,$$

$$\tfrac{1}{2}[\max_x f_{2k}(x, y_0) - \max_x f_{2k-1}(x, y_0)] + \tfrac{1}{2}[\max_x -f_{2k-1}(x, y_0) - \max_x -f_{2k}(x, y_0)] = \| g_k \|.$$

By Lemma 1.2, each bracketed expression is at most $\| g_k \|$, and hence by the equality condition of Lemma 1.2, there exists an x_0 such that $-g_k(x_0) = \| g_k \|$, $f_{2k}(x_0, y_0) = \max_x f_{2k}(x, y_0)$, and $f_{2k-1}(x_0, y_0) = \max_x f_{2k-1}(x, y_0)$. From the definition of g_k, $g_k(x_0) = \frac{1}{2} \max_y f_{2k-1}(x_0, y) + \frac{1}{2} \min_y f_{2k-1}(x_0, y)$ whence

$$\begin{aligned} \tau_{2k-1} &\geqslant -\min_y f_{2k-1}(x_0, y) = -2g_k(x_0) + \max_y f_{2k-1}(x_0, y) \\ &\geqslant -2g_k(x_0) + f_{2k-1}(x_0, y_0) = -g_k(x_0) + f_{2k}(x_0, y_0) \\ &= \| g_k \| + \tau_{2k}. \end{aligned}$$

Case (ii) is similar. ∎

LEMMA 1.4 [Aumann, 2]. *In the Diliberto–Straus algorithm, we have* $g_n \to 0$ *and* $h_n \to 0$.

Proof. Diliberto and Straus [3] stated that $\| g_2 \| \geqslant \| h_2 \| \geqslant \| g_3 \| \geqslant \cdots$. One proves this easily by Eq. (1) in the previous proof.
Thus we can define $\theta = \lim \| f_{n+1} - f_n \|$. Because of equicontinuity (see [3] or Section 2, below), there is a convergent subsequence, say $f_{n_k} \to f^*$. Applying the algorithm to f^*, we obtain a sequence f_m^*. Since the operations in the algorithm are continuous, $f_m^* = \lim_k f_{n_k+m}$. Hence $\| f_{m+1}^* - f_m^* \| = \lim_k \| f_{n_k+m+1} - f_{n_k+m} \| = \theta$. By Lemma 1.3 (applied to f_m^*) we have $\tau_m^* \leqslant \tau_{m+1}^* - \theta, m = 2, 3, \ldots$. By induction, this leads to $0 \leqslant \tau_m^* \leqslant \tau_1^* - (m-1)\theta$, whence $\theta = 0$. ∎

If f is fixed, we define operators A and B by putting

$$AF = \mathscr{M}_x(f - F), \qquad BF = \mathscr{M}_y(f - F).$$

From Lemma 1.1 we have immediately:

LEMMA 1.5.

$$\begin{aligned} \| A\phi - A\psi \| &\leqslant \| \phi - \psi \|, \\ \| B\phi - B\psi \| &\leqslant \| \phi - \psi \|, \\ \| AB\phi - AB\psi \| &\leqslant \| \phi - \psi \|. \end{aligned}$$

THEOREM 1.1. *The sequence* $\{f_n\}$ *produced by the Diliberto–Straus algorithm converges uniformly.*

Proof. Let $G_n = \sum_1^n g_i$ and $H_n = \sum_1^n h_i$. Then

$$\begin{aligned} G_n &= G_{n-1} + g_n = G_{n-1} + \mathscr{M}_y f_{2n-1} = G_{n-1} + \mathscr{M}_y(f - G_{n-1} - H_{n-1}) \\ &= \mathscr{M}_y(f - H_{n-1}) = BH_{n-1}. \end{aligned}$$

Similarly, $H_n = AG_n$. Hence $H_n = ABH_{n-1}$.

Select g and h so that $f - g - h$ is level (see [3, Theorem 7] or Theorem 2.3, below). Then $0 = \mathscr{M}_x(f - g - h) = Ag - h$, so $h = Ag$. Similarly, $g = Bh$. Hence $h = ABh$. Since $\| H_n - h \| = \| ABH_{n-1} - ABh \| \leqslant \| H_{n-1} - h \|$ by Lemma 1.5, we see by induction that the sequence $\{H_n\}$ is bounded. It is equicontinuous, as shown in [3, Theorem 4] or by the argument in Theorem 2.3 of this paper. Hence by the Ascoli theorem, there is a convergent subsequence, say $H_{n_k} \to H$.

Since $G_n = BH_{n-1}$, we see that G_{n_k+1} converges, say to G. Since $g_n \to 0$ [3, Theorem 6] we conclude that $G_{n_k} \to G$. It follows that the sequence $f_{2n_k+1} = f - G_{n_k} - H_{n_k}$ converges to $f - G - H$. As in [3, Theorem 7], $f - G - H$ is level. As in the argument above for h, $ABH = H$. By repeating another argument used above, $\| H_n - H \| \leqslant \| H_{n-1} - H \|$. This means that the sequence $\{H_n\}$ can have no cluster point other than H. Hence $H_n \to H$, $G_n \to G$, and $f_{2n+1} \to f - G - H$. Then $f_{2n} \to f - G - H$ and $f_n \to f - G - H$. ▮

2. New Proofs of Two Theorems of Diliberto and Straus

It is convenient to term a function f *horizontally level* if $\mathscr{M}_x f = 0$, *vertically level* if $\mathscr{M}_y f = 0$, and simply *level* if $\mathscr{M}_x f = \mathscr{M}_y f = 0$. The elements f_{2n} in the algorithm are vertically level since $\mathscr{M}_y f_{2n} = \mathscr{M}_y(f_{2n-1} - g_n) = \mathscr{M}_y f_{2n-1} - g_n = 0$. Similarly $\mathscr{M}_x f_{2n+1} = 0$. The following two lemmas can be found in [4]. Proofs are included here because of their brevity.

Lemma 2.1 [Golomb [4]]. *If $\mathscr{M}_y f = 0$ and $g \in C(X)$, then $\| f + g \| = \| f - g \|$.*

Proof. Select $p = (x_0, y_0)$ so that $(f + g)(p) = \sigma \| f + g \|$, $\sigma = \pm 1$. Since f is vertically level, there is a point $q = (x_0, y_1)$ such that $\sigma[f(p) + f(q)] \leqslant 0$. Then $\| f - g \| \geqslant \sigma[g(q) - f(q)] = \sigma g(p) - \sigma f(q) \geqslant \sigma[g(p) + f(p)] = \| f + g \|$. Apply the argument to $-g$ in order to get $\| f + g \| \geqslant \| f - g \|$. ▮

Lemma 2.2 [Golomb [4]]. *For all n, $\| f_n - 2f_{n+1} \| = \| f_n \|$.*

Proof. If $n = 2k - 1$ then by Lemma 2.1, $\| f_n - 2f_{n+1} \| = \| f_{2k-1} - f_{2k} - f_{2k} \| = \| g_k - f_{2k} \| = \| g_k + f_{2k} \| = \| f_{2k-1} \|$. The proof for even n is similar. ▮

Lemma 2.3. *For $n \geqslant 2$, $\| g_n \| \leqslant \frac{1}{2} \| f_{2n-2} \|$ and $\| h_n \| \leqslant \frac{1}{2} \| f_{2n-1} \|$.*

Proof. By Lemma 2.2, $\| f_n - 2f_{n+1} \| \leqslant \| f_n \|$, whence

$$f_n - \| f_n \| \leqslant 2f_{n+1} \leqslant f_n + \| f_n \|.$$

If $n = 2k$, we apply $\mathscr{M}_y$ to each term in this inequality and use the monotonicity of $\mathscr{M}_y$, obtaining

$$\mathscr{M}_y f_{2k} - \| f_{2k} \| \leqslant 2\mathscr{M}_y f_{2k+1} \leqslant \mathscr{M}_y f_{2k} + \| f_{2k} \|.$$

Since f_{2k} is vertically level, $\mathscr{M}_y f_{2k} = 0$, and the previous inequality reduces to

$$-\| f_{2k} \| \leqslant 2g_{k+1} \leqslant \| f_{2k} \|.$$

The proof of the other inequality is similar. ▮

LEMMA 2.4. *If $n > N$ then* $2f_n - \|f_N\| \leqslant f_{n-1} \leqslant 2f_n + \|f_N\|$.

Proof. By Lemma 2.2 and the monotonicity of $\|f_n\|$,

$$\|f_{n-1} - 2f_n\| = \|f_{n-1}\| \leqslant \|f_N\|.$$

Hence

$$-\|f_N\| \leqslant f_{n-1} - 2f_n \leqslant \|f_N\|$$

and

$$2f_n - \|f_N\| \leqslant f_{n-1} \leqslant 2f_n + \|f_N\|. \quad \blacksquare$$

LEMMA 2.5 [5, Lemma 2.2]. *If f differs from a vertically level function by a function of y alone then* $\max_y[f(x_1, y) - f(x_2, y)] \geqslant 0$ *for all x_1 and x_2.*

An ordered set of points $[p_1, p_2, \ldots]$ in $X \times Y$ is called a *path* if $p_i = (x_i, y_i)$, $x_{2i-1} = x_{2i}$, $y_{2i+1} = y_{2i}$ [or $x_{2i+1} = x_{2i}$ and $y_{2i-1} = y_{2i}$]. If the path is finite, if it has an even number of terms, say $[p_1, p_2, \ldots, p_{2n}]$, and if $y_1 = y_{2n}$, we call it a *closed path*, l. (We do not assume that the points p_i are distinct.) Corresponding to l there is a linear functional π_l defined by

$$\pi_l(f) = \frac{1}{2n} \sum_{i=1}^{2n} (-1)^i f(p_i).$$

The principal properties of these functionals, as established by Diliberto and Straus, are these:

THEOREM. (1) $\pi_l(\phi) = 0$ *for all* $\phi \in M$.

(2) $\|\pi_l\| \leqslant 1$.

(3) $\operatorname{dist}(f, M) = \sup_l \pi_l(f)$.

Although most of the results of this paper are valid for arbitrary compacta X and Y, it is convenient to retain the terminology of $\mathbb{R}^2$. For example, "$p - q$ is horizontal" means that $p = (x_0, y_0)$ and $q = (x_1, y_0)$.

THEOREM 2.1. *For any two integers $N \geqslant 2$ and $k \geqslant 1$,*

$$\frac{k+2}{k+1} \operatorname{dist}(f, M) \geqslant \|f_N\| - 2^k(\|f_N\| - \|f_{N+k}\|).$$

Thus it follows immediately that $\lim \|f_N\| = \operatorname{dist}(f, M)$.

Proof. Fix N and k, and let $n = N + k$. It is convenient but not necessary to assume that n is even. Then f_n is vertically level. Hence there exist points

p_0 and q_0 such that $p_0 - q_0$ is vertical and $f_n(p_0) = -f_n(q_0) = \|f_n\|$. By Lemma 2.4,

$$f_{n-1}(p_0) \geqslant 2f_n(p_0) - \|f_N\| = 2\|f_n\| - \|f_N\|,$$
$$f_{n-1}(q_0) \leqslant 2f_n(q_0) + \|f_N\| = -2\|f_n\| + \|f_N\|.$$

Since f_{n-1} is horizontally level, there exist points p_1 and q_1 such that $p_0 - q_1$ is horizontal, $q_0 - p_1$ is horizontal, and

$$f_{n-1}(p_1) \geqslant 2\|f_n\| - \|f_N\|,$$
$$f_{n-1}(q_1) \leqslant -2\|f_n\| + \|f_N\|.$$

We shall prove by induction that the following assertion, $A(r)$, is true for $r = 1, 2,..., k$.

$A(r)$: There exist points $p_0, p_1, p_2,..., p_r$ and $q_0, q_1,..., q_r$ forming a path such that

$$f_{n-r}(p_i) \geqslant 2^r\|f_n\| - (2^r - 1)\|f_N\| \text{ and } f_{n-r}(q_i) \leqslant -2^r\|f_n\| + (2^r - 1)\|f_N\|.$$

The preceding remarks have established $A(1)$. Now suppose that $A(r)$ has been established for a particular r in $\{1, 2,..., k-1\}$. In proving $A(r+1)$, we suppose first that r is even. Then $n - r - 1$ is odd and f_{n-r-1} is horizontally level. Select p_{r+1} so that $p_{r+1} - q_r$ is horizontal and $f_{n-r-1}(p_{r+1}) \geqslant -f_{n-r-1}(q_r)$. By Lemma 2.4 and the induction hypothesis, $A(r)$, we have

$$\begin{aligned} f_{n-r-1}(p_{r+1}) &\geqslant -f_{n-r-1}(q_r) \geqslant -2f_{n-r}(q_r) - \|f_N\| \\ &\geqslant -2[-2^r\|f_n\| + (2^r - 1)\|f_N\|] - \|f_N\| = 2^{r+1}\|f_n\| - (2^{r+1} - 1)\|f_N\|. \end{aligned}$$

The choice and analysis of q_{r+1} is similar. For indices $i = 0, 1,..., r$ we have by Lemma 2.4 and the induction hypothesis

$$\begin{aligned} f_{n-r-1}(p_i) &\geqslant 2f_{n-r}(p_i) - \|f_N\| \\ &\geqslant 2[2^r\|f_n\| - (2^r - 1)\|f_N\|] - \|f_N\| \\ &= 2^{r+1}\|f_n\| - (2^{r+1} - 1)\|f_N\|. \end{aligned}$$

A similar analysis applies to the points q_i. If r is odd, a similar proof can be given.

Thus $A(r)$ is true for $r = k$, and there exist points $p_0,..., p_k$, $q_0,..., q_k$ such that $f_N(p_i) \geqslant 2^k\|f_n\| - (2^k - 1)\|f_N\| = \|f_N\| - 2^k(\|f_N\| - \|f_n\|)$ and similarly for q_i. Now complete a closed path by constructing p_{k+1} and q_{k+1} in such a way that $f_N(p_{k+1}) \geqslant f_N(q_{k+1})$. This is possible by Lemma 2.5. Then we have

$$\begin{aligned} \operatorname{dist}(f, M) \geqslant \pi_l(f_N) &= \frac{1}{2k+4}\left\{\sum_{i=0}^{k} [f_N(p_i) - f_N(q_i)] + f_N(p_{k+1}) - f_N(q_{k+1})\right\} \\ &\geqslant \frac{2k+2}{2k+4}\{\|f_N\| - 2^k(\|f_N\| - \|f_n\|)\}. \quad \blacksquare \end{aligned}$$

COROLLARY 2.1. $\frac{3}{2}$ dist$(f, M) \geqslant 2 \|f_{n+1}\| - \|f_n\|$, *if* $n \geqslant 2$.

THEOREM 2.2. *If the Diliberto–Straus algorithm is applied on a domain which is a subset of* $X \times Y$ *then for* $n \geqslant 2$ *and* $k \geqslant 1$,

$$\operatorname{dist}(f, M) \geqslant \frac{k+1}{k+2}\{\|f_n\| - 2^k(\|f_n\| - \|f_{n+k}\|)\} - \frac{1}{k+2}\|f_n\|.$$

Hence dist$(f, M) = \lim_{n\to\infty} \|f_n\|$.

Proof. In the proof of the preceding theorem, the only change to be made is at the end. Lemma 2.5 is not applicable, but the path can be completed by any two convenient points p_{k+1} and q_{k+1}. In the estimate, use

$$f_N(p_{k+1}) - f_N(q_{k+1}) \geqslant -2\|f_N\|. \quad ∎$$

LEMMA 2.6 [5]. *For functions of one variable, the averaging operator* $\mathscr{M}f = \frac{1}{2}\sup f(x) + \frac{1}{2}\inf f(x)$ *has these properties*:

(1) $|\mathscr{M}v| \leqslant \|v\|$,

(2) $|\mathscr{M}v_1 - \mathscr{M}v_2| \leqslant \|v_1 - v_2\|$,

(3) $\|v - \mathscr{M}v\| = \|v\| - |\mathscr{M}v|$.

DEFINITIONS. Fixing $f \in C(X \times Y)$, we define mappings A, B, and S on M as follows:

(1) $A\phi = \mathscr{M}_x(f - \phi)$,

(2) $B\phi = \mathscr{M}_y(f - \phi)$,

(3) $S\phi = \phi + A\phi + B(\phi + A\phi)$.

We define also

$$\Delta(x_1, y_1, x_2, y_2) = \sup_x |f(x, y_1) - f(x, y_2)| + \sup_y |f(x_1, y) - f(x_2, y)|,$$

$$K = \{\phi \in M: \|f - \phi\| \leqslant \|f\|, |\phi(x_1, y_1) - \phi(x_2, y_2)| \leqslant \Delta(x_1, y_1, x_2, y_2)\}.$$

A standard compactness argument shows that $\{f(x, \cdot): x \in X\}$ and $\{f(\cdot, y): y \in Y\}$ are equicontinuous sets in $C(Y)$ and $C(X)$, respectively. Hence Δ can be made as small as we wish by restricting (x_2, y_2) to a neighborhood of (x_1, y_1).

THEOREM 2.3. *The set* K *is nonempty, convex, and compact. The map* S *is continuous and carries* K *into* K. *Hence* S *has fixed points in* K. *If* ϕ *is any fixed point of* S *in* M *then* $\phi \in K$, $f - \phi$ *is level, and* $\|f - \phi\| = $ dist(f, M).

Proof. The maps A, B, S are continuous by Lemma 2.6. For any $\phi \in M$, we have by Lemma 2.6,

$$\begin{aligned} \| f - S\phi \| &= \| f - \phi - A\phi - \mathscr{M}_y(f - \phi - A\phi)\| \\ &\leqslant \| f - \phi - A\phi \| = \| f - \phi - \mathscr{M}_x(f - \phi)\| \\ &\leqslant \| f - \phi \|. \end{aligned}$$

This shows that if $\phi \in K$ then $\| f - S\phi \| \leqslant \| f \|$.

If $\phi(x, y) = g(x) + h(y)$, then

$$\begin{aligned} S\phi &= g + h + \mathscr{M}_x(f - g - h) + B(\phi + A\phi) \\ &= g + Ag + B(g + Ag) \\ &= Ag + BAg. \end{aligned}$$

Now let $\phi \in M$ and $u = S\phi$. If $\phi(x, y) = g(x) + h(y)$, then by Lemma 2.6,

$$\begin{aligned} &| u(x_1, y_1) - u(x_2, y_2)| \\ &\quad = |(Ag)(x_1, y_1) - (Ag)(x_2, y_2) + (BAg)(x_1, y_1) - (BAg)(x_2, y_2)| \\ &\quad = |(Ag)(y_1) - (Ag)(y_2) + (BAg)(x_1) - (BAg)(x_2)| \\ &\quad \leqslant | \mathscr{M}_x(f - g)(y_1) - \mathscr{M}_x(f - g)(y_2)| \\ &\qquad + | \mathscr{M}_y(f - Ag)(x_1) - \mathscr{M}_y(f - Ag)(x_2)| \\ &\quad \leqslant \sup_x |(f - g)(x, y_1) - (f - g)(x, y_2)| \\ &\qquad + \sup_y |(f - Ag)(x_1, y) - (f - Ag)(x_2, y)| \\ &\quad = \sup_x | f(x, y_1) - f(x, y_2)| + \sup_y | f(x_1, y) - f(x_2, y)|. \end{aligned}$$

This completes the proof that S carries K into K.

For the convexity of K, let ϕ_1 and ϕ_2 belong to K, and let $0 \leqslant \lambda \leqslant 1$. Put $\phi = \lambda\phi_1 + (1 - \lambda)\,\phi_2$. Then clearly, $\| f - \phi \| \leqslant \| f \|$. To show that $| \phi(x_1, y_1) - \phi(x_2, y_2)| \leqslant \Delta(x_1, y_1, x_2, y_2)$ we compute as follows, with $p = (x_1, y_1)$ and $q = (x_2, y_2)$:

$$\begin{aligned} | \phi(p) - \phi(q)| &= | \lambda\phi_1(p) + (1 - \lambda)\,\phi_2(p) - \lambda\phi_1(q) - (1 - \lambda)\,\phi_2(q)| \\ &\leqslant \lambda\, | \phi_1(p) - \phi_1(q)| + (1 - \lambda)|\phi_2(p) - \phi_2(q)| \\ &\leqslant \lambda\Delta + (1 - \lambda)\Delta = \Delta(x_1, y_1, x_2, y_2). \end{aligned}$$

A simple argument shows that M is a closed subspace. We have proved that K is bounded and equicontinuous. In order to prove K closed, let $\phi_n \in K$ with $\phi_n \to \phi$. (Pointwise convergence suffices.) The two conditions for membership in K are then obviously satisfied by ϕ.

By the Ascoli theorem, K is compact. By the Schauder Fixed-Point Theorem, S has a fixed point in K.

Let ϕ be a fixed point of S in M. Then $A\phi + B(\phi + A\phi) = 0$. Since $A\phi$ is a function of y alone and $B(\phi + A\phi)$ is a function of x alone, we have $A\phi = c$ and $B(\phi + A\phi) = -c$ for some constant c. It follows that $-c = B(\phi + A\phi) = B(\phi + c) = -c + B(\phi)$. Hence $0 = B(\phi) = \mathscr{M}_y(f - \phi)$. The equation $\mathscr{M}_x(f - \phi) = A\phi = 0$ follows from Lemma 2.7 (below). Hence $f - \phi$ is level, and $\|f - \phi\| = \text{dist}(f, M)$. ∎

LEMMA 2.7. *If* $\mathscr{M}_x f = 0$ *and* $\mathscr{M}_y f = c$, *then* $c = 0$.

Proof. By Lemma 2.6,

$$\begin{aligned}\|f\| &= \|f - c - \mathscr{M}_x f + c\| \\ &= \|(f - c) - \mathscr{M}_x(f - c)\| \leqslant \|f - c\| = \|f - \mathscr{M}_y f\| \\ &= \max_x \{\max_y |f(x, y)| - |(\mathscr{M}_y f)(x)|\} = \|f\| - |c|. \quad ∎\end{aligned}$$

Remark. If ϕ is a discontinuous function of the form $h(y) + g(x)$, then $S\phi$ is a continuous function satisfying $\|f - S\phi\| \leqslant \|f - \phi\|$. Hence, as pointed out by Diliberto and Straus, the approximation of a continuous f cannot be improved by allowing discontinuous $g(x) + h(y)$.

Remark. The mapping S is defined so that if $f_{2n} = f - \phi$ then $f_{2n+2} = f - S\phi$.

Remark. The fixed-point theorem is not necessary in proving the existence of a best approximation to f in M. Once S has been proved to have the properties $\|f - S\phi\| \leqslant \|f - \phi\|$ and $S(K) \subset K$, the existence follows from compactness of K and continuity of $\|f - \phi\|$. The fixed-point argument produces a best approximation ϕ such that $f - \phi$ is level, a stronger result.

3. SPECIAL RESULTS IN THE DISCRETE CASE

Our objective here is to investigate the structure of the set of best approximations when X and Y are finite sets. A unicity theorem for best approximations is one by-product. Insofar as possible, definitions and results are given for general X and Y.

DEFINITION 3.1. An *extremal path* for a function f is a closed path l such that $\pi_l(f) = \text{dist}(f, M)$. (Recall that the inequality $\leqslant$ is always valid here.) The union of all extremal paths for f, considered simply as sets, will be denoted by $E(f)$. The following lemma is from [5].

LEMMA 3.1. *If* $\|f\| = \|f + \phi\| = \operatorname{dist}(f, M)$ *then* $\phi(p) = 0$ *for* $p \in E(f)$. (Here $\phi \in M$.)

DEFINITION 3.2. The *null points* belonging to f are the elements of the set

$$\mathcal{N}(f) = \bigcap \{Z(\phi) \colon \phi \in M \text{ and } \|f^* + \phi\| = \|f^*\|\}$$

where f^* is chosen so that $f^* - f \in M$ and $\|f^*\| = \operatorname{dist}(f, M)$, and $Z(\phi)$ denotes the set $\{p \colon \phi(p) = 0, p \in X \times Y\}$. (The null points are points where any two best approximations of f must agree.)

LEMMA 3.2. *The definition of null points is independent of the choice of* f^*.

Proof. Let $f' - f \in M$, $\|f'\| = \operatorname{dist}(f, M) = r$, and $p \in \bigcap\{Z(\phi)\colon \|f^* + \phi\| = r\}$. If $\|f' + \phi\| = r$ then

$$r = \|f' + \phi\| = \|f^* + (f' - f + \phi)\| = \|f^* + (f' - f^*)\|.$$

Hence $(f' - f^* + \phi)(p) = (f' - f^*)(p) = 0$ and $\phi(p) = 0$. Thus $p \in \bigcap\{Z(\phi)\colon \|f' + \phi\| = r\}$. ∎

LEMMA 3.3. *If* $f - f' \in M$ *then* f *and* f' *have the same null points and extremal paths.*

Proof. If $f - f' \in M$ then $\pi_l(f) = \pi_l(f')$ for all closed paths l. Also, $\operatorname{dist}(f, M) = \operatorname{dist}(f', M)$. Hence the extremal paths are the same. If $f^* - f \in M$ and $\operatorname{dist}(f, M) = \|f^*\|$, then $f^* - f' \in M$ and $\operatorname{dist}(f', M) = \|f^*\|$. By the preceding lemma, f and f' have the same null points. ∎

THEOREM 3.1. *Let* X *and* Y *be finite. To each function* f *defined on* $X \times Y$ *there corresponds a function* f^* *such that* $f^* - f \in M$, f^* *is level, and* $\operatorname{crit}(f^*) = E(f)$ (*i.e., each critical point of* f^* *is on an extremal path of* f.) $\operatorname{crit} f = \{p \colon |f(p)| = \|f\|\}$.

Proof. By Theorem 3.2 of [3] there exists an f' such that $f' - f \in M$, $\operatorname{crit}(f') = E(f)$, and $\|f'\| = \operatorname{dist}(f, M)$. Let $f' = f_1$ be the starting point of the Diliberto–Straus algorithm. Let f^* be the limit of the generated sequence $\{f_n\}$. Then $f^* - f \in M$ and f^* is level. It remains to be shown that $\operatorname{crit}(f^*) = E(f)$.

Let X_1 and Y_1 denote, respectively, the projections of E onto the X and Y sets. Put $X_0 = X \backslash X_1$ and $Y_0 = Y \backslash Y_1$. Let c denote the maximum of $|f'(p)|$ on the set $L = (X \times Y_0) \cup (X_0 \times Y)$. Since $\operatorname{crit}(f') \subset X_1 \times Y_1$, we have $c < \|f'\|$.

We now prove that for $n \geqslant 1$, $f_{2n+1} = f'$ on $X_1 \times Y_1$ and $|f_{2n+1}| < c$ on L. Since the step from f_1 to f_3 is just like the step from f_{2n-1} to f_{2n+1}, we need to consider f_3 only.

By the definitions of X_1 and g_1, $g_1 = 0$ on X_1 and $f_2 = f'$ on $X_1 \times Y$. Likewise $h_1 = 0$ on Y_1 and $f_3 = f_2$ on $X \times Y_1$. Consequently $f_3 = f'$ on $X_1 \times Y_1$.

Now let $p = (x_0, y_0)$. Using Lemma 2.6, we obtain the following.

(1) If $p \in X_0 \times Y$ then $|f_2(p)| \leqslant \max_y |f_2(x_0, y)| \leqslant \max_y |f_1(x_0, y)| \leqslant c$.

(2) If $p \in X_1 \times Y_0$ then $|f_2(p)| = |f_1(p)| \leqslant c$.

(3) By (1) and (2), if $p \in X \times Y_0$ then $|f_2(p)| \leqslant c$.

(4) By (1), if $p \in X_0 \times Y_1$ then $|f_3(p)| = |f_2(p)| \leqslant c$.

(5) By (3), if $p \in X \times Y_0$ then $|f_3(p)| \leqslant \max_x |f_3(x, y_0)| \leqslant \max_x |f_2(x, y_0)| \leqslant c$. ∎

LEMMA 3.4. *Let Q be a subset of $X \times Y$ such that*

$$[(x, y) \in Q, (u, y) \in Q, (u, v) \in Q] \Rightarrow (x, v) \in Q.$$

Then there exist pairwise disjoint families $X_\alpha \subset X$ and $Y_\alpha \subset Y$ such that $Q = \bigcup_\alpha (X_\alpha \times Y_\alpha)$.

Proof. Put $X_0 = \{x \in X: (x, y) \notin Q \text{ for all } y \in Y\}$,

$$Y_0 = \{y \in Y: (x, y) \notin Q \text{ for all } x \in X\}.$$

On $X \backslash X_0$ we introduce an equivalence relation $\sim$ by putting $x_1 \sim x_2$ if and only if there is a $y \in Y$ such that $(x_1, y) \in Q$ and $(x_2, y) \in Q$. Let $\{X_\alpha\}$ be the family of equivalence classes determined by this relation. Define

$$Y_\alpha = \{y \in Y: (x, y) \in Q \quad \text{for all} \quad x \in X_\alpha\}.$$

We prove that the Y_α are pairwise disjoint. Suppose that $y \in Y_\alpha \cap Y_\beta$. Then $(x, y) \in Q$ for all $x \in X_\alpha \cup X_\beta$. Take $x_\alpha \in X_\alpha$ and $x_\beta \in X_\beta$. Then $(x_\alpha, y) \in Q$ and $(x_\beta, y) \in Q$. Hence $x_\alpha \sim x_\beta$ and $X_\alpha = X_\beta$.

We prove that $Q \supset \bigcup_\alpha (X_\alpha \times Y_\alpha)$. Let $(\xi, \eta) \in X_\alpha \times Y_\alpha$. By the definition of Y_α, $(x, \eta) \in Q$ for all $x \in X_\alpha$. Hence in particular $(\xi, \eta) \in Q$.

We prove that $Q \subset \bigcup_\alpha (X_\alpha \times Y_\alpha)$. Let $(\xi, \eta) \in Q$. Then $\xi \in X \backslash X_0$ and $\xi \in X_\alpha$ for some α. Let x be an arbitrary element of X_α. Then $\xi \sim x$. Hence there is a $y \in Y$ such that $(x, y) \in Q$ and $(\xi, y) \in Q$. By the hypothesis on Q, $(x, \eta) \in Q$. This shows that $(x, \eta) \in Q$ for all $x \in X_\alpha$, and that $\eta \in Y_\alpha$. Hence $(\xi, \eta) \in X_\alpha \times Y_\alpha$. ∎

LEMMA 3.5. *Let X and Y be discrete spaces, and let S be a subset of $X \times Y$ such that when three vertices of a rectangle belong to S so does the fourth.*

Let $p \in (X \times Y)\backslash S$. Then there is a function $\phi(x, y) = g(x) + h(y)$ such that $\phi \mid S = 0$ and $\phi(p) = 1$.

Proof. By Lemma 3.4, there exist sets X_α and Y_α such that

$$X = X_0 \cup \bigcup_\alpha X_\alpha, \qquad Y = Y_0 \cup \bigcup_\alpha Y_\alpha,$$

$$S = \bigcup_\alpha (X_\alpha \times Y_\alpha).$$

Let $p = (x_0, y_0)$. If $x_0 \in X_0$, we let $g(x_0) = 1$ and let $g(x) = 0$ for all other $x \in X$. Let $h = 0$. Then this $g + h$ has the desired properties.

If $x_0 \notin X_0$, then $x_0 \in X_\beta$ for some β. Since $p \notin S$, $y_0 \notin Y_\beta$. Let $g \mid X_\beta = 1 = -h \mid Y_\beta$ and let g vanish on all the other component sets of X. Let h vanish on all the other component sets of Y. This $g + h$ has the desired properties. ∎

THEOREM 3.2. *If X and Y are finite sets then the set of null points of f is the smallest set S having these properties:*

(1) *S contains every extremal path for f*; i.e., $S \supset E(f)$.

(2) *If S contains three vertices of a rectangle, then it contains the fourth.*

Proof. By Lemma 3.3 and Theorem 3.1, there is no loss of generality in assuming that f is level, and $\operatorname{crit}(f) = E(f)$. We begin by showing that $N(f)$, the set of null points, has properties (1) and (2). Let $\phi \in M$ and $\|f + \phi\| = \|f\|$. By Lemma 3.1, ϕ vanishes on $E(f)$. Since this is true for every such ϕ, $E(f) \subset N(f)$. This is (1). Now suppose that $p_1, ..., p_4$ are the vertices of a rectangle and that p_1, p_2, p_3 belong to N. Then $\phi(p_1) = \phi(p_2) = \phi(p_3) = 0$. If l is the closed path made from $p_1, ..., p_4$ then $\pi_l(\phi) = 0$, and therefore $\phi(p_4) = 0$. Again, this is true for all such ϕ, and so $p_4 \in N(f)$. This is (2).

In the second half of the proof we show that N is contained in any set S having properties (1) and (2). Let $p \notin S$. By Lemma 3.5, there exists $\phi \in M$ such that $\phi \mid S = 0$ and $\phi(p) \neq 0$. Since $\operatorname{crit}(f) \subset S$, ϕ vanishes on $\operatorname{crit}(f)$. Hence an $\epsilon > 0$ exists for which $\|f + \epsilon\phi\| = \|f\|$. Since $\phi(p) \neq 0$, p is not a null point. Thus the complement of S is contained in the complement of N. ∎

THEOREM 3.3. *Let Q be a subset of $X \times Y$ such that when three vertices of a rectangle belong to Q so does the fourth. If the sets X_0 and Y_0 of Lemma 3.4 are finite, and if there are only finitely many equivalence classes X_α say k, then*

$$\dim\{\phi \in M: \phi \mid Q = 0\} = \#X_0 + \#Y_0 + k - 1.$$

Proof. Select k arbitrary numbers $\alpha_1, \ldots, \alpha_k$ and define $g(x) = -h(y) = \alpha_i$ if $(x, y) \in X_i \times Y_i$, $1 \leqslant i \leqslant k$. Let g and h be defined arbitrarily on X_0 and Y_0, respectively. Put $\phi(x, y) + h(y)$. Obviously, $\phi \in M$. If $(x, y) \in X_i \times Y_i$ then $\phi(x, y) = g(x) + h(y) = 0$. Hence $\phi \mid Q = 0$. Since ϕ is not changed if a constant is added to g and subtracted from h, the dimension of the set in question is at least $m \equiv \#X_0 + \#Y_0 + k - 1$.

On the other hand, if $\phi(x, y) = g(x) + h(y)$ and $\phi \mid Q = 0$, then $g(x) + h(y) = 0$ when $(x, y) \in X_i \times Y_i$. Hence $g(x)$ is a constant α_i on X_i and $h(y) = -\alpha_i$. Thus a set of m functions ϕ can be constructed, taking only values 0 and 1, to span $\{\phi: \phi \mid Q = 0\}$. Hence the dimension of this set is no greater than m. ∎

LEMMA 3.6. *Let f be a function on $X \times Y$, X and Y being finite sets. Let Q be the set of null points (Definition 3.2). Let $X_0 \cup X_1 \cup \cdots \cup X_k$ and $Y_0 \cup Y_1 \cup \cdots \cup Y_k$ be the decompositions of X and Y described in Lemma 3.4, relative to the set Q. Then*

$$\dim\{\phi \in M: \|f + \phi\| = \operatorname{dist}(f, M)\} = \#X_0 + \#Y_0 + k - 1.$$

Proof. By Lemma 3.3 and Theorem 3.1, there is no loss of generality in assuming that $\|f\| = \operatorname{dist}(f, M)$ and $\operatorname{crit}(f) = E(f)$. By Theorem 3.2, Q satisfies the hypotheses of Lemma 3.4. By the definition of Q, $\phi \mid Q = 0$ whenever $\|f + \phi\| = \|f\|$. Thus the set whose dimension we wish to compute is a subset of $\{\phi \in M: \phi \mid Q = 0\}$, and by Theorem 3.3 the former has dimension at most $m = \#X_0 + \#Y_0 + k - 1$.

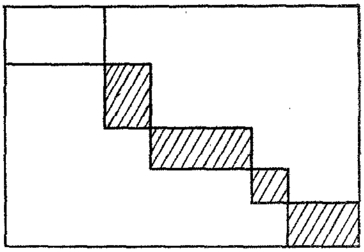
On the other hand, if $\phi \in M$ and $\phi \mid Q = 0$, then because the domain is discrete and $\operatorname{crit}(f) \subset Q$, we can determine an $\epsilon \neq 0$ for which $\|f + \epsilon\phi\| = \|f\|$. Thus the dimension to be computed is at least m. ∎

THEOREM 3.4. *Let X and Y be finite sets. A function f on $X \times Y$ has a unique best approximation in M if and only if every point of $X \times Y$ is a null point of f.*

Proof. By the preceding lemma, unicity of the best approximation occurs exactly when $\#X_0 + \#Y_0 + k - 1 = 0$. Since $k \geqslant 1$, $\#X_0 \geqslant 0$, and $\#Y_0 \geqslant 0$, the condition is that $k = 1$, $\#X_0 = 0$, and $\#Y_0 = 0$. Hence X_0 and Y_0 are void. ∎

The matrix formulation of Lemma 3.4 is as follows.

LEMMA 3.7. *Let A be a matrix of* 0's *and* 1's *such that whenever three* 0's *are vertices of a rectangular pattern, the fourth vertex is also a zero. Then by row and column permutations A can be put into the block form*



where the shaded area is occupied solely by 0's *and the unshaded area is occupied solely by* 1's.

Proof. By row and column interchanges, the rows and columns consisting entirely of 1's are moved to the top and left. Let these rows and columns be described by the inequalities $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$. In the next step permutations involving only indices $i > n$ and $j > m$ are performed to obtain a maximal rectangle of 0's in the upper left-hand position of the sub-matrix. Let this rectangle be described by $n < i \leqslant p$, $m < j \leqslant q$. If there is an element $a_{\nu\mu} = 0$ with $n < \nu \leqslant p$ and $\mu > q$, then by the hypothesis on A, we would have $a_{i\mu} = 0$ for $n < i \leqslant p$. An interchange of column μ with column $q + 1$ would enlarge the maximal rectangle of 0's. In this way we show that elements to the right of and below 0-blocks are all 1's. The proof is completed by repetition of this process. ∎

4. Special Problems on Finite Domains

Here a special case of the Diliberto–Straus algorithm is considered—one in which the generally nonlinear process becomes linear. The approximation problem on any Cartesian product of finite sets can always be solved by a finite sequence of the linear problems discussed here.

Assume that C and Y are two finite sets, each containing exactly n points. The domain, D, of our functions is a subset of $X \times Y$ having these properties:

(1) The points of D can be labelled in such a way that the result is a path: $D = [p_1, p_2, ..., p_{2n}]$ $(n = \#X = \#Y)$.

(2) Each horizontal or vertical line of $X \times Y$ contains exactly two points of D.

We can assume the labelling of D is chosen so that $p_i = (x_i, y_i)$, $x_{2i} = x_{2i+1}$, $y_{2i} = y_{2i-1}$, and $x_{2n} = x_1$.

Now let f be any function defined on D. We intend to subtract from f a function $\phi(x, y) = g(x) + h(y)$ so that $f - \phi$ is level on D. This will produce the best approximation ϕ at once. The "levelling equations" are (for $1 \leqslant i \leqslant n$)

$$\begin{aligned} f(p_{2i}) - g(p_{2i}) - h(p_{2i}) + f(p_{2i+1}) - g(p_{2i+1}) - h(p_{2i+1}) &= 0 \\ f(p_{2i-1}) - g(p_{2i-1}) - h(p_{2i-1}) + f(p_{2i}) - g(p_{2i}) - h(p_{2i}) &= 0 \end{aligned} \tag{3}$$

In these equations $P_{2n+1} = p_1$. With the abbreviations

$$\begin{aligned} u_i &= g(p_{2i}) = g(p_{2i+1}), \\ v_i &= h(p_{2i}) = h(p_{2i-1}), \\ 2b_i &= f(p_{2i+1}) + f(p_{2i}), \\ 2c_i &= f(p_{2i}) + f(p_{2i-1}) \end{aligned}$$

the system (3) can be written in the form

$$\begin{aligned} u_i + \tfrac{1}{2}(v_i + v_{i+1}) &= b_i , \\ v_i + \tfrac{1}{2}(u_{i-1} + u_i) &= c_i . \end{aligned}$$

In matrix form this is

$$\begin{bmatrix} I_n & B \\ B^T & I_n \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} b \\ c \end{bmatrix} \tag{4}$$

with $b_{ij} = \frac{1}{2}$ if $(0 \leqslant j - i \leqslant 1)$ or $(i = n$ and $j = 1)$, and $b_{ij} = 0$ otherwise.

Now it is necessary to prove that the following matrix is of rank $n - 1$.

$$A = \frac{1}{4}\begin{bmatrix} 2 & -1 & 0 & 0 & 0 & \cdots & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & & \vdots & \vdots \\ -1 & 0 & 0 & 0 & 0 & \cdots & -1 & 2 \end{bmatrix}.$$

Since any constant vector is in the null space of A, the rank of A is at most $n - 1$. We will show that the $(n-1) \times (n-1)$ matrix A^0 obtained from A by removing the last row and column is nonsingular. Suppose that $z = (z_1 ,..., z_{n-1})$ and that $zA^0 = 0$. This system of equations is

$$\begin{aligned} &2z_1 - z_2 = 0, \\ &-z_i + 2z_{i+1} - z_{i+2} = 0, \qquad 1 \leqslant i \leqslant n - 3, \\ &-z_{n-2} + 2z_{n-1} = 0, \end{aligned}$$

These equations can be employed one-by-one to express each z_i in terms of z_1, the result being $z_i = iz_1$ $(1 \leqslant i \leqslant n - 1)$. The last equation then states that $-(n-2)\, z_1 + 2(n-1)\, z_1 = 0$, whence $z_1 = 0$. Hence $z = 0$ and A^0 is nonsingular.

The equation

$$\begin{bmatrix} I & 0 \\ -B^T & I \end{bmatrix}\begin{bmatrix} I & B \\ B^T & I \end{bmatrix} = \begin{bmatrix} I & B \\ 0 & A \end{bmatrix}$$

implies that the second factor has rank $2n - 1$. The vector $\beta = [1,..., 1, -1,..., -1]$ is orthogonal to its columns. In order to prove that system (4) is consistent, it suffices to prove that β is orthogonal to $[b, c]^T$. The required calculation is

$$\sum b_i - \sum c_i = \tfrac{1}{2} \sum [f(p_{2i+1}) + f(p_{2i})] - \tfrac{1}{2} \sum [f(p_{2i}) + f(p_{2i-1})] = 0. \quad \blacksquare$$

THEOREM 4.1. *The best-approximation operator for the domain D is linear, and the linear system of equations* (4) *which determines the best approximation is consistent.*

Of course, the consistency of the system of equations also follows from an existence theorem for the ϕ such that $f - \phi$ is level.

The foregoing analysis can be used to compute the best approximation in one step on any domain D having the property that every horizontal or vertical line through a point of D contains exactly one other point of D. In this case, D can be decomposed into a disjoint union of paths on each of which $f - \phi$ can be made level by solving linear equations.

The formal way of decomposing D is to define an equivalence relation $p \sim q$ to mean that p can be joined to q by a polygonal line whose vertices are in D and whose line segments are alternately horizontal and vertical. The equivalence classes give the decomposition $D = \bigcup D_i$. This can be done whether X and Y are finite or infinite sets. If $M(D)$ denotes the space of functions $\phi(x, y) = g(x) + h(y)$ on D, we have

THEOREM 4.2. *Corresponding to the decomposition $D = \bigcup D_i$, $M(D)$ has the direct-sum decomposition*

$$M(D) = \sum \oplus M(D_i).$$

Here it is understood that if $\phi = \sum \phi_i$ with $\phi_i \in M(D_i)$ then $\phi(p) = \phi_i(p)$ if $p \in D_i$. Furthermore, $\| \phi \|_\infty = \sup_i \| \phi_i \|_\infty$.

Proof. If $\phi \in M$, let $\phi_i(p) = \phi(p)$ on D_i and $\phi_i(p) = 0$ elsewhere. Obviously $\phi = \sum \phi_i$. If $\phi_i \in M(D_j)$ then $\phi_i = \phi \mid D_j$ for some $\phi \in M$. Then $\phi_i \mid D_i = (\phi \mid D_j) \mid D_i = 0$ by disjointness of D_i and D_j. $\blacksquare$

Reference [6] came to our attention after we completed our work. It contains another proof of convergence of the Diliberto–Straus algorithm. References [1, 7] contain an abstract account of algorithms having the form of the Diliberto–Straus algorithm.

REFERENCES

1. B. ATLESTAN, F. E. SULLIVAN, Iteration with best approximation operators, *Rev. Roumaine Math. Pures Appl.* **21** (1976), 125–131. MR 53 # 6188.
2. G. AUMANN, Über approximative Nomographie, II, *Bayer. Akad. Wiss. Math.-Nat. Kl. S.-B.* (1959), 103–109. MR 22 # 6968.
3. S. P. DILIBERTO AND E. G. STRAUS, On the approximation of a function of several variables by the sum of functions of fewer variables, *Pacific J. Math.* **1** (1951), 195–210. MR 13, 334.
4. M. GOLOMB, Approximation by functions of fewer varibles, *in* "On Numerical Approximation" (R. Langer, Ed.), pp. 275–327, Univ. of Wisconsin Press, Madison, Wisconsin, 1959. MR 21, 962.
5. M. v. GOLITSCHEK AND E. W. CHENEY, On the algorithm of Diliberto and Straus for approximating bivariate functions by univariate ones, *Numer. Functional Anal. Optimization* **1** (1979), 341–363.
6. C. T. KELLEY, A note on the approximation of functions of several variables by sums of functions of one variable, *J. Approximation Theory*, in press.
7. F. SULLIVAN, A generalization of best approximation operators, *Ann. Mat. Pura Appl* **107** (1975), 245–261. MR 53 # 3564.

Additional Literature Citations to Mathematical Reviews

M. B. A. BABAEV, 43 # 6634, 35 # 7048, 35 # 7049, 45 # 2377, 45 # 4024, 43 # 2398, 36 # 4228, 26 # 6659.
R. C. BUCK, 51 # 13536, 52 # 8356, 46 # 7746.
A. S. CERETELI, 34 # 6401, 45 # 4029, 37 # 3255, 35 # 2040, 35 # 4655, 34 # 8046, 40 # 7679.
L. COLLATZ, 54 # 3240, 54 # 3254.
L. FLATTO, 34 # 1766.
S. JA. HAVINSON, 41 # 7351.
I. I. IBRAGIMOV AND M.-B. A. BABAEV, 43 # 5219, 45 # 7357.
V. M. MORDASEV, 38 # 3666, 39 # 4571.
V. P. MOTORNYI, 29 # 2582, 33 # 7763.
JU. P. OFMAN, 23 # A2684.
S. Z. RAFALSON, 45 # 2383.
T. J. RIVLIN AND R. J. SIBNER, 32 # 4450.
D. A. SPRECHER, 37 # 669, 34 # 8049, 41 # 2272.
A. I. VAINDINER, 41 # 8890, 48 # 12966.