

Two Formulas for the General Multivariate Polynomial Which Interpolates a Regular Grid on a Simplex

By Coert Olmsted

Abstract. Two formulas are exhibited for the multivariate Lagrange shape polynomials which interpolate a regular grid on a simplex in \mathbb{R}^n .

1. Introduction. The idea of a shape function (or Lagrange) basis for an approximating function subspace is fundamental to the finite element method, spline theory and interpolation procedures in general. Given a set of points (or nodes) in the domain of the approximants, a *shape function* is an approximant associated with a given node and which assumes the value 1 at that node and the value zero at all other nodes in the set. If we have a shape function for each node, we say that the set of shape functions is *biorthonormal* to the node set. It is then very natural to attempt to force an approximating function to interpolate the node set (i.e., to take on arbitrary values at the given points) by constructing it as a linear combination of the shape functions. The coefficients in the linear combination are the arbitrary values to be assumed and the biorthonormal property ensures, in an obvious and elementary way, that this linear combination will indeed interpolate those values at the given nodes. Less obvious, but equally elementary, is the fact that biorthonormality of the shape functions implies that they are unique and linearly independent and so form a basis for the approximating subspace.

To state all this precisely, let X be a vector space over an ordered field F of character zero, and let \mathcal{F} be a family of functions defined on X and with values in F , which is itself a vector space over F by means of the operations induced from F . Let \mathcal{H} be an (approximating) subspace of \mathcal{F} of finite dimension N and let the node set $P = \{x^{(i)}\}_{i=1}^N$ be a subset of X . Using straightforward arguments of elementary linear algebra, as is done by Davis in Chapter 2 of [1] and Thacher [5], [6], it is easy to prove the following

FINITE LAGRANGE INTERPOLATION THEOREM. *The following four statements are equivalent:*

(a) *The node set P can be interpolated uniquely by functions of \mathcal{H} , i.e., for any set of*

Received October 30, 1984; revised September 30, 1985.

1980 *Mathematics Subject Classification.* Primary 41A05, 41A63.

Key words and phrases. Multivariate polynomial finite simplicial interpolation, total degree polynomial space Lagrange basis.

N values $\{a_i\} \subset F$, there is a unique function $p \in \mathcal{H}$ such that $p(x^{(i)}) = a_i$ for $i = 1, \dots, N$.

(b) For any basis $\{u_i\}_{i=1}^N$ of \mathcal{H} the Gram matrix $G = (u_j(x^{(i)}))_{i,j=1}^N$ is nonsingular.

(c) There exists a Lagrange basis at P for \mathcal{H} , i.e., there is a subset $\{l_i\}_{i=1}^N \subset \mathcal{H}$ which is biorthonormal to P . Such a system is necessarily unique and linearly independent in \mathcal{H} and so forms a basis for \mathcal{H} .

(d) The node set P does not lie in the null set of any nonzero function of \mathcal{H} , i.e., for every $f \in \mathcal{H}$, $f(P) = \{0\}$ implies $f = 0$.

The first condition is, of course, what we want. The second and third are what we can compute, while the fourth logical variant gives useful heuristics both for the abstract algebra of the interpolation scheme, and for the analytic geometry of the node set.

Although the proof of this theorem is elementary, it is instructive to work out one of the many possible cycles of implications in order to obtain some useful formulas. We use boldface notation to represent a column N -vector with components in F ,

$$\mathbf{v} = (v_1, v_1, \dots, v_N)^T,$$

and denote the standard matrix and vector linear products by juxtaposition.

(a) \Rightarrow (d). The contrapositive of this implication follows easily since a nonzero interpolant which vanishes on the node set may be added to any solution of the interpolation problem to obtain a different function which also interpolates. This spoils uniqueness in (a).

(b) \Rightarrow (a). This is the use of the Gram matrix which, if nonsingular, may be inverted to solve the linear system expressing the desired interpolant in terms of a given basis. Explicitly, this is

$$(1.1) \quad p(x) = \mathbf{u}(x)^T G^{-1} \mathbf{a}.$$

(c) \Rightarrow (b). A basis $\{l_j\}$ is biorthonormal to P if $l_j(x^{(i)}) = \delta_{ij}$, the Kronecker delta. In this case, $G = (\delta_{ij}) = I$ which is trivially nonsingular. Then (1.1) becomes

$$(1.2) \quad p(x) = \mathbf{1}(x)^T \mathbf{a}.$$

(d) \Rightarrow (c). This closing implication is not as obvious, but it also yields a useful formula. We construct a Lagrange basis as follows. For all $u, v \in \mathcal{H}$, define

$$\phi(u, v) = \sum_{k=1}^N u(x^{(k)}) v(x^{(k)}).$$

The field properties of F and (d) show that ϕ is an inner product on \mathcal{H} . By the Gram-Schmidt procedure we may construct a basis $\{u_i\}$ for \mathcal{H} which is orthonormal with respect to ϕ . That is,

$$(1.3) \quad \sum_{k=1}^N u_i(x^{(k)}) u_j(x^{(k)}) = \phi(u_i, u_j) = \delta_{ij}, \quad 1 \leq i, j \leq N.$$

Let G be the Gram matrix of this orthonormal basis. Then, by (1.3),

$$G^T G = \left(\sum_k u_i(x^{(k)}) u_j(x^{(k)}) \right) = (\delta_{ij}) = I.$$

This, incidentally, proves (b). Also, we have $G^{-1} = G^T$ (i.e., G is orthogonal) and so

$$(1.4) \quad (\delta_{ij}) = I = GG^{-1} = GG^T = \left(\sum_k u_k(x^{(i)}) u_k(x^{(j)}) \right).$$

Now define

$$(1.5) \quad l_i(x) = \sum_k u_k(x^{(i)}) u_k(x), \quad 1 \leq i \leq N.$$

Then, by (1.4),

$$l_i(x^{(j)}) = \sum_k u_k(x^{(i)}) u_k(x^{(j)}) = \delta_{ij},$$

and $\{l_i\}$ is biorthonormal to P . This concludes the proof.

The clear economy of (1.2) prompts us to seek out effective methods for determining and evaluating the Lagrange shape functions. Indeed, while (1.5) works only for a particular orthonormal basis, we can, given any basis $\{u_i\}$, take $a_j = \delta_{ij}$, $i, j = 1, \dots, N$ in (1.1), to get

$$(1.6) \quad l_i(x) = \mathbf{u}(x)^T G^{-1} \mathbf{e}_i,$$

where \mathbf{e}_i is the conventional i th unit basis vector in F^N . Thus, we can represent our Lagrange basis in terms of any given basis by means of the columns of the inverse of its Gram matrix.

Unfortunately, for conventional bases and most practical node sets, the Gram matrix and the Gram-Schmidt biorthogonalization are unwieldy or possibly ill-conditioned. We turn, therefore, to a particular case of some generality, where F is the real numbers, X is n -dimensional Euclidean space, \mathcal{H} is the space of multivariate polynomials of total degree m , and P is the intersection of the right unit simplex with the square lattice $(1/m)\mathbb{N}^n$.

2. Notation and Definitions.

2.1. *The Multivariate Polynomial Function Space.* Consider real Euclidean n -space

$$\mathbb{R}^n = \{x = (x_1, \dots, x_n) : x_i \in \mathbb{R}\},$$

with the standard vector and metric structure. Consider also n -tuples of nonnegative integers $\kappa = (\kappa_1, \dots, \kappa_n)$. If $x \in \mathbb{R}^n$, we write

$$(2.1) \quad x^\kappa \text{ for the number } \prod_{i=1}^n x_i^{\kappa_i}.$$

Following convention, we also write $|\kappa|$ for $\sum_{i=1}^n \kappa_i$. Then, if $|\kappa| \leq m$, we say that x^κ is a monomial function of degree $\leq m$ on \mathbb{R}^n . We may now write the general polynomial function of degree m on \mathbb{R}^n as

$$p_n^m(x) = \sum_{|\kappa| \leq m} a_\kappa x^\kappa,$$

where the a_κ are real coefficients. We define the multi-index set for the coefficients and exponents as

$$\mathcal{X} = \left\{ \kappa = (\kappa_1, \dots, \kappa_n) : \sum_{i=1}^n \kappa_i \leq m \right\}$$

where, since the κ_i are nonnegative, the sum condition implies that $\kappa_i \leq m$ for $1 \leq i \leq n$.

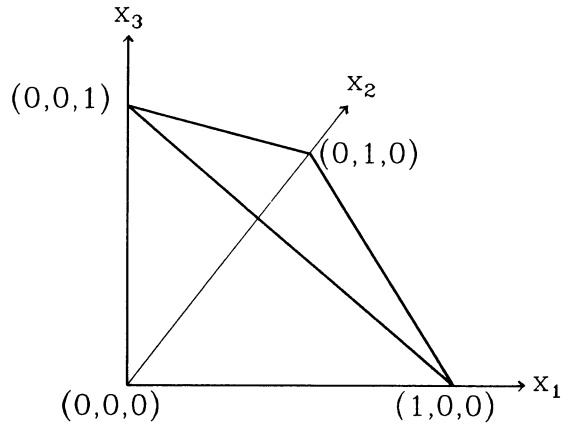


FIGURE 1

The simplex V in three-dimensional space ($n = 3$).

If we consider all such polynomials as a subset of the linear function space $C^\infty(\mathbb{R}^n)$, we note that it is a subspace of finite dimension spanned by the monomial functions, (2.1). That they are linearly independent is a consequence of a theorem of multivariate algebra which we paraphrase from Jacobson [3, p. 112].

THEOREM. *A multivariate polynomial over an infinite field is zero for all evaluations if and only if all its coefficients are zero.*

2.2. *The Node Set.* Consider the right unit simplex $V \subset \mathbb{R}^n$. If $x \in \mathbb{R}^n$, we write

$$x_0 = 1 - \sum_{i=1}^n x_i.$$

Then our simplex may be defined as

$$V = \{(x_1, \dots, x_n) : 0 \leq x_i \leq 1, 0 \leq i \leq n\}.$$

The $n + 1$ st constraint, $0 \leq x_0 \leq 1$, implies $\sum_{i=1}^n x_i \leq 1$, and so restricts x to a corner of the unit hypercube defined by the n other constraints. This corner contains the origin at its right vertex and is truncated by the hyperplane $x_0 = 0$ which forms the slant face in the simplex. The perpendicular faces are defined by $x_i = 0$, $1 \leq i \leq n$. The case $n = 3$ gives the simple geometry illustrated in Figure 1.

We now select a finite set of node points $P \subset V$ by restricting our domain to a regular orthogonal grid of spacing $1/m$ in each coordinate direction. Hence

$$P = \{(1/m)\kappa : \kappa \in \mathcal{X}\}.$$

These are just those points whose coordinates are nonnegative multiples of $1/m$ and which, since $|\kappa| \leq m$, are confined to the simplex. Thus \mathcal{X} also indexes our node set in a very natural way.

If we define κ_0 analogously to x_0 by

$$\kappa_0 = m - \sum_{i=1}^n \kappa_i,$$

we have immediately

$$\sum_{i=0}^n \kappa_i = m.$$

Define

$$\mathcal{X}_0 = \left\{ (\kappa_0, \kappa_1, \dots, \kappa_n) : \sum_{i=0}^n \kappa_i = m \right\}.$$

The mapping $(m - |\kappa|, \kappa_1, \dots, \kappa_n) \leftrightarrow \kappa$ is clearly a bijection between \mathcal{X}_0 and \mathcal{X} so that the two index sets are equinumerous. A simple combinatorial argument (Feller [2, II.5.2]) shows that this size is $N = \binom{n+m}{m} = \binom{n+m}{n}$. Thus there are N of the monomial basis functions defined in Subsection 2.1 so that the dimension of the polynomial subspace also is N .

2.3. *The First Shape Function Formula.* For our purposes we need to define a nodal shape polynomial of degree m for each point in P . To this end, let $(1/m)\kappa \in P$, and write

$$(2.2) \quad l_\kappa^m(x) = \prod_{i=0}^n \binom{mx_i}{\kappa_i} = \prod_{i=0}^n \frac{mx_i(mx_i - 1) \cdots (mx_i - \kappa_i + 1)}{\kappa_i(\kappa_i - 1) \cdots 2 \cdot 1},$$

where $\binom{a}{p}$ is the conventional binomial coefficient defined for $a \in \mathbb{R}$ and integer $p \geq 0$. To check that this is indeed a shape function, we observe that

(a) l_κ^m is of degree m . Clearly $\binom{mx_i}{\kappa_i}$ is of degree κ_i so l_κ^m is of degree $\sum_{i=0}^n \kappa_i = m$. In fact, l_κ^m is a product of exactly m linear factors.

(b) $l_\kappa^m((1/m)\kappa) = 1$. Substitution of κ_i/m for x_i in (2.2) yields this result immediately since $\binom{\kappa_i}{\kappa_i} = 1$.

(c) l_κ^m is zero at all nodes other than $(1/m)\kappa$. To show this, we need a short

LEMMA. *If κ' and κ are distinct elements of \mathcal{X}_0 , then there is a subscript q such that $\kappa'_q < \kappa_q$. (Note that this relationship is necessarily symmetrical so that also there is a $p \neq q$ such that $\kappa_p < \kappa'_p$.) Here we abuse notation slightly by writing κ for an element of \mathcal{X}_0 as well as \mathcal{X} . It will always be clear from the context whether we mean $\kappa = (\kappa_1, \dots, \kappa_n)$ or $\kappa = (\kappa_0, \kappa_1, \dots, \kappa_n)$.*

Proof of Lemma. Assume the contrary, i.e., for all q , $\kappa'_q \geq \kappa_q$. Then, since $\kappa' \neq \kappa$, we must have some p such that $\kappa'_p > \kappa_p$. These two inequalities imply $\sum_{i=0}^n \kappa'_i > \sum_{i=0}^n \kappa_i$, contrary to the sum condition defining \mathcal{X}_0 . Now, to demonstrate (c) above, let $\kappa' \neq \kappa$ and, by the Lemma, $\kappa'_q < \kappa_q$. Then the numerator of $\binom{\kappa'_q}{\kappa_q}$ is

$$\kappa'_q(\kappa'_q - 1) \cdots (\kappa'_q - \kappa_q) \cdots (\kappa'_q - (\kappa_q - 1)) = 0,$$

and $l_\kappa^m((1/m)\kappa')$ has at least one zero factor and so is zero.

3. Computing the First Formula. We have noted that l_κ^m has exactly m factors. If $n > m$ then some of the members of the product in (2.2) must be unity. These factors correspond to the components of κ which are zero.

We notice that, while the formula notation of (2.2) is simple, compact and classical, because of the unit factors its implementation produces a cumbersome algorithm. Furthermore, if we want to differentiate (2.2), even more redundant computational steps are introduced as the nested products are proliferated by the chain rule. It would be best if the m factors could be exhibited explicitly.

4. The Node Labels.

4.1. *Introductory Example.* For this purpose we introduce an alternate method of indexing the nodes, used by Leung and Ghaderpanah [4], which has the desired advantage of consisting of multi-indices of length m . While the set \mathcal{X} is closely associated with the coordinates of the nodes, our set \mathcal{N} of labels will be related more to the topology of their placement on the simplex. This is illustrated for the simple case $n = 3, m = 2$ in Figure 2. For labels we use ordered pairs of integers in the range 0, 1, 2, 3. The 4 vertices are labeled with (i, i) where i is the subscript of the coordinate axis passing through the vertex. The $n + 1$ st vertex, the origin, is assigned $(0, 0)$. The remaining 6 nodes are at the midpoints of the edges and so can be labeled (i, j) , where the edge is between vertex (i, i) and (j, j) . For uniqueness, we require $i < j$. Thus, we use all pairs of integers (i, j) such that $0 \leq i < j \leq 3$ of which there are $\binom{3+2}{2} = 10$. By examining cases of higher degree and dimension, where the nodes populate the faces and subcells of the simplex in a regular way, we can extend the labeling scheme to the general case.

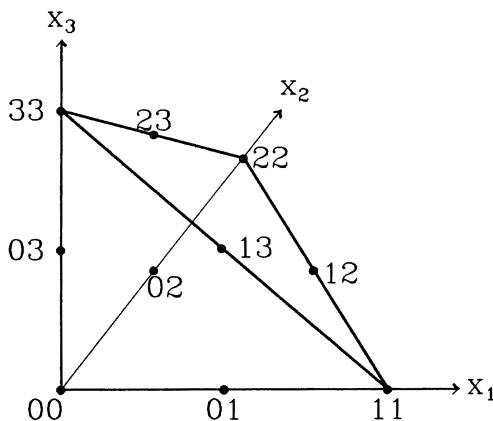


FIGURE 2

The node labels for the quadratic in three space ($n = 3, m = 2, N = 10$).

4.2. *Node Labels in Terms of Node Coordinates.* Thus, define

$$\mathcal{N} = \{ \nu = (\nu_1, \dots, \nu_m) : 0 \leq \nu_1 \leq \nu_2 \leq \dots \leq \nu_m \leq n \}.$$

That there are $N = \binom{n+m}{m}$ elements of \mathcal{N} can be proved by a straightforward induction argument on m using elementary properties of the binomial coefficient. For our purposes, however, we need more than an enumeration. In order to express our shape function in terms of the node labels, we need to construct a correspondence between \mathcal{X} and \mathcal{N} . In fact, we can define a correspondence between \mathcal{X}_0 and \mathcal{N} by setting components of ν to be the subscripts of the nonzero components of κ .

More precisely, given $\kappa = (\kappa_0, \kappa_1, \dots, \kappa_n)$, let $\kappa_{s(1)}, \kappa_{s(2)}, \dots, \kappa_{s(p)}$, be the positive components in order that $0 \leq s(1) < s(2) < \dots < s(p) \leq n$. By the sum condition on \mathcal{X}_0 ,

$$\sum_{r=1}^p \kappa_{s(r)} = m,$$

and so we must have $p \leq \min(m, n + 1)$. Thus, the subscripts s are in the correct range, there are not too many of them, and they are in the proper order to serve as values for the components of ν . Further, we are tying them to the nonzero components of κ which are just those which contribute nonunit factors to the product in formula (2.2). What remains in setting up a unique correspondence is to show how the s 's are distributed among the ν 's when $p < m$. We do this by repeating $s(q)$, as a component of ν , $\kappa_{s(q)}$ times. Thus, define the cumulative sum of the positive components of κ as a function of the s index

$$c(q) = \begin{cases} \sum_{r=1}^q \kappa_{s(r)} & \text{if } q \geq 1, \\ 0 & \text{if } q = 0. \end{cases}$$

Then we may define the components of ν as

$$\nu_j = s(q) \quad \text{if } c(q - 1) < j \leq c(q).$$

More explicitly,

$$\nu_j = \begin{cases} s(1) & \text{if } 1 \leq j \leq \kappa_{s(1)}, \\ s(2) & \text{if } \kappa_{s(1)} < j \leq \kappa_{s(1)} + \kappa_{s(2)}, \\ \dots & \\ s(p) & \text{if } \sum_{r=1}^{p-1} \kappa_{s(r)} < j \leq \sum_{r=1}^p \kappa_{s(r)} = m. \end{cases}$$

Since $s(q)$ and $c(q)$ are strictly increasing integers with q , we can see that the ν_j are nondecreasing with j as it ranges over all integers from 1 to m . Thus all the conditions defining \mathcal{N} are met.

4.3. *Equivalence of Node Labels and Coordinates.* If we regard ν_j as a function of j , it is a nondecreasing integer-valued step function of p steps where the height of each step is given by the subscripts of the nonzero components of κ and the width of each step is the integer value of that component. This is illustrated in Figure 3 by an example.

The step function concept of ν enables us to show that our correspondence is injective. Thus, for $\kappa \neq \kappa'$, $\kappa, \kappa' \in \mathcal{X}_0$, we wish to show that the corresponding labels ν and ν' are distinct. Consider three exhaustive cases:

- (a) $p \neq p'$. Then there will be a different number of steps in ν' than in ν .
- (b) $p = p'$ but for some q , $s(q) \neq s'(q)$. Then ν and ν' will differ in height at the q th step.
- (c) $p = p'$, for all q , $s(q) = s'(q)$. Then to have $\kappa \neq \kappa'$ there must be some q such that $\kappa_{s(q)} \neq \kappa_{s'(q)}$ and the widths of the q th steps will be different.

That our correspondence is surjective is more easily seen since we may write the inverse correspondence as

$$\kappa_i = \sum_{j=1}^m \delta(i, \nu_j), \quad 0 \leq i \leq n,$$

where δ is the Kronecker delta, now using functional argument notation.

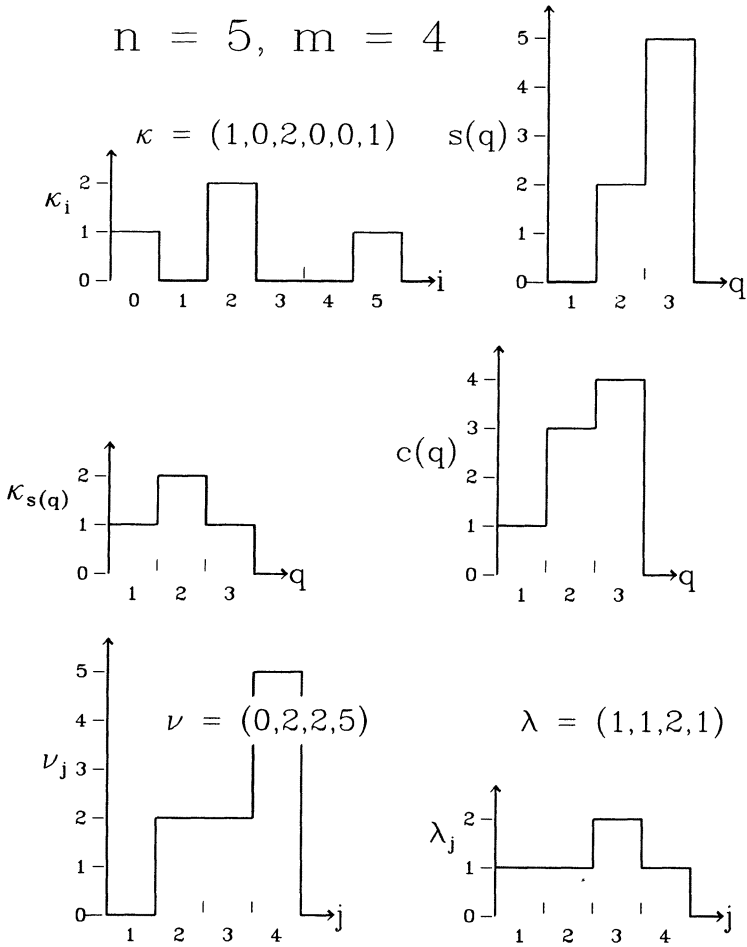


FIGURE 3

Step function graphs of the intermediate stages in the correspondence between an example node coordinate index, the node label, and the local multiplicity.

5. The Shape Functions in Terms of Node Labels.

5.1. *The Second Formula.* The summation of the δ function over ν gives the multiplicity of occurrences of i in ν which is just the value of the i th component of κ from the definition. Let the local multiplicity, λ_j , of a node label component, ν_j , be the number of its previous occurrences in ν ,

$$\lambda_j = \sum_{r=1}^j \delta(\nu_r, \nu_j).$$

Thus, as j increases, λ_j will count the cumulative occurrences of identical values for ν_j from 1 to its maximum multiplicity which is some component κ_i of κ .

Thus λ_j steps out the values $1, 2, \dots, \kappa_j$ as long as ν_j is constant in ν . Since there is a string of constant ν_j 's in ν for every nonzero κ_i , we see that the transformation

of formula (2.2) becomes

$$(5.1) \quad l_\nu^m(x) = \prod_{j=1}^m (mx_{\nu_j} - \lambda_j + 1) / \lambda_j.$$

For the reader who has meticulously followed the mappings from κ to ν to λ defined above, it will be apparent that the formulas (2.2) and (5.1) represent the identical expression, even to the order of the factors. For the less than meticulous, we can show that the two formulas produce the same polynomial function by proving that the second polynomials are biorthonormal to P . The identity of the two polynomial functions then follows from the uniqueness of the Lagrange basis. The proof necessarily echoes the proof for formula (2.2).

Thus, let

$$l_\nu^m((1/m)\kappa) = \prod_{j=1}^m (\kappa_{\nu_j} - \lambda_j + 1) / \lambda_j.$$

From our remarks above we have for each distinct component of ν ,

$$\lambda_j = 1, 2, \dots, \kappa_{\nu_j},$$

while

$$\kappa_{\nu_j} - \lambda_j + 1 = \kappa_{\nu_j}, \kappa_{\nu_j} - 1, \dots, 1.$$

Thus numerator and denominator consist of the same numbers in reverse order so that the entire quotient is 1.

Now let $\kappa \in \mathcal{X}_0$ and $\kappa' \neq \kappa$. By the Lemma of Subsection 2.3 there exists a subscript q such that $\kappa'_q < \kappa_q$. From the definitions of ν and λ there is a j such that

$$\left. \begin{array}{l} \nu_i = q \\ \text{and} \\ \lambda_i = \kappa_q - i + 1 \end{array} \right\} \text{ for } j - \kappa_q + 1 \leq i \leq j.$$

Since $\kappa'_q < \kappa_q$, there is one such i such that $\lambda_i = \kappa'_q + 1$ and $\nu_i = q$. For this i the factor in

$$l_\nu^m((1/m)\kappa') = \prod_{j=1}^m (\kappa'_{\nu_j} - \lambda_j + 1) / \lambda_j$$

is

$$(\kappa'_{\nu_i} - \lambda_i + 1) / \lambda_i = (\kappa'_q - \lambda_i + 1) / \lambda_i = 0.$$

This completes the proof of biorthonormality.

Acknowledgments. The author would like to thank Professor Gary Gislason for supporting this research, carefully reading several drafts, and providing many helpful discussions. Financial support was provided in part by the Geophysical Institute of the University of Alaska.

Geophysical Institute
 University of Alaska
 Fairbanks, Alaska 99775-0800

1. P. J. DAVIS, *Interpolation and Approximation*, Dover, New York, 1975.
2. W. FELLER, *An Introduction to Probability Theory and Its Applications*, Vol. I, Wiley, New York, 1957.

3. N. JACOBSON, *Lectures in Abstract Algebra*, Van Nostrand, New York, 1951.
4. K. V. LEUNG & S. S. GHADERPANAHI, "An application of the finite element approximation method to find the complex zeros of the modified Bessel function $K_n(z)$," *Math. Comp.*, v. 33, 1979, pp. 1299–1306.
5. H. C. THACHER, JR., "Derivation of interpolation formulas in several independent variables," *Ann. New York Acad. Sci.*, v. 86, 1960, pp. 758–775.
6. H. C. THACHER, JR. & W. E. MILNE, "Interpolation in several variables," *J. Soc. Indust. Appl. Math.*, v. 8, 1960, pp. 33–42.