A Construction of Nonnegative Approximate Quadratures*

By Philip J. Davis

1. Introduction. In a paper which appeared in 1957, V. Tchakaloff [1] proved the following theorem. *Let B be a closed bounded set in the plane with positive area. Let $\phi_1, \phi_2, \cdots, \phi_N$ be N linearly independent and continuous functions of x, y in B, of which one does not vanish in B. Then we can find N points $P_i:(x_i, y_i)$ lying in B and N weights $w_i \geqq 0$ such that*

$$(1.1) \qquad \iint_B \phi_j dxdy = \sum_{i=1}^{N} w_i \phi_j(P_i) , \qquad j = 1, 2, \cdots, N .$$

Tchakaloff's demonstration is a very beautiful one, involving the theory of convex bodies. A separating hyperplane is employed and a nonconstructive proof is obtained. The theorem is valid for weighted integrals of dimension $d \geqq 1$.

Equivalent results on finite moment spaces were obtained earlier by various authors. See, e.g., Karlin and Studden [2, Chapter II]. Tchakaloff's independent work appears to be the first to formulate the result explicitly as in (1.1), thereby stressing its numerical analysis aspect.

This result is interesting for numerical analysis because: (1) Quadrature rules with nonnegative weights are more favorable than rules with mixed weights in that they lead to more stable computations; (2) Interpolating quadrature formulas determined by brute force methods do not often yield weights that are of one sign.

The purpose of the present paper is to give an alternative proof of Tchakaloff's theorem which is constructive in its nature. The present proof is also a more "elementary" one than Tchakaloff's in that it makes use only of the familiar raw materials of elementary numerical analysis.

Extensions and numerical applications will be published subsequently by the author and by M. W. Wilson.

2. An Alternate Proof of Tchakaloff's Theorem. In this proof we limit ourselves to integrals of dimension $d = 2$ and to functions $\phi_1, \phi_2, \cdots, \phi_N$ that are monomials (i.e., powers) in x, y. This limitation will still enable us to exhibit the essential features of the method.

We begin with a number of very simple lemmas.

Lemma 1. *Let $\phi_1(x, y) = 1$, $\phi_2(x, y) = x$, $\phi_3(x, y) = y$, $\phi_4(x, y) = x^2$, $\phi_5(x, y) = xy$, $\phi_6(x, y) = y^2$, $\cdots$ be an arrangement of the powers $x^i y^j$, $0 \leqq i, j < \infty$. For any integer $N \geqq 1$, the functions $\phi_1, \cdots, \phi_N$ are linearly independent. That is, if $f(x, y) = \sum_{i=1}^{N} a_i \phi_i(x, y) \equiv 0$ in a region R, then $a_i = 0$, $i = 1, 2, \cdots, N$.*

Proof. Call $m + n$ the degree of the monomial $x^m y^n$. We have $\partial^{m+n} x^m y^n / \partial x^m \partial y^n = m!n!$, and $\partial^{m+n} x^{m'} y^{n'} / \partial x^m \partial y^n = 0$ if $m' + n' = m + n$ but $(m', n') \neq (m, n)$, or if

Received January 3, 1967.

* The results contained in this paper were obtained in the course of research sponsored by the National Science Foundation under Grant GP-4213.

$m' + n' < m + n$. Assume that $f \equiv 0$ in R. Now let $a_j x^m y^n$ be a monomial of highest degree in f. Then, $\partial^{m+n} f / \partial x^m \partial y^n \equiv 0 = a_j m! n!$. Hence $a_j = 0$. Now iterate this process and conclude that all the coefficients vanish.

COROLLARY. *If $\mathcal{P}_N$ designates the linear space of functions $\sum_{i=1}^N a_i \phi_i$, then $\mathcal{P}_N$ is of dimension N.*

LEMMA 2. *Let B be a region in the x, y plane. Then, we can find points $P_1 = (x_1, y_1), \cdots, P_N = (x_N, y_N)$ in B such that*

$$\begin{vmatrix} \phi_1(P_1) & \phi_2(P_1) & \cdots & \phi_N(P_1) \\ \vdots & & & \vdots \\ \phi_1(P_N) & \phi_2(P_N) & \cdots & \phi_N(P_N) \end{vmatrix} \neq 0 \,. \tag{2.1}$$

Proof. Select any point in B as P_1. Then $\phi_1(P_1) \neq 0$. Consider the function

$$\begin{vmatrix} \phi_1(P_1) & \phi_2(P_1) \\ \phi_1(P) & \phi_2(P) \end{vmatrix} = g(P) \,.$$

This is a linear combination of $\phi_1(P)$ and $\phi_2(P)$. If $g(P) \equiv 0$ in B, it would follow from Lemma 1 that $\phi_1(P_1) = 0$ and $\phi_2(P_1) = 0$. This is impossible. Hence there is a P_2 such that $g(P_2) \neq 0$. Consider next the function

$$\begin{vmatrix} \phi_1(P_1) & \phi_2(P_1) & \phi_3(P_1) \\ \phi_1(P_2) & \phi_2(P_2) & \phi_3(P_2) \\ \phi_1(P) & \phi_2(P) & \phi_3(P) \end{vmatrix} = h(P) \,.$$

This is a linear combination of $\phi_1(P)$, $\phi_2(P)$, $\phi_3(P)$. If $h(P) \equiv 0$, all coefficients would be zero. But the coefficient of $\phi_3(P)$ is $g(P_2) \neq 0$. In this way we may proceed step by step.

It should be observed that if $Q_1, \cdots, Q_N$ are N distinct points in B, it does not necessarily follow (as in the case of polynomials of one variable) that $|\phi_i(Q_j)| \neq 0$. However, the following may be asserted.

COROLLARY. *Given N points $Q_1, Q_2, \cdots, Q_N$ in B, and given $\epsilon > 0$. Then we can find N points $P_1, P_2, \cdots, P_N$ such that $|Q_i - P_i| \leqq \epsilon$, $i = 1, 2, \cdots, N$ and $|\phi_i(P_j)| \neq 0$.*

Proof. Select $P_1 = Q_1$. The above argument for $g(P)$ yields a point P_2 in any neighborhood of Q_2 such that $g(P_2) \neq 0$. We may now proceed step by step.

LEMMA 3. *Given a rectangle R: $x_1 \leqq x \leqq x_2$, $y_1 \leqq y \leqq y_2$ and a fixed integer $N \geqq 1$. We can find an integer $k(N)$ and $k(N)$ points $P_1, P_2, \cdots, P_{k(N)}$, and $k(N)$ weights $w_1 > 0, w_2 > 0, \cdots, w_{k(N)} > 0$ such that*

$$\iint_R \phi_j dxdy = \sum_{i=1}^{k(N)} w_i \phi_j(P_i) \,, \qquad j = 1, 2, \cdots, N \,. \tag{2.2}$$

Proof. This can be accomplished in many ways. For example, one can use a product rule of Gauss rules of sufficiently high order. To be more specific, let the highest power of x and y in $\phi_1, \phi_2, \cdots, \phi_N$ be respectively $p(N)$ and $q(N)$. Determine $p^*(N)$ and $q^*(N)$ such that $2p^*(N) - 1 \geqq p(N)$ and $2q^*(N) - 1 \geqq q(N)$. Set $k(N) = p^*(N) q^*(N)$ and form the product rule of Gauss rules of order $p^*(N)$ in x and $q^*(N)$ in y. This product rule will integrate over R exactly all monomials $x^i y^j$, $0 \leqq i \leqq 2p^* - 1$, $0 \leqq j \leqq 2q^* - 1$, and a fortiori $\phi_1, \phi_2, \cdots, \phi_N$. The points

$P_1, P_2, \cdots$ are (x_α, y_β) where $\{x_\alpha\}$ and $\{y_\beta\}$ are the Gaussian abscissas along the x and y axes respectively. The weights are products of Gaussian weights and hence are positive.

COROLLARY. *By taking Gauss rules of odd order, one of the points P_i will be the center of the rectangle.*

LEMMA 4. *Given a region B and N distinct points $P_1, P_2, \cdots, P_N$ lying in the interior of B. Then we can find N squares S_i: $x_{1i} \leqq x \leqq x_{2i}$, $y_{1i} \leqq y \leqq y_{2i}$, $i = 1, 2, \cdots, N$, sufficiently small and placed in such a manner that*

(a) $S_i \subset B$,

(b) $S_i \cap S_j = 0$ *if* $i \neq j$.

Proof. Take, e.g., P_i as the center of the squares and take the diameter of the squares less than $\frac{1}{2} \min_{1 \leqq i,j \leqq N} |P_i - P_j|$.

LEMMA 5. *Given a bounded region B and N distinct points $P_1, P_2, \cdots, P_N$ in B. Given a $\delta > 0$. Then we can find an integer s ($s = s(B; P_1, \cdots, P_N; \delta)$) and s rectangles $R_1, R_2, \cdots, R_s$ with sides parallel to the x and y axes such that*

(a) *The R_i include the squares S_i already constructed in Lemma* 4,

(b) $R_i \subset B$,

(c) $R_i \cap R_j = 0$ *if* $i \neq j$,

(d) area $B - \sum_{i=1}^{s}$ area $R_i \leqq \delta$.

Proof. Take $R_1 = S_1, \cdots, R_N = S_N$. For the remaining rectangles, pack $B - \bigcup_{i=1}^{N} S_i$ with nonoverlapping rectangles sufficiently densely so that the area of B is approximated by $\sum_1^s$ area R_i to within δ. The exact details here do not have to be spelled out.

Remark. If N is held fixed but if $\delta \to 0$, note that the first N rectangles, the first of which contains P_1, the second P_2, etc., may be kept fixed.

THEOREM. *Let B be a bounded region in the $x-y$ plane. Let $N \geqq 1$ be fixed. Then we can find points $T_1, T_2, \cdots, T_N$ in B and nonnegative weights $w_1, w_2, \cdots, w_N$ such that*

$$\iint_B \phi_j dxdy = \sum_{i=1}^{N} w_i \phi_j(T_i) , \qquad j = 1, 2, \cdots, N . \tag{2.3}$$

Proof. I. Select N points $P_1, \cdots, P_N$ in B such that $|\phi_i(P_j)| \neq 0$. This is possible by Lemma 2. Pack B with rectangles $R_1, R_2, \cdots, R_s$ as in Lemma 5. The relevant δ will be specified shortly. The first N rectangles will be squares $S_1, \cdots, S_N$. Over each rectangle R_i, define a positive quadrature rule as in Lemma 3. In $S_1, \cdots, S_N$ make sure that one of the nodal points in the respective squares is the center of the square, i.e., P_i, where $i = 1, 2, \cdots, N$.

Let u_i, U_i ($i = 1, 2, \cdots, n$) designate the weights and respective locations that occur in all the rules defined over all the rectangles. Note that $u_i > 0$. Note further that we may arrange the order so that $U_1 = P_1, U_2 = P_2, \cdots, U_N = P_N$.

For any matrix $A = (a_{ij})$, let $||A||$ designate the matrix norm $\max_i \sum_j |a_{ij}|$ and for a vector $v = (v_1, \cdots)$ let $||v||$ designate the compatible vector norm $\max_i |v_i|$. Let $M = \max_{1 \leqq j \leqq N} \sup_{P \epsilon B} |\phi_j(P)|$. Now select δ such that

$$0 < \delta < \min_{1 \leqq i \leqq N} \frac{u_i}{M||(\phi_j(P_i))^{-1}} . \tag{2.4}$$

Use such a value of δ in packing B with rectangles. (Note the order of procedure here. N is given. Determine $P_1, \cdots, P_N$. Put squares S_i around P_i, $i = 1, 2, \cdots, N$, and in each square define a product Gauss rule of which one node is P_i and the corresponding weight is u_i. Next determine δ from (2.4) and use it to form a packing of B by rectangles into which we also insert a positive quadrature.)

If $R = \bigcup_{i=1}^{s} R_i$, we have

$$\iint_R \phi_j dxdy = \sum_{i=1}^{n} u_i \phi_j(U_i) , \qquad j = 1, 2, \cdots, N . \tag{2.5}$$

Let

$$\epsilon_j = \iint_B \phi_j dxdy - \iint_R \phi_j dxdy , \qquad j = 1, 2, \cdots, N . \tag{2.6}$$

Then,

$$|\epsilon_j| \leqq \iint_{B-R} |\phi_j| dxdy \leqq M \iint_{B-R} dxdy \leqq M\delta , \qquad j = 1, 2, \cdots, N . \tag{2.7}$$

Now (2.5) can be rewritten as

$$\iint_B \phi_j dxdy - \epsilon_j = \sum_{i=1}^{N} u_i \phi_j(P_i) + \sum_{i=N+1}^{n} u_i \phi_j(U_i) . \tag{2.8}$$

Now consider the $N \times N$ system in variables $t = (t_1, \cdots, t_N)$

$$\sum_{i=1}^{N} t_i \phi_j(P_i) = \epsilon_j , \qquad j = 1, 2, \cdots, N . \tag{2.9}$$

If $\epsilon = (\epsilon_1, \epsilon_2, \cdots, \epsilon_N)'$, (2.9) has the solution

$$t = [\phi_j(P_i)]^{-1} \epsilon \tag{2.10}$$

and hence by (2.7) and (2.4),

$$\begin{aligned} \|t\| &\leqq \|[\phi_j(P_i)]^{-1}\| \|\epsilon\| \\ &\leqq \|[\phi_j(P_i)]^{-1}\| M\delta \\ &< \min_{1 \leqq i \leqq N} u_i . \end{aligned} \tag{2.11}$$

Hence,

$$\max_{1 \leqq i \leqq N} |t_i| < \min_{1 < i < N} u_i . \tag{2.12}$$

Combining (2.9) and (2.8) we obtain

$$\begin{aligned} \iint_B \phi_j dxdy &= \sum_{i=1}^{N} (u_i + t_i) \phi_j(P_i) + \sum_{N+1}^{n} u_i \phi_j(U_i) \\ &= \sum_{i=1}^{n} u_i' \phi_j(U_i) \qquad j = 1, 2, \cdots, N \end{aligned} \tag{2.13}$$

where, in view of (2.12), $u_i' > 0$.

II. The object of part I was to produce a quadrature formula (2.13) with positive weights. The abscissas or nodes are $U_1, U_2, \cdots, U_n$, where n may be very much larger than N. We shall next show that we may reduce n to N by using an appropriate subset of $\{U_i\}$. This can be done by a method of E. Steinitz [3].

The linear space $\mathcal{P}_N$ of functions $\sum_{j=1}^{N} a_i\phi_i$ defined on a region R is of dimension N. Hence, the algebraic dual space (the space of all linear functionals defined on $\mathcal{P}_N$) is also of dimension N. Among the n linear functionals

$$L_i(f) = f(U_i)\,, \tag{2.14}$$

at most N can be linearly independent. Hence if $n > N$, we must have

$$a_1L_1 + \cdots + a_nL_n = 0 \tag{2.15}$$

where not all the a's vanish and, in fact, one of the a's may be assumed to be positive. Define

$$L = u_1'L_1 + u_2'L_2 + \cdots + u_n'L_n\,, \qquad (u_i' > 0) \tag{2.16}$$

where the u_i' are from (2.13), and set

$$\sigma = \max_{1 \leqq i \leqq n} \frac{a_i}{u_i'}\,. \tag{2.17}$$

Note that $\sigma > 0$, $\sigma u_i' - a_i \geqq 0$, and furthermore, $\sigma u_i' - a_i = 0$ for at least one i. From (2.16) and (2.15) we obtain

$$L = \frac{\sigma u_1' - a_1}{\sigma} L_1 + \frac{\sigma u_2' - a_2}{\sigma} L_2 + \cdots + \frac{\sigma u_n' - a_n}{\sigma} L_n\,. \tag{2.18}$$

Thus, L has been expressed as a linear combination of at most $n - 1$ of $L_1, \cdots, L_n$ with nonnegative coefficients. Iterating this process, we arrive at

$$L = w_1L_1 + w_2L_2 + \cdots + w_NL_N\,, \qquad w_i \geqq 0\,. \tag{2.19}$$

Hence, from (2.13)

$$\iint_B \phi_j dxdy = \sum_{i=1}^{N} w_i\phi_j(U_i')\,, \qquad w_i \geqq 0\,, \qquad j = 1, 2, \cdots, N\,, \tag{2.20}$$

where the U_i' $(i = 1, 2, \cdots, N)$ are a subset of U_i $(i = 1, 2, \cdots, n)$.

Brown University
Mathematics Department
Providence, Rhode Island

1. V. Tchakaloff, "Formules de cubatures mécaniques à coefficients non negatifs," *Bull. Sci. Math.*, v. 81, 1957, pp. 123–134. MR **20** #1145.

2. S. J. Karlin & W. J. Studden, *Tchebycheff Systems*, Interscience, New York, 1966.

3. E. Steinitz, "Über bedingt konvergente Reihen und konvexe Systeme," *J. Reine Angew. Math.*, v. 143, 1913, pp. 128–175.