ADVANCES IN APPLIED MATHEMATICS 10, 348–357 (1989)

Biorthonormal Systems, Partial Fractions, and Hermite Interpolation

LUIS VERDE-STAR

Departamento de Matemáticas, Universidad Autónoma Metropolitana, Apdo. Postal 55-534, México, D.F. 09340, México

Using some properties of dual bases in finite dimensional vector spaces we obtain elementary linear algebra proofs of the partial fractions decomposition and the Hermite interpolation theorems. We also obtain an explicit expression for the inverse of a confluent Vandermonde matrix, an algebraic version of the residue theorem for rational functions, and several inverse pairs of change of basis matrices on a space of polynomials.

1. INTRODUCTION

The usual proofs of the partial fractions decomposition theorem use either complex analysis methods, as in Henrici [4, Theorem 4.4h, p. 218], or they use induction and divisibility properties of polynomials, as in Birkhoff and MacLane [1, Theorem 18, p. 80].

On the other hand, for the Hermite interpolation theorem, besides the well-known complex variables proof, there exist several linear algebra proofs which are in general nonconstructive. See Davis [2, Example 6, p. 29].

In the present paper we present an elementary linear algebra proof of the partial fractions decomposition theorem, and we obtain as a corollary the Hermite interpolation theorem. We also obtain explicit expressions for the basic Hermite interpolation polynomials as linear combinations of powers.

We construct several pairs of dual bases on the vector space $\mathscr{P}$ of polynomials of degree at most equal to N and its algebraic dual space. Such dual pairs of bases are also called biorthonormal systems on $\mathscr{P}$. We get first a pair of dual bases by a simple modification of the standard power basis and its dual, which consists of Taylor functionals. Using only some basic properties of dual bases we obtain a simple proof of the partial fractions

0196-8858/89 $7.50

decomposition theorem. Then, applying Leibniz's rule we get another biorthonormal system, which gives us immediately a constructive proof of the Hermite interpolation theorem.

Another important biorthonormal system is the one formed by the basis of Horner polynomials and its dual. Thus dual pair allows us to find explicit formulas for the coefficients of the Hermite interpolation polynomials. This is equivalent to finding an explicit formula for the inverse of a confluent Vandermonde matrix.

The dual of the Horner basis yields a simple formula for the sum of the residues of a proper rational function, which is very useful for the efficient computation of partial fractions decompositions of rational functions with high order poles.

A very good reference for partial fractions is Mahoney and Sivazlian [5]. For the properties of dual bases that we consider in the next section see Nomizu [6], or Davis [2].

2. Dual Bases

In this section we present some elementary properties of dual bases in finite dimensional vector spaces. We also introduce some notation and terminology.

Let N be a nonnegative integer and let $\mathscr{E}$ be a complex vector space of dimension $N+1$. The elements of $\mathscr{E}^{N+1}$ will be considered as column vectors and will be denoted by boldface letters, for example $\mathbf{u} = (u_0, u_1, \dots, u_N)^{\mathrm{T}}$. A vector $\mathbf{u}$ in $\mathscr{E}^{N+1}$ is called an ordered basis for $\mathscr{E}$ if $\{u_0, u_1, \dots, u_N\}$ is linearly independent. The algebraic dual space of $\mathscr{E}$ is denoted by $\mathscr{F}$. A basic result in linear algebra says that $\mathscr{F}$ has dimension $N+1$.

For $\mathbf{u}$ in $\mathscr{E}^{N+1}$ and $\mathbf{f}$ in $\mathscr{F}^{N+1}$ we define the Gramian matrix $G(\mathbf{f}, \mathbf{u})$ by

$$G(\mathbf{f}, \mathbf{u}) = \left[f_i u_j \right], \qquad 0 \le i,\ j \le N, \tag{2.1}$$

where the index i corresponds to rows and j to columns.

The proof of the following proposition is straightforward and we omit it.

Proposition 2.1. *Let* $\mathbf{u}$ *be an element of* $\mathscr{E}^{N+1}$ *and let* $\mathbf{f}$ *be an element of* $\mathscr{F}^{N+1}$, *then*

(i) $G(\mathbf{f}, \mathbf{u})$ *is nonsingular if and only if* $\mathbf{f}$ *and* $\mathbf{u}$ *are ordered basis for* $\mathscr{F}$ *and* $\mathscr{E}$, *respectively.*

(ii) *If* $\mathbf{v} = A\mathbf{u}$ *and* $\mathbf{g} = B\mathbf{f}$ *then*

$$G(\mathbf{g}, \mathbf{v}) = BG(\mathbf{f}, \mathbf{u})A^{\mathrm{T}}. \tag{2.2}$$

The ordered bases $\mathbf{u}$ and $\mathbf{f}$ of $\mathscr{E}$ and $\mathscr{F}$ respectively are called dual bases if and only if $G(\mathbf{f},\mathbf{u}) = I$. A pair of dual bases $(\mathbf{f},\mathbf{u})$ is also called a biorthonormal system on $\mathscr{E}$. This is the terminology used by Davis [2].

PROPOSITION 2.2. (i) *Given an ordered basis* $\mathbf{u}$ *of* $\mathscr{E}$ *there exists a unique ordered basis* $\mathbf{f}$ *of* $\mathscr{F}$ *such that* $(\mathbf{f},\mathbf{u})$ *is a biorthonormal system on* $\mathscr{E}$. *The symmetric statement with* $\mathbf{u}$ *and* $\mathbf{f}$ *interchanged also holds.*

(ii) *Let* $\mathbf{u}$ *and* $\mathbf{f}$ *be dual bases. Then for every* w *in* $\mathscr{E}$ *and every* h *in* $\mathscr{F}$ *we have*

$$w = \sum_{k=0}^{N} f_k w u_k \tag{2.3}$$

and

$$h = \sum_{k=0}^{N} h u_k f_k. \tag{2.4}$$

(iii) *Let* $\mathbf{u}$ *and* $\mathbf{f}$ *be dual bases and let* $\mathbf{v} = A\mathbf{u}$ *and* $\mathbf{g} = B\mathbf{f}$. *Then* $\mathbf{g}$ *and* $\mathbf{v}$ *are dual bases if and only if* $B^{-1} = A^{\mathrm{T}}$.

Proof. Part (i) is a standard result on dual bases. See Davis [2, Theorem 1.2.1] or any book on linear algebra. Part (ii) is an immediate consequence of the relations $f_i u_j = \delta_{i,j}$. Part (iii) follows from Eq. (2.2).

3. BIORTHONORMAL SYSTEMS ON SPACES OF POLYNOMIALS

We denote by $\mathscr{P}$ the complex vector space of polynomials in the variable z with degree at most equal to N. The standard power basis of $\mathscr{P}$ is the ordered basis $\mathbf{s} = (s_0, s_1, \ldots, s_N)^{\mathrm{T}}$, where $s_k(z) = z^k$, for $0 \le k \le N$. The dual space of $\mathscr{P}$ is denoted by $\mathscr{F}$.

In order to simplify the notation we define the operators $d^k = D^k/k!$, for $k \ge 0$, where D denotes the usual differentiation operator. With this notation Leibniz's rule is

$$d^k(uv) = \sum_{j=0}^{k} d^j u\, d^{k-j} v.$$

For any complex number x we denote by E_x the linear functional of evaluation at $z = x$ defined by $E_x p(z) = p(x)$, for $p \in \mathscr{P}$. The dual basis of $\mathbf{s}$ is the basis of Taylor functionals at $x = 0$, $\mathbf{T}_0 = (T_0, T_1, \ldots, T_N)^{\mathrm{T}}$, where $T_k p = E_0\, d^k p(z)$, for $0 \le k \le N$, since

$$T_k s_j(z) = \binom{j}{k} E_0 z^{j-k} = \delta_{j,k}, \qquad 0 \le k,\ j \le N. \tag{3.1}$$

Another pair of dual bases is obtained as follows. For any number x the shifted powers $u_j(z) = (z-x)^j$ form a basis of $\mathscr{P}$. Its dual basis is the set of Taylor functionals $T_{x,k} = E_x\, d^k$, for $0 \le k \le N$, because

$$T_{x,k} u_j(z) = \binom{j}{k} E_x (z-x)^{j-k} = \delta_{j,k}, \qquad 0 \le j,\ k \le N. \tag{3.2}$$

In some cases a dual basis may be constructed directly, without the use of matrix inversion. The shifted powers are one example. Another typical example is the Lagrange interpolation problem, where, given the functionals $E_{z_0}, E_{z_1}, \ldots, E_{z_N}$, with distinct z_i's, we want to find the dual basis $\{l_0, l_1, \ldots, l_N\}$. This is easily done defining

$$l_k(z) = \prod_{\substack{i=0 \\ j \ne k}}^{N} \frac{(z - z_j)}{(z_k - z_j)}, \qquad 0 \le k \le N, \tag{3.3}$$

called the basic Lagrange interpolation polynomials for the points $z_0, z_1, \ldots, z_N$. It is clear that $E_{z_j} l_k = \delta_{j,k}$, and therefore we have a biorthonormal system.

We present next an important example motivated by the decomposition of rational functions into partial fractions.

Let $z_0, z_1, \ldots, z_n$ be distinct real or complex numbers. Let $m_0, m_1, \ldots, m_n$ be positive integers such that $\sum_i m_i = N + 1$. Define

$$w(z) = \prod_{i=0}^{n} (z - z_i)^{m_i}. \tag{3.4}$$

The partial fractions decomposition theorem says that for any $p \in \mathscr{P}$ there exist numbers $a_{i,k}$ such that

$$\frac{p(z)}{w(z)} = \sum_{i=0}^{n} \sum_{k=0}^{m_i - 1} \frac{a_{i,k}}{(z - z_i)^{m_i - k}}. \tag{3.5}$$

Multiplying both sides by $w(z)$ we get $p(z)$ expressed as a linear combination of the polynomials $q_{i,k}(z) = w(z)(z - z_i)^{k - m_i}$. We will show that these polynomials form a basis for $\mathscr{P}$, and we will construct the dual basis. We introduce some notation first.

We define the index set

$$J := \{(i, k) : 0 \le i \le n, 0 \le k \le m_i - 1\},$$

and we will always consider the elements of J ordered as

$$(0,0),(0,1),\dots,(0, m_0 - 1),(1,0),(1,1),\dots,$$
$$(1, m_1 - 1),\dots,(n,0),(n,1),\dots,(n, m_n - 1).$$

Let us define the polynomials

$$w_i(z) := \prod_{\substack{j=0\\ j\neq i}}^{n} (z - z_j)^{m_j}, \qquad 0 \le i \le n, \tag{3.6}$$

and

$$q_{i,k}(z) := w_i(z)(z - z_i)^k, \qquad (i, k) \in J. \tag{3.7}$$

Note that the $q_{i,k}$ are elements of $\mathscr{P}$; i.e., their degree is at most equal to N.

Since $q_{i,k}$ is a shifted power multiplied by w_i, in order to construct the elements of the dual basis we may try to eliminate first the factor w_i, and then proceed as in the case of shifted powers. Thus we define the linear functionals $L_{i,k}$ as

$$L_{i,k}p = E_{z_i} d^k \left(\frac{p(z)}{w_i(z)} \right), \qquad (i, k) \in J,\ p \in \mathscr{P}. \tag{3.8}$$

PROPOSITION 3.1. *The ordered bases* $\mathbf{L} = (L_{i,k} : (i, k) \in J)^{\mathrm{T}}$ *and* $\mathbf{q} = (q_{i,k} : (i, k) \in J)^{\mathrm{T}}$ *form a biorthonormal system on* $\mathscr{P}$. *That is,*

$$L_{r,s} q_{i,k} = \delta_{(r,s),(i,k)}, \qquad (r, s),(i, k) \in J.$$

Proof. If $r \neq i$, Leibniz's rule gives us

$$L_{r,s} q_{i,k} = E_{z_r} d^s \left(w_i(z) \frac{(z - z_i)^k}{w_r(z)} \right)$$
$$= E_{z_r} d^s \left((z - z_r)^{m_r} (z - z_i)^{k - m_i} \right) = 0,$$

since z_r is a root of w_i with multiplicity m_r, and $s < m_r$. If $r = i$ then

$$L_{i,s} q_{i,k} = E_{z_r} d^s (z - z_i)^k = \delta_{s,k}. \qquad \square$$

We obtain immediately

COROLLARY 3.2. *For any $p \in \mathscr{P}$ we have*

$$p(z) = \sum_{(i,k) \in J} L_{i,k} p q_{i,k}(z), \tag{3.9}$$

and

$$\frac{p(z)}{w(z)} = \sum_{(i,k) \in J} L_{i,k} p (z - z_i)^{k - m_i}. \tag{3.10}$$

Note that (3.10) is the partial fractions decomposition formula.

In order to simplify the notation we denote the Taylor functionals $E_{z_i} d^j$ by $T_{i,j}$, for $(i, j) \in J$.

Applying Leibniz's rule to (3.8) we get

$$L_{i,k} p = \sum_{j=0}^{k} T_{i,k-j} \left(\frac{1}{w_i(z)} \right) T_{i,j} p = \sum_{j=0}^{k} L_{i,k-j} 1 T_{i,j} p, \qquad p \in \mathscr{P}, \tag{3.11}$$

where 1 denotes the constant polynomial with value 1. Therefore

$$L_{i,k} = \sum_{j=0}^{k} L_{i,k-j} 1 T_{i,j}, \qquad (i,k) \in J. \tag{3.12}$$

Writing (3.12) in matrix notation we obtain

$$\mathbf{L} = \operatorname{diag}(B_0, B_1, \ldots, B_n) \mathbf{T}, \tag{3.13}$$

where $\mathbf{T} = (T_{i,j} : (i, j) \in J)^{\mathrm{T}}$, and for $0 \le i \le n$, B_i is the lower triangular Toeplitz block

$$B_i = \begin{pmatrix} L_{i,0}1 & & & & \\ L_{i,1}1 & L_{i,0}1 & & & \\ L_{i,2}1 & L_{i,1}1 & L_{i,0}1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ L_{i,m_i-1}1 & L_{i,m_i-2}1 & L_{i,m_i-3}1 & \cdots & L_{i,0}1 \end{pmatrix}.$$

Since $L_{i,0}1 = 1/w_i(z_i) \neq 0$, B_i is invertible and hence $\mathbf{T}$ is an ordered basis for $\mathscr{F}$. Its dual basis is the set of basic Hermite interpolation polynomials associated to the points $\{z_i\}$ with multiplicities $\{m_i\}$. If we denote the basic Hermite polynomials by $H_{i,k}(z)$, then, by Proposition 2.2 (iii) and (3.13) we have

$$\mathbf{H} = \operatorname{diag}(B_0^{\mathrm{T}}, B_1^{\mathrm{T}}, \ldots, B_n^{\mathrm{T}}) \mathbf{q}, \tag{3.14}$$

which is equivalent to the scalar equations

$$H_{i,k}(z) = \sum_{j=k}^{m_i-1} L_{i,j-k}1q_{i,j}(z), \qquad (i,k) \in J. \tag{3.15}$$

It is easy to see that

$$H_{i,k}(z) = q_{i,k}(z) \sum_{j=0}^{m_i-1-k} L_{i,j}1(z-z_i)^j, \qquad (i,k) \in J. \tag{3.16}$$

Note that the sum in (3.16) is the Taylor polynomial of degree $m_i - 1 - k$ of $1/w_i$ at $z = z_i$. Note also that $H_{i,k}$ has degree N for each $(i,k) \in J$.

Applying Proposition 2.2(ii) to the biorthonormal system $(\mathbf{T},\mathbf{H})$ we obtain the Hermite interpolation formula

$$p(z) = \sum_{(i,k)\in J} T_{i,k}pH_{i,k}(z), \qquad p \in \mathcal{P}. \tag{3.17}$$

We introduce next another important basis for $\mathcal{P}$, closely related to $w(z)$, which will allow us to express the polynomials $q_{i,k}$ and $H_{i,k}$ as linear combinations of powers.

Suppose that

$$w(z) = \sum_{k=0}^{N+1} b_k z^{N+1-k}. \tag{3.18}$$

Using the identity

$$\frac{z^{k+1} - t^{k+1}}{z - t} = \sum_{j=0}^{k} z^j t^{k-j}, \qquad k \geq 0, \tag{3.19}$$

and a simple rearrangement of terms, it is easy to see that

$$w[z,t] := \frac{w(z) - w(t)}{z - t} = \sum_{k=0}^{N} v_k(t) z^{N-k}, \tag{3.20}$$

where

$$v_k(t) = \sum_{j=0}^{k} b_j t^{k-j}, \qquad 0 \leq k \leq N+1. \tag{3.21}$$

The polynomials v_k are called the Horner polynomials of w because they satisfy the recurrence relation

$$v_{k+1}(t) = tv_k(t) + b_{k+1}, \qquad 0 \leq k \leq N, \tag{3.22}$$

and $v_{N+1} = w$. Note that (3.22) is Horner's algorithm for the computation of $w(t)$.

Since the polynomial $q_{i,k}$ is obtained dividing w by a power of $z - z_i$, Horner's algorithm gives us the coefficients in the expression for $q_{i,k}$ as a linear combination of powers of z.

PROPOSITION 3.3.

$$q_{i,k}(t) = T_{i,m_i-1-k}w[t,z], \qquad (i,k) \in J, \tag{3.23}$$

where the Taylor functional T_{i,m_i-1-k} acts with respect to the variable z.

Proof. Since

$$T_{i,j}\{w(t) - w(z)\} = \begin{cases} 0, & \text{if } 0 < j < m_i; \\ w(t), & \text{if } j = 0, \end{cases}$$

by Leibniz's rule we have

$$\begin{aligned} T_{i,k}w[t,z] &= E_{z_i}d^k\{(w(t) - w(z))(t-z)^{-1}\} \\ &= w(t)(t-z_i)^{-1-k} = q_{i,m_i-1-k}(t). \end{aligned}$$

COROLLARY 3.4.

$$\text{(i)} \qquad w[t,z] = \sum_{(i,k)\in J} q_{i,m_i-1-k}(t)H_{i,k}(z). \tag{3.24}$$

$$\text{(ii)} \qquad H_{i,k}(t) = L_{i,m_i-1-k}w[t,z], \tag{3.25}$$

where the functional acts with respect to z.

$$\text{(iii)} \qquad H_{i,k}(z) = \sum_{j=0}^{N} L_{i,m_i-1-k}v_j z^{N-j}. \tag{3.26}$$

$$\text{(iv)} \qquad H_{i,k}(z) = \sum_{j=0}^{N} L_{i,m_i-1-k}s_j v_{N-j}(z). \tag{3.27}$$

$$\text{(v)} \qquad q_{i,k}(z) = \sum_{j=k}^{N} T_{i,m_i-1-k}v_j z^{N-j}. \tag{3.28}$$

$$\text{(vi)} \qquad q_{i,k}(z) = \sum_{j=k}^{N} T_{i,m_i-1-k}s_j v_{N-j}(z). \tag{3.29}$$

$$\text{(vii)} \qquad \frac{w[z,t]}{w(z)} = \sum_{(i,k)\in J} H_{i,k}(t)(z-z_i)^{-1-k}. \tag{3.30}$$

Proof. Part (i) follows from Proposition 3.3, since (**T**, **H**) is a biorthonormal system on $\mathscr{P}$. Part (ii) follows from (i), since (**L**, **q**) is a biorthonormal system. Statements (iii) and (iv) are immediate consequences of (ii) and (3.20). Part (vii) clearly follows from (i) and clarifies the relationship between the basic Hermite polynomials and partial fractions decomposition.

The Hermite interpolation formula (3.17) gives us

COROLLARY 3.5.

$$\text{(i)} \qquad s_j(z) = \sum_{(i,k)\in J} \binom{j}{k} z_i^{j-k} H_{i,k}(z), \qquad 0 \le j \le N. \tag{3.31}$$

$$\text{(ii)} \qquad q_{i,j}(z) = \sum_{k=j}^{m_i-1} T_{i,k-j} w_i H_{i,k}(z), \qquad (i,j) \in J. \tag{3.32}$$

$$\text{(iii)} \qquad v_j(z) = \sum_{(i,k)\in J} T_{i,k} v_j H_{i,k}(z), \qquad 0 \le j \le N. \tag{3.33}$$

PROPOSITION 3.6. (i) *The dual basis of* **v** *is* $\mathbf{R} = (R_0, R_1, \ldots, R_N)^{\mathrm{T}}$, *where*

$$R_k = \sum_{(i,j)\in J} \binom{N-k}{j} z_i^{N-k-j} L_{i,m_i-1-j}, \qquad 0 \le k \le N. \tag{3.34}$$

(ii) *For any p in $\mathscr{P}$ we have*

$$\sum_{i=0}^{n} \left(\text{Residue at } z_i \text{ of } \frac{p}{w} \right) = T_N p = d^N p(0). \tag{3.35}$$

Proof. Part (i) follows immediately from Proposition 2.2(iii) and (3.29).

Since $R_N v_k = \delta_{N,k}$, for $0 \le k \le N$, and v_k is a monic polynomial for each k, it is clear that $R_N p = 0$ for any p whose degree is less than or equal to $N-1$, and also that $R_N z^N = 1$. Therefore $R_N p = T_N p = d^N p(0)$ for any p in $\mathscr{P}$.

On the other hand, taking $k = N$ in (3.34) we obtain

$$R_N = \sum_{i=0}^{n} L_{i,m_i-1},$$

and the definition of the functionals $L_{i,j}$ gives

$$L_{i,m_i-1} p = E_{z_i} d^{m_i-1} \left(\frac{p}{w_i} \right) = \text{Residue of } \frac{p}{w} \text{ at } z_i.$$

This completes the proof of (3.35).

Remark. Note that (3.35) is equivalent to a theorem of Hazony and Riley [3].

COROLLARY 3.7. *Let V be the matrix of order $N+1$ whose jth row consists of the numbers $\binom{j}{k} z_i^{j-k}$, for $(i, k) \in J$, in the usual order. Then the inverse of V is the matrix whose (i, k)th row consists of the numbers $L_{i, m_i - 1 - k} v_{N-k}$, for $0 \leq j \leq N$.*

Proof. This is just the matrix interpretation of the pair of inverse relations (3.31) and (3.26).

Remarks. The matrix V is a generalized Vandermonde matrix. In [8] we use Corollary 3.7 in order to construct an efficient algorithm for the computation of V^{-1}.

Some of the propositions presented above simplify the proofs of the main results in our previous paper [7].

Note that we have obtained other pairs of inverse relations; for example, (3.15) and (3.32) form an inverse pair, and so do (3.27) and (3.32).

REFERENCES

1. G. BIRKHOFF AND S. MACLANE, "A survey of Modern Algebra," 3rd ed., Macmillan Co., New York, 1965.
2. P. J. DAVIS, "Interpolation and Approximation," Dover, New York, 1975.
3. D. HAZONY AND J. RILEY, Evaluating residues and coefficients of high order poles, *IRE Trans. Automat. Control* **4** (1959), 132–136.
4. P. HENRICI, "Applied and Computational Complex Analysis," Vol. 1, Wiley, New York, 1974.
5. J. F. MAHONEY AND B. D. SIVAZLIAN, Partial fractions expansion: A review of computational methodology and efficiency, *J. Comput. Appl. Math.* **9** (1983), 247–269.
6. K. NOMIZU, "Fundamentals of Linear Algebra," 2nd ed., Chelsea, New York, 1979.
7. L. VERDE-STAR, Inverses of generalized Vandermonde matrices, *J. Math. Anal. Appl.* **131** (1988), 341–353.
8. L. VERDE-STAR, An algorithm for the inversion of generalized Vandermonde matrices, *Cuadrn. Anál. Apl. UAM-I* **3**, No. 2 (1988).