JOURNAL OF APPROXIMATION THEORY 6, 283–290 (1972)

Birkhoff Interpolation and the Problem of Free Matrices*

G. G. LORENTZ

Department of Mathematics, University of Texas, Austin, Texas 78712

Communicated by Oved Shisha

DEDICATED TO PROFESSOR J. L. WALSH ON THE OCCASION OF HIS 75TH BIRTHDAY

1. INTRODUCTION

An *incidence matrix* for the polynomials of degree n is an $m \times (n+1)$ matrix

$$E = (\epsilon_{kl}), \qquad k = 1,..., m, \qquad l = 0,..., n \tag{1.1}$$

with elements ϵ_{kl} that take values 0 and 1. A *scheme* S is the set consisting of an incidence matrix E and of m points $a \leqslant x_1 < x_2 < \cdots < x_m \leqslant b$; a *Birkhoff interpolation problem* is the problem of finding a polynomial P of degree n that satisfies, for the given data b_{kl}, the condition

$$P^{(l)}(x_k) = b_{kl}, \; (k, l) \in e \tag{1.2}$$

(e is the set of pairs (k, l) for which $\epsilon_{kl} = 1$). (Named after G. D. Birkhoff, who submitted the paper [2] to the American Mathematical Society at the age of 20).

Schoenberg [5] proposed the problem to describe all *free* (or *poised*) *matrices* E, for which the problem (1.2) has a solution for each choice of the x_k and the b_{kl}. We can assume that the set e has $|e| \leqslant n+1$ elements; if $|e| = n+1$, the problem always has a solution if and only if each polynomial P of degree n that vanishes on the scheme S [that is, satisfies the homogeneous Eq. (1.2)] is identically zero.

Let M_l denote the number of 1's in the rows $j = 0,..., l$ of E. Of importance are the following conditions:

$$M_l \geqslant l + 1, \; l = 0, 1,..., n \tag{1.3}$$

* Supported by contract no. AFOSR-69-1826 of the office of Scientific Research, U. S. Air Force. Submitted to the American Math. Society, Nov. 3, 1969.

(the Pólya condition) and

$$M_l \geqslant l + 2, \; l = 0, 1, \ldots, n-1 \tag{1.4}$$

(the strong Pólya condition). Each free matrix satisfies (1.3).

A *supported sequence* of E is a maximal sequence of 1's in a row of E,

$$\epsilon_{i_0 j_0} = \cdots = \epsilon_{i_0 J} = 1,$$

which is supported: there exist (i_1, j_1), (i_2, j_2) for which $i_1 < i_0 < i_2$, $j_1, j_2 < j_0$ and $\epsilon_{i_1 j_1} = \epsilon_{i_2 j_2} = 1$. Atkinson and Sharma [1] (see also [4]) proved that E is free if it satisfies (1.3) and if each of its supported sequences is even (that is, it has an even number of elements). They proposed the conjecture that if E satisfies (1.4), their condition is also necessary for E to be free. This proved to be incorrect [4].

In this note we describe a wide class of nonfree matrices E. Although technically more difficult, the proof of our main result is based on ideas that appear in Theorem 2 of the paper [4].

2. Remarks about Identities

We shall relate our problem to the existence of certain identities for polynomials P of degree n. There does not seem to exist a theory of such identities. They have been of importance also for the problem of monotone approximation [3].

Proposition. *A scheme given by the points $x_0 < \cdots < x_m$ and an incidence matrix E is not free if and only if there exists a nontrivial identity*

$$\sum_{(i,j) \in e} a_{ij} P^{(j)}(x_i) = 0, \tag{2.1}$$

valid for all polynomials P of degree n.

Proof. We consider the $n + 1$-dimensional space R^{n+1} with points $\xi = (\xi_0, \ldots, \xi_n)$; in particular, let

$$\xi_{ij} = \{n \cdots (n-j+1)x_i^{n-j}, (n-1) \cdots (n-j)x_i^{n-j-1}, \ldots, j!, 0, \ldots, 0\}, \quad (i,j) \in e. \tag{2.2}$$

The scheme is not free precisely when the points (2.2) are linearly dependent; this is equivalent to the existence of constants a_{ij}, not all zero, with the property that in R^{n+1},

$$\sum a_{ij} \xi_{ij} = 0. \tag{2.3}$$

Applying here any functional $L(\xi) = a_0\xi_0 + \cdots + a_n\xi_n$, and noticing that $L(\xi_{ij}) = P^{(j)}(x_i)$ for the corresponding $P(x) = a_0x^n + \cdots + a_n$, we see that (2.1) is equivalent to (2.3) for all P.

EXAMPLE. For P of degree 2 one shows that an identity of type (2.1), which contains a value of P itself, must be of the form

$$P'\left(\frac{a+b}{2}\right)(b-a) = P(b) - P(a). \tag{2.4}$$

It follows that the only nonfree matrix for polynomials of degree 2 that satisfies (1.3) is the matrix

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Several "strange" identities of type (2.1) have been constructed in [3, Section 5]. In particular,

1. For n even, k odd and $1 \leqslant k \leqslant n - 1$, there exists an identity (2.1) for polynomials of degree n that contains $n - 2$ values of P and 2 values of $P^{(k)}$. Here the total number of nonzero terms in (2.1) is n [3, Theorem 14].

2. If n is odd, there exists an identity (2.1) containing $(n + 3)/2$ values of P, $(n - 1)/2$ values of P' [3, Theorem 15].

3. If $3 \leqslant m \leqslant n + 2$ is of the same parity with n, there exists an identity (2.1) with altogether $(n + m)/2 + 1$ terms, m of them values of P, and $(n - m)/2 + 1$ first derivatives [3, Theorem 16]. In particular, if n is odd, there could be 3 values of P_n, and $(n - 1)/2$ values of P', altogether only $(n + 5)/2$ values.

3. THE MAIN RESULT

THEOREM. *Let E be an incidence matrix which satisfies* (1.3) *and has a row with exactly one supported odd sequence. Then E is not free.*

We shall use the following known facts about polynomials:

LEMMA 1 (Rolle's theorem). *If $\alpha < \beta$ are two consecutive real roots of a polynomial P, then the number of the roots of the derivative P' in (α, β) is odd.*

LEMMA 2 [4, 6]. *If $d = d_n = (4n^2)^{-1}$, $n > 0$, and if $\alpha < \beta$ are two roots of a polynomial P of degree n, then $P'(\xi) = 0$ for some $\xi = [\alpha + 2dl, \beta - 2dl]$, $l = \beta - \alpha$.*

Let $p_j \leqslant k \leqslant q_j$, $j = 1, 2,...$, be the locations of the supported sequences of the i-th row of E, and let $j = j_0$ correspond to the odd sequence. We write $p = p_{j_0}$, $q = q_{j_0}$.

We define a very special scheme S for the matrix E. We consider the points of $(-1, 1)$,

$$y_{-i} = -1 + N^{-i}, \quad y_i = 1 - N^{-i}, \quad i = 1, 2,... \quad N = d^{-n-1}. \tag{3.1}$$

We define S by assigning to the $i-1$-st row of E the point y_{-2}, to the $i+1$-st row $- y_2$, to the $i-2$-nd row the point y_{-3}, and so on. To the i-th row we assign the variable point λ, $y_{-1} \leqslant \lambda \leqslant y_1$.

Let E' be the matrix obtained from E by replacing the value $\epsilon_{iq} = 1$ by 0, and S' the corresponding scheme. Since each system of n homogeneous linear equations with $n+1$ unknowns has a nontrivial solution, there exist polynomials P of degree n, that vanish on S' without vanishing identically.

We fix one of these polynomials P, and study its "Rolle zeros". These are the zeros of P and of its derivatives which are specified by S', and also those that can be derived from them by the use of Rolle's theorem.

More precisely, the *Rolle zeros* of P are defined inductively in k for each derivative $P^{(k)}$, $0 \leqslant k \leqslant n$. The Rolle zeros of $P^{(0)}$ are the zeros of P given by the scheme S'. Let the Rolle zeros of $P^{(k-1)}$ be known. We define those of $P^{(k)}$ (and some zeros of the higher derivatives) in the following way. Let α, β be two consecutive Rolle zeros of $P^{(k-1)}$. It may happen that (α, β) contains an *even* number (counting their multiplicity) of zeros of $P^{(k)}$, specified by S'. Then, by Lemma 1, (α, β) contains an additional zero of $P^{(k)}$. If it is possible, we select this zero ξ to be different from all zeros previously known. If it is impossible (case of *degeneracy*), then there must be a multiple root ξ, specified by S', and this root must have a multiplicity at least one unit larger than specified. In this case ξ is added as a root of a corresponding higher derivative of P. In all cases, we say that ξ has been obtained by combining α and β.

The Rolle zeros of $P^{(k)}$ are all ξ obtained in this way together with the zeros of $P^{(k)}$ specified by S'. (There may be several possible choices of Rolle zeros). On (y_{-2}, y_2), degeneracy can occur only if $\alpha < \lambda < \beta$ and only if $k = p_j$ for one of the supported sequences. In case of degeneracy λ will be the Rolle zeros of $P^{(q_j+1)}$, $j \neq j_0$, and if $j = j_0$, of $P^{(q)}$. If this happens, we shall say that there is a *loss of a zero* for the k-th derivative $k = p_j$, and a *gain of a zero* for $k = q_{j+1}$ (or $k = q$).

The points (3.1) have been so selected that *there is no degeneracy for* $\xi < y_{-2}$ *or* $\xi \geqslant y_2$. This follows from Lemma 2 (see [4] or Lemma 3 below).

We count the number of Rolle zeros of $P^{(k)}$. For $k = 0$, P_n has exactly $m_0 = M_0 \geqslant 2$ zeros. Let $k < q$. By induction in k we see that $P^{(k)}$ has

$M_k - k \geqslant 2$ zeros, unless there is a loss. A loss can happen only at $\xi = \lambda$, and then λ will be a zero until the next gain. After the gain, there will be again $M_k - k \geqslant 2$ Rolle zeros. This will continue until $k = q$. From here on, we have to replace M_k by $M_k - 1$, since the 1 at the place (i, q) has been replaced by 0 in the matrix E'. Thus, by (1.4), the number of Rolle zeros of $P^{(k)}$ will be $\geqslant 1$, $k \leqslant n - 1$. This will be even true in case of a loss, for then, until the next gain, λ will provide a known zero. We have shown: *The number of Rolle zeros of $P^{(k)}$ in $(-1, 1)$ is independent of the position of λ in $[y_{-1}, y_1]$, except for*

$$p_j \leqslant k \leqslant q_j, j \neq j_0 \text{ or } p \leqslant k < q. \tag{3.2}$$

For $0 \leqslant k \leqslant n - 1$, $P^{(k)}$ has either at least two zeros, or at least the zero λ.

Assume now that we somehow have found an additional zero of $P^{(k)}$ in (y_{-2}, y_2), or have proved that one of the Rolle zeros (other than λ) in this interval is a double zero. Then the above count gives one additional zero for each derivative, now even for $P^{(n)}$. Then P must identically vanish, a contradiction. Thus, *all Rolle zeros in (y_{-2}, y_2), other than λ, are simple, and there are no other zeros in this interval.*

It also follows that each polynomial of degree $n - 1$ that vanishes on S', is identically zero. Hence our P are of degree exactly n. We norm P by making the highest coefficient to be 1, and obtain then, for each λ, $y_{-1} \leqslant \lambda \leqslant y_1$, a unique polynomial $P(x, \lambda) = x^n + a_1(\lambda)x^{n-1} + \cdots + a_n(\lambda)$, vanishing on S'. The coefficients a_i are obtainable from n linear equations with n unknown $a_1, \ldots, a_n$, which have a unique solution for each λ. Hence the determinant of the system is not zero, $y_{-1} \leqslant \lambda \leqslant y_1$, so that $a_1, \ldots, a_n$ are *continuous functions of* λ. Our intention is to study the roots of $P^{(q)}(x; \lambda)$; under certain conditions they also will be continuous functions of λ, and choosing λ properly, we can make one of them equal to λ. We then will have a nonzero polynomial P vanishing on S, and thus prove our theorem.

About the distribution of Rolle zeros of $P(x, \lambda)$ for different λ, $y_{-1} \leqslant \lambda \leqslant y_1$, we have the following:

LEMMA 3. *For each n, there exists a number δ, $0 < \delta < y_2 - y_1$ with the following properties*:

(i) *All Rolle zeros of P in (y_{-2}, y_2) actually lie in $J_1 = (y_{-1} - \delta, y_1 + \delta)$; all Rolle zeros that can be derived from Rolle zeros in J_1 lie again in J_1.*

(ii) *Let $\lambda = y_{-1}$ or $\lambda = y_1$. Then all Rolle zeros of the interval (y_{-1}, y_1) lie actually in $J_2 = (y_{-1} + \delta, y_1 - \delta)$; all Rolle zeros that can be derived from these lie again in J_2; each Rolle zero obtained from two zeros, one negative and the other positive, lies in J_2.*

(iii) *Let* $\lambda = y_1$. *Then all Rolle zeros of the interval* (y_1, y_2) *are actually in* $J_3 = (y_1, y_1 + \delta)$. *A Rolle zero derived from a zero to the left of* y_1 *and one to the right of* y_1 *lies in* J_2.

(iv) *There is no degeneracy for the Rolle zeros in* $[-1, y_{-2}]$ *and* $[y_2, 1]$; *Rolle zeros derived from a zero in one of these intervals, and a zero outside, lie in* J_1.

Proof. We prove only (i), the other proofs being similar (compare also [4]). The proof is by induction in k. For $k = 0$, there are no Rolle zeros of P_n in (y_{-2}, y_2), except perhaps λ. A Rolle zero ξ of P_n' could come form a zero $\alpha \leqslant y_{-2}$ and a zero $\beta \geqslant y_{-1}$, or from a zero $\alpha \leqslant y_1$ and a zero $\beta \geqslant y_2$. For all these positions, $\alpha = -1$, $\beta = y_{-1}$ and $\alpha = y_1$, $\beta = 1$ and the intervals of Lemma 2 provide a lower and an upper bound for ξ. It follows that $\xi \in [-1 + 2d(1 - y_1), 1 - 2d(1 - y_1)]$. Similarly, Rolle zeros of P_n'', derivable from these ξ, or by combining a point $\leqslant y_{-2}$ with a point $\geqslant y_2$, lie in the interval $[-1 + 4d^2(1 - y_1), 1 - 4d^2(1 - y_1)]$, and so on. For all k, Rolle zeros of (y_{-2}, y_2) lie in $[-1 + 2^n d^n(1 - y_1), 1 - 2^n d^n(1 - y_1)]$. Thus, for the purpose of (i), one can take $\delta = d^{n+1}(1 - 2^n d^n) < y_2 - y_1$.

We shall count certain categories of Rolle zeros in $[-1, +1]$, especially those of $P^{(q)}$. We shall show that the number of some of them is independent of the position of λ in $[y_{-1}, y_1]$, and that the number of others, in the contrary, changes with λ. The desired information can be derived from Lemma 3. We introduce the following notations. Let m_k', l_k, m_k'' denote the number of 1's in the k-th column of E, and, respectively, the first $i - 1$, the i-th, or the last $n - i$ rows. Let X_k', R_k, X_k'', $R_k(\lambda)$ denote the number of Rolle zeros of $P^{(k)}$, respectively, in the intervals $[-1, y_{-2}]$, (y_{-2}, y_2), $[y_2, 1]$, (λ, y_2); in particular, let $R_q = r$.

We can show that R_k is independent of λ, if k does not belong to the intervals (3.2). This has been shown above for the total number of Rolle zeros of $P^{(k)}$. It is sufficient to add that X_k', X_k'' are independent of λ. For $k = 0$ this is clear immediately. For the general case it follows from Lemma 3(iv) that

$$X_k' = (X_{k-1}' - 1)_+ + m_k', \quad X_k'' = (X_{k-1}'' - 1)_+ + m_k'' \tag{3.3}$$

and our statement follows by induction.

All zeros, except λ, of $P^{(q)}$ in (y_{-2}, y_2) are simple. This is true also of λ itself (λ can become a Rolle zero of $P^{(q)}$ only as a zero gain), because S' does not specify λ as a zero of $P^{(q+1)}$. For $P^{(q)}$ we can use the following known (compare [3]) lemma:

LEMMA 4. *Let* $Q(x, \lambda)$ *be a polynomial in* x, *that depends continuously on the real argument* λ. *If all zeros of* Q *in* $[a, b]$ *lie in* $[a + \delta, b - \delta]$, $\delta > 0$, *and are simple, then their number r does not depend upon λ. Enumerated in order of their magnitude*

$$x_1(\lambda) < \cdots < x_r(\lambda), \tag{3.4}$$

they are continuous functions of λ.

We shall now prove that

$$R_k(y_1) < R_k(y_{-1}), \qquad p \leqslant k \leqslant q. \tag{3.5}$$

First let $y = y_i$. By Lemma 3(iii), Rolle zeros of $P^{(k)}$ in (y_1, y_2) are obtainable only by combining Rolle zeros of $P^{(k-1)}$ in this interval and λ with each other or with zeros of $[y_2, 1]$. Therefore

$$R_k(y_1) = (l_{k-1} + R_{k-1}(y_1) + \epsilon''_{k-1} - 1)_+, \tag{3.6}$$

where $\epsilon''_{k-1} = 1$ if $X''_{k-1} > 0$, $\epsilon''_{k-1} = 0$ otherwise.

For $\lambda = y_{-1}$, Rolle zeros of $P^{(k)}$ in (λ, y_2) are described by Lemma 3(ii), (iii). This time we have

$$R_k(y_{-1}) = (\epsilon'_{k-1} + R_{k-1}(y_{-1}) + \epsilon''_{k-1} - 1)_+, \tag{3.7}$$

where $\epsilon'_{k-1} = 1$ if there are Rolle zeros of $P^{(k-1)}$ in $[-1, y_{-1}]$, and $= 0$ otherwise. Thus $\epsilon'_{k-1} \geqslant l_{k-1}$, and from (3.6) and (3.7) we derive $R_k(y_{-1}) \geqslant R_k(y_1)$. Moreover, we have here the strict inequality if for this k,

$$\epsilon'_{k-1} = 1 > l_{k-1} = 0 \qquad \text{and} \qquad R_{k-1}(y_{-1}) + \epsilon''_{k-1} \geqslant 1. \tag{3.8}$$

Once the properties (3.8) have been established for some k, inequality (3.5) will continue to hold for all larger k.

The reason for this is as follows. According to Lemma 3(ii), (3.8) produces a zero ξ of $P^{(k)}$ in J_2. This ξ, combined with any other zero, produces again a zero in J_2. There are zeros to combine with this ξ, for $P^{(k)}$ has either at least two zeros, or else $\lambda \neq \xi$ as a zero. Thus there will be a zero ξ' of $P^{(k+1)}$ in J_2. This shows that $R_{k+1}(y_{-1}) > 0$. But then by (3.7) and (3.6), $R_{k+1}(y_{-1}) = \epsilon_k' + R_k(y_{-1}) + \epsilon''_k - 1 > R_{k+1}(y_1)$. Similarly for $k + 2$ and so on.

It remains to establish (3.8). We show that for arbitrary λ, $y_{-1} \leqslant \lambda \leqslant y_1$, there is a $k < p$, for which: λ is not a root of $P^{(k)}$, but there are Rolle zeros of $P^{(k)}$ both to the right and the left of λ. Let $\epsilon_{i_1 j_1} = \epsilon_{i_2 j_2} = 1$ be the two elements that support $\epsilon_{ip} = 1$. Then $P^{(j_1)}$ has a Rolle zero to the left of λ, and $P^{(j_2)}$ a zero to the right of λ, with $j_1 < p$, $j_2 < p$. Assume that there are ν, $j_2 < \nu < p$, for which there is no Rolle zero of $P^{(\nu)}$ to the right of λ. Let ν be the smallest such integer. Then there is just one Rolle zero of $P^{(\nu-1)}$ to the right of λ. Moreover, λ is not a zero itself. But then $P^{(\nu-1)}$ must have at least two Rolle zeros, hence there is at least one zero to the left of λ.

Thus, we can assume that for each ν, $j_2 < \nu < p$, there is a zero to the right of λ, and similarly, for each ν, $j_1 < \nu < p$, there is a zero to the left of λ. Then we can take $k = p - 1$.

Now it is easy to complete the proof. Consider the curves (3.4), $\mu = x_i(\lambda)$, $i = 1,..., r$, and the straight line $\mu = \lambda$. For $\lambda = y_{-1}$ there are more curves (3.4) above the straight line than for $\lambda = y_1$. Therefore, for some λ_0, $y_{-1} < \lambda_0 < y_1$, the straight line intersects one of the curves. This means that $P^{(q)}(x, \lambda_0)$ has the zero $x = \lambda_0$. Hence P vanishes on S.

References

1. K. Atkinson and A. Sharma, A partial characterization of poised Hermite–Birkhoff interpolation problems. *SIAM J. of Numer. Anal.* **6** (1969), 230–235.
2. G. D. Birkhoff, General mean value and remainder theorems with applications to mechanical differentiation and integration, *Trans. Amer. Math. Soc.* **7** (1906), 107–136.
3. G. G. Lorentz and K. L. Zeller, Monotone approximation by algebraic polynomials, *Trans. Amer. Math. Soc.* **149** (1970), 1–18.
4. G. G. Lorentz and K. L. Zeller, Birkhoff interpolation, *SIAM J. of Numer. Anal.* **8** (1971), 43–48.
5. I. J. Schoenberg, On Hermite–Birkhoff interpolation, *J. Math. Anal. Appl.* **16** (1966), 538–543.
6. P. Turan, On rational polynomials, *Acta Univ. Szeged.* **11** (1946), 106–113.