

ON THE THEORY OF BIORTHOGONAL POLYNOMIALS

A. ISERLES AND S. P. NØRSETT

ABSTRACT. Let $\varphi(x, \mu)$ be a distribution in $x \in \mathbf{R}$ for every μ in a real parameter set Ω . Subject to additional technical conditions, we study m th degree monic polynomials p_m that satisfy the biorthogonality conditions

$$\int_{-\infty}^{\infty} p_m(x) d\varphi(x, \mu_l) = 0, \quad l = 1, 2, \dots, m, \quad m \geq 1,$$

for a distinct sequence $\mu_1, \mu_2, \dots \in \Omega$. Necessary and sufficient conditions for existence and uniqueness are established, as well as explicit determinantal and integral representations. We also consider loci of zeros, existence of Rodrigues-type formulae and reducibility to standard orthogonality. The paper is accompanied by several examples of biorthogonal systems.

1. Introduction. The theme of the present paper is a generalization of the classical concept of orthogonality into the following framework: Let $\varphi(x, \mu)$ be, for every μ in a real parameter set Ω , a *distribution* in x —a real monotonically nondecreasing, right-continuous function with all moments bounded and with an infinite number of points of increase. Given $m \geq 1$ and m distinct points $\mu_1, \dots, \mu_m \in \Omega$, we consider m th degree monic polynomials $p_m \equiv p_m(x; \mu_1, \dots, \mu_m)$ that satisfy the *biorthogonality* conditions

$$\int_{-\infty}^{\infty} p_m(x; \mu_1, \dots, \mu_m) d\varphi(x, \mu_l) = 0, \quad l = 1, 2, \dots, m.$$

Biorthogonal polynomials include, as a special case, the familiar orthogonal polynomials. Hence, it is legitimate to ask what are the known properties of orthogonal polynomials that are shared in the present, more general, setting. We address ourselves to this question, exploring existence and uniqueness, explicit representation, location of zeros and the existence of Rodrigues-type formulae and of recurrence relations. A future paper will be devoted to extensions of the Christoffel-Darboux formula and to a general recurrence relation which holds for all biorthogonal polynomials.

Our motivation in the present work is threefold. Firstly, by elucidating properties of biorthogonal polynomials we throw light on orthogonal polynomials and show that some familiar properties of the latter are not generic, being rather a manifestation of biorthogonality. Secondly, biorthogonal polynomials are applicable to a whole range of problems in computational and applied analysis. Indeed, our original motivation in studying this construct came from a problem in designing multistep numerical methods for ordinary differential equations [Iserles & Nørsett, 1987a].

Received by the editors March 15, 1986.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 42C05; Secondary 26C05, 33A35.

Since then, applications have been found to studying zeros of polynomials [Iserles & Nørsett, 1987b] and to rational approximation of Stieltjes functions [Iserles & Saff, 1987]. A forthcoming paper of the authors, joint with P. E. Koch, will present applications to numerical quadrature. Finally, biorthogonal polynomials deserve attention on their own merit and present a whole spectrum of intriguing and challenging problems.

In §2 we define formally biorthogonal polynomials. We introduce the concept of *regularity*, which is equivalent to existence and uniqueness. Finally, we present two explicit formulae for general biorthogonal polynomials, one determinantal and one integral.

§3 is devoted to the study of loci of zeros. By introducing the *interpolation property* we generalize a standard technique from the theory of orthogonal polynomials to the present framework.

It is well known that classical orthogonal polynomials are characterized by the existence of a Rodrigues formula. In §4 we demonstrate that a similar concept, that of a *Rodrigues-type formula*, exists for some sets of biorthogonal polynomials. We characterize all such sets. This entails discussion of inclusion and equivalence relations among pairs of the form $\{\varphi, \Omega\}$.

A property that characterizes orthogonal polynomials is a three-term recurrence relation. In §5 we ask whether such a relation can be valid for a biorthogonal system—in that case, by the Favard theorem [Chichara, 1978], the system will be orthogonal (in the usual sense) with respect to some distribution function, possibly with a different support. We show that this is highly unlikely and provide restrictive necessary and sufficient conditions.

Finally, in §6 we present several different examples of biorthogonal systems of polynomials. Some of these are straightforward generalizations of familiar orthogonal polynomials, while others are of a different nature. We examine their properties and, in most cases, present explicit formulae.

We mention in passing that the phrase “biorthogonal polynomials” has been already used in the literature in reference to two polynomial sequences, $\{r_n\}_{n=0}^\infty$ and $\{s_n\}_{n=0}^\infty$, say, which are mutually orthogonal with respect to some distribution φ :

$$\int_{-\infty}^{\infty} r_n(x) s_m(x) d\varphi(x) = c_{n,m} \delta_{n,m}, \quad n, m \geq 0,$$

where $\delta_{n,m}$ is the Kronecker delta [Pastro, 1984; Szegő, 1982]. Indeed, polynomials that are biorthogonal in *both* senses will be featured in a forthcoming paper. However, no confusion should arise between the two concepts.

2. Existence and uniqueness. There are three equivalent ways to introduce biorthogonal polynomials. First, consider a distribution $\varphi(x, \mu)$, $x \in (a, b)$, $\mu \in \Omega$, where $(a, b) \subseteq \mathbf{R}$ and Ω is a real set. A sequence $\{\mu_l\}_{l=1}^\infty$ of points in Ω will be called a *D-sequence* if $\mu_l \neq \mu_k$ for all $l \neq k$. We henceforth assume that Ω contains at least one *D-sequence*. The set of all *D-sequences* in Ω will be denoted by Ω^D .

Let $\pi_m[x]$ denote the set of all m th degree polynomials. Given $\{\mu_l\} \in \Omega^D$, we say that monic $p_m \in \pi_m[x]$, $p_m \equiv p_m(x; \mu_1, \dots, \mu_m)$, is *biorthogonal* if

$$(1) \quad \int_a^b p_m(x; \mu_1, \dots, \mu_m) d\varphi(x, \mu_l) = 0, \quad 1 \leq l \leq m.$$

Moreover, $p_0(x) \equiv 1$. The set $\{p_m\}_{m=0}^\infty$, where $p_m \equiv p_m(x; \mu_1, \dots, \mu_m)$ —if it exists—will be called the *biorthogonal polynomial system* (BOPS) with respect to the distribution φ .

An alternative way is to consider generalized positive functions $\omega(x, \mu)$, instead of distributions $\varphi(x, \mu)$. (1) now has the form

$$\int_a^b p_m(x; \mu_1, \dots, \mu_m) \omega(x, \mu_l) dx = 0, \quad 1 \leq l \leq m.$$

ω can be interpreted as the Radon-Nikodym derivative of φ [Halmos, 1950].

Finally, biorthogonality may be introduced by considering a family of positive linear functionals \mathcal{L}_μ , $\mu \in \Omega$. Condition (1) now reads

$$\mathcal{L}_{\mu_l} p_m(\cdot; \mu_1, \dots, \mu_m) = 0, \quad 1 \leq l \leq m,$$

and the link with the original definition is provided via the Riesz representation theorem.

In the sequel we will, as a rule, use the formulation of distributions.

The distribution φ is called *regular* if for every $\{\mu_l\} \in \Omega^D$ there exists a unique BOPS.

THEOREM 1. *Let*

$$I_k(\mu) := \int_a^b x^k d\varphi(x, \mu), \quad \mu \in \Omega, \quad k \geq 0,$$

and

$$(2) \quad D_m(\mu_1, \dots, \mu_m) := \det \begin{bmatrix} I_0(\mu_1) & I_1(\mu_1) & \cdots & I_{m-1}(\mu_1) \\ I_0(\mu_2) & I_1(\mu_2) & \cdots & I_{m-1}(\mu_2) \\ \vdots & \vdots & & \vdots \\ I_0(\mu_m) & I_1(\mu_m) & \cdots & I_{m-1}(\mu_m) \end{bmatrix}, \quad m \geq 1.$$

Regularity is equivalent to the condition that $D_m(\mu_1, \dots, \mu_m) \neq 0$ for every $m \geq 1$ and every $\{\mu_l\} \in \Omega^D$. Furthermore, in the case of regularity p_m possesses the explicit representation

$$(3) \quad p_m(x; \mu_1, \dots, \mu_m) = \frac{1}{D_m(\mu_1, \dots, \mu_m)} \det \begin{bmatrix} I_0(\mu_1) & I_1(\mu_1) & \cdots & I_m(\mu_1) \\ I_0(\mu_2) & I_1(\mu_2) & \cdots & I_m(\mu_2) \\ \vdots & \vdots & & \vdots \\ I_0(\mu_m) & I_1(\mu_m) & \cdots & I_m(\mu_m) \\ 1 & x & \cdots & x^m \end{bmatrix}.$$

PROOF. Let $p_m(x) = x^m + \sum_{k=0}^{m-1} \sigma_k x^k$. Hence (1) is equivalent to

$$I_m(\mu_l) + \sum_{k=0}^{m-1} \sigma_k I_k(\mu_l) = 0, \quad 1 \leq l \leq m.$$

The theorem follows at once by Kramer's rule. \square

COROLLARY. φ is regular if and only if for all $m \geq 0$ and $\mu \in \Omega$ the set $\{I_0(\mu), I_1(\mu), \dots, I_m(\mu)\}$ is a Chebyshev system.

PROOF. Follows at once from (2) by the standard definition of a Chebyshev system [Karlin & Studden, 1966]. \square

The representation (3) is similar to the standard determinantal form of monic orthogonal polynomials [Chichara, 1978]. Our second representation has a different character.

THEOREM 2. Subject to regularity, it is true that

$$(4) \quad D_m(\mu_1, \dots, \mu_m) = \int_a^b \cdots \int_a^b \prod_{1 \leq i < j \leq m} (x_j - x_i) d\varphi(x_1, \mu_1) d\varphi(x_2, \mu_2) \cdots d\varphi(x_m, \mu_m)$$

and

$$(5) \quad p_m(x; \mu_1, \dots, \mu_m) = \frac{\int_a^b \cdots \int_a^b \prod_{1 \leq i < j \leq m} (x_j - x_i) \prod_{i=1}^m (x - x_i) d\varphi(x_1, \mu_1) \cdots d\varphi(x_m, \mu_m)}{D_m(\mu_1, \dots, \mu_m)}$$

PROOF. We write $d_m(x_1, \dots, x_m) := \prod_{1 \leq i < j \leq m} (x_j - x_i)$ as a Vandermonde determinant,

$$d_m(x_1, \dots, x_m) = \det \begin{bmatrix} 1 & x_1 & \cdots & x_1^{m-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_m & \cdots & x_m^{m-1} \end{bmatrix}.$$

Multiplying the last expression by $d\varphi(x_1, \mu_1)$ and integrating in x_1 from a to b yields

$$\int_a^b d_m(x_1, \dots, x_m) d\varphi(x_1, \mu_1) = \det \begin{bmatrix} I_0(\mu_1) & I_1(\mu_1) & \cdots & I_{m-1}(\mu_1) \\ 1 & x_2 & \cdots & x_2^{m-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_m & \cdots & x_m^{m-1} \end{bmatrix}.$$

Subsequent multiplications by $d\varphi(x_l, \mu_l)$ and integration in x_l , $l = 2, 3, \dots, m$, lead to the identity (4). Similarly, (5) can be obtained from (3) by starting with $d_{m+1}(x_1, \dots, x_m, x)$, instead of $d_m(x_1, \dots, x_m)$. \square

Subject to regularity, we define the m th generator as the integral

$$H_m(\mu) \equiv H_m(\mu; \mu_1, \dots, \mu_m) := \int_a^b p_m(x; \mu_1, \dots, \mu_m) d\varphi(x, \mu), \quad \mu \in \Omega.$$

LEMMA 3. (a) Let φ be regular. Then it is true that

$$(6) \quad H_m(\mu) = D_{m+1}(\mu_1, \dots, \mu_m, \mu) / D_m(\mu_1, \dots, \mu_m), \quad m \geq 1.$$

(b) If for every $x \in (a, b)$ $\varphi(x, \mu) \in C^1(\Omega)$, where Ω is an interval, then $H_m(\mu) \in C^1(\Omega)$ and there exists a function $h_m(\mu) \equiv h_m(\mu; \mu_1, \dots, \mu_m)$ such that $h_m(\mu) \neq 0$ for all $\mu \in \Omega$, $h_m \in C^1(\Omega)$ and

$$(7) \quad H_m(\mu) = h_m(\mu) q_m(\mu), \quad \text{where } q_m(\mu) := \prod_{l=1}^m (\mu - \mu_l).$$

PROOF. Identity (6) follows at once by multiplying (5) by $d\varphi(x, \mu)$ and integrating in x .

To prove (7) we note that $H_m(\mu) \in \text{Sp}\{I_0(\mu), I_1(\mu), \dots, I_m(\mu)\}$. Since φ is in $C^1(\Omega)$, so are the $I_k(\mu)$'s. By regularity $\{I_0(\mu), \dots, I_m(\mu)\}$ is a Chebyshev set, and hence the number of roots of the equation $\sum_{k=0}^m \alpha_k I_k(\mu)$, where $\alpha_0, \dots, \alpha_m$ are any real constants, in Ω is at most m [Powell, 1981]. Note that the number of roots is well defined, since all $I_k(\mu)$'s are in $C^1(\Omega)$. From the definition of biorthogonality

$$H_m(\mu_l) = 0, \quad 1 \leq l \leq m.$$

Hence H_m has m simple zeros in μ_1, \dots, μ_m , and does not vanish elsewhere in Ω . \square

Generators will be useful in §6, in the derivation of explicit forms of biorthogonal polynomials.

3. Loci of the zeros. The standard proof that zeros of orthogonal polynomials reside in the support of the underlying distribution and are distinct impinges on the fact that it is possible to interpolate with polynomials; in other words, successive powers of x form a Chebyshev system. In the present section we exploit a similar argument to investigate zeros of biorthogonal polynomials.

Throughout this section we assume that $d\varphi(x, \mu) = \omega(x, \mu) d\alpha(x)$, where α is a distribution, independent of μ . We say that ω has the *interpolation property* if for all $m \geq 1$, $\{\mu_l\} \in \Omega^D$, distinct $x_1, \dots, x_m \in (a, b)$, and $y_1, \dots, y_m \in \mathbf{R}$ there exist real constants β_1, \dots, β_m so that

$$(8) \quad \sum_{l=1}^m \beta_l \omega(x_k, \mu_l) = y_k, \quad 1 \leq k \leq m.$$

LEMMA 4. ω possesses the interpolation property if and only if

$$E_m \begin{pmatrix} x_1, \dots, x_m \\ \mu_1, \dots, \mu_m \end{pmatrix} \neq 0$$

for every $m \geq 1$, distinct sequence $x_1, x_2, \dots \in (a, b)$ and $\{\mu_l\} \in \Omega^D$. Here, in accordance with [Karlin & Studden, 1966],

$$E_m \begin{pmatrix} x_1, \dots, x_m \\ \mu_1, \dots, \mu_m \end{pmatrix} := \det \begin{bmatrix} \omega(x_1, \mu_1) & \omega(x_1, \mu_2) & \cdots & \omega(x_1, \mu_m) \\ \omega(x_2, \mu_1) & \omega(x_2, \mu_2) & \cdots & \omega(x_2, \mu_m) \\ \vdots & \vdots & & \vdots \\ \omega(x_m, \mu_1) & \omega(x_m, \mu_2) & \cdots & \omega(x_m, \mu_m) \end{bmatrix}.$$

PROOF. By standard solution of linear algebraic equations (8). \square

COROLLARY. ω has the interpolation property if and only if, subject to possible rearrangement of x_1, x_2, \dots , the function $\omega(x, \mu)$ is strictly totally positive of all orders $m \geq 1$.

PROOF. Follows at once from the definition of strict total positivity [Karlin & Studden, 1966]. \square

Lemma 4 provides a handy method to check for the interpolation property. It will be used extensively in §6.

The interpolation property is important because of the following result.

THEOREM 5. *If ω is C^1 in μ , possesses the interpolation property and corresponds to a regular distribution then each $p_m(x; \mu_1, \dots, \mu_m)$ has m distinct zeros in (a, b) .*

PROOF. We extend the standard argument that demonstrates that all the zeros of an orthogonal polynomial are distinct and reside within the support of the underlying distribution.

Because of the interpolation property and Lemma 4 the function ω is positive almost everywhere for $x \in (a, b)$. Hence

$$\int_a^b p_m(x) \omega(x, \mu_1) d\alpha(x) = 0$$

implies that p_m changes sign at least once in (a, b) . Let $\zeta_1 < \zeta_2 < \dots < \zeta_n$ be all the points in (a, b) where p_m changes sign and ζ_{n+1} be any point on (a, b) such that $p_m(\zeta_{n+1}) \neq 0$. Of course $n \leq m$.

Let us assume that $n \leq m - 1$. Due to the interpolation property there exist real $\beta_1, \dots, \beta_{n+1}$ such that

$$(9) \quad \sum_{l=1}^{n+1} \beta_l \omega(\zeta_k, \mu_l) = \delta_{n+1, k}, \quad 1 \leq k \leq n+1.$$

Certainly, $\beta_1^2 + \beta_2^2 + \dots + \beta_{n+1}^2 > 0$. Let us consider the function

$$f(x) := \sum_{l=1}^{n+1} \beta_l \omega(x, \mu_l).$$

By Lemma 4 $\{\omega(x, \mu_l)\}_{l=1}^{n+1}$ is a Chebyshev system. Thus, since ω is in $C^1(\Omega)$, f has at most n zeros in (a, b) [Powell, 1981]. However, (9) implies that $f(\zeta_k) = 0$, $1 \leq k \leq n$. Consequently, without loss of generality,

$$p_m(x) f(x) \geq 0, \quad x \in (a, b)$$

and is positive on a set of positive measure. Therefore

$$\sum_{l=1}^{n+1} \beta_l H_m(\mu_l) = \int_a^b p_m(x) f(x) d\alpha(x) > 0,$$

because α is a distribution.

This contradicts the biorthogonality condition, since $H_m(\mu_l) = 0$, $1 \leq l \leq m$, and $n+1 \leq m$. Therefore the assumption that $n \leq m-1$ is false and p_m has m distinct zeros in (a, b) . \square

It is worthwhile to note two points about the scope of the last theorem. Firstly, it is perfectly possible for the interpolation property to hold in the absence of regularity; cf. Example G. In that case biorthogonal polynomials, in general, do not exist. Secondly—and this is much more important—regularity does not imply the interpolation property; cf. Example E. Hence, there is an important gap in the totality of biorthogonal polynomials that is not covered by Theorem 5. It is, indeed, perfectly possible for some of the zeros of p_m to lie outside the support of φ (cf. Example E). The characterization of all distributions φ that give rise to regular BOPS with all zeros distinct and within the support is presently an open problem.

4. Rodrigues-type formulae. It is well known that the classical orthogonal polynomials (Jacobi, Laguerre and Hermite) are the only orthogonal polynomials that possess a Rodrigues representation of the form

$$p_m(x) = c_m \{w(x)\}^{-1} \frac{d^m}{dx^m} \{(\rho(x))^m w(x)\},$$

where w is the underlying weight function, c_m is a normalizing constant and $\rho \in \pi_2[x]$ is independent of m [Cryer, 1970; Chichara, 1978]. In the present section we explore a Rodrigues-type representation for biorthogonal polynomials. It will transpire that formulae of this kind exist for essentially just two BOPS, given by $\varphi(x, \mu) = x^\mu/\mu$, $x \in (0, 1)$, $\Omega = (0, \infty)$, and by $\varphi(x, \mu) = -e^{-x/\mu}$, $x \in (0, \infty)$, $\Omega = (0, \infty)$.

Let $g(x, \mu)$ be a function which is smoothly differentiable, monotone and integrable in $x \in (a, b)$ for all $\mu \in \Omega$. We define the differential operator $T_\mu: C^1(a, b) \rightarrow L(a, b)$, $\mu \in \Omega$, by

$$(10) \quad T_\mu f(x) := \frac{(\partial/\partial x)\{g(x, \mu)f(x)\}}{\partial g(x, \mu)/\partial x} = f(x) + \frac{g(x, \mu)}{\partial g(x, \mu)/\partial x} f'(x).$$

The function g is said to be *admissible* if both following conditions are satisfied:

$$(I) \quad T_\mu: \pi_m[x] \rightarrow \pi_m[x] \quad \text{for all } \mu \in \Omega \text{ and } m \geq 0;$$

$$(II) \quad T_\mu T_\nu = T_\nu T_\mu \quad \text{for every } \mu, \nu \in \Omega.$$

THEOREM 6A. Let $b < \infty$ and φ be continuous in x . Define g as

$$g(x, \mu) := \int_a^x d\varphi(x, \mu).$$

If g is admissible then the Rodrigues-type formula

$$(11) \quad p_m(x; \mu_1, \dots, \mu_m) = c_m \left(\prod_{l=1}^m T_{\mu_l} \right) \{(b-x)^m\}, \quad m \geq 1,$$

is valid. Here

$$c_m = (-1)^m \prod_{l=1}^m (1 + \beta(\mu_l)m)^{-1}, \quad m \geq 1,$$

where

$$(12) \quad \beta(\mu) := \frac{\partial}{\partial x} \left\{ \frac{g(x, \mu)}{\partial g(x, \mu)/\partial x} \right\}.$$

PROOF. We define p_m by (11) and demonstrate that it obeys the definition of a biorthogonal polynomial. The proof will then follow by regularity.

First we note that it follows from (I) that $p_m \in \pi_m$. Moreover, admissibility implies that

$$\begin{aligned} H_m(\mu_l) &= c_m \int_a^b \frac{(\partial/\partial x)\{g(x, \mu_l) \prod_{k \neq l} T_{\mu_k}(b-x)^m\}}{\partial g(x, \mu_l)/\partial x} d\varphi(x, \mu_l) \\ &= c_m g(x, \mu_l) \prod_{k \neq l} T_{\mu_k}(b-x)^m \Big|_a^b, \quad 1 \leq l \leq m. \end{aligned}$$

Thus, since $\prod_{k \neq l} T_{\mu_k}(b-x)^m$ vanishes for $x = b$ and $g(x, \mu)$ vanishes for $x = a$, $H_m(\mu_l) = 0$ for all $1 \leq l \leq m$, hence biorthogonality.

It remains to demonstrate that p_m is monic. Since

$$T_\mu x^m = x^m + m \frac{g(x, \mu)}{\partial g(x, \mu)/\partial x} x^{m-1}$$

admissibility implies that $g(x, \mu)/(\partial g(x, \mu)/\partial x)$ is linear in x , $\gamma(\mu) + x\beta(\mu)$, say, where β is given by (12). Hence

$$T_\mu(b-x)^m = (-1)^m(1+m\beta(\mu))x^m(1+O(x^{-1})) \quad (x \gg 0).$$

Consequently, the coefficient of x^m in $\prod_{i=1}^m T_{\mu_i}(b-x)^m$ is c_m^{-1} . This concludes the proof. \square

The present theorem seemingly fails when $\beta(\mu_l) = -1/m$ for some l , since then c_m becomes unbounded. However, it is then obvious from the proof that $\prod_{l=1}^m T_{\mu_l}(b-x)^m$ is biorthogonal to φ for μ_1, \dots, μ_m , but belongs to $\pi_{m-1}[x]$. This is ruled out by regularity.

THEOREM 6B. *Let $a < -\infty$ and φ be regular and continuous in x . Set*

$$g(x, \mu) := \int_x^b d\varphi(x, \mu).$$

If g is admissible then the Rodrigues-type formula

$$p_m(x, \mu_1, \dots, \mu_m) = c_m \left(\prod_{l=1}^m T_{\mu_l} \right) \{(x-a)^m\}, \quad m \geq 1,$$

is valid. Here

$$c_m = \prod_{l=1}^m (1 + \beta(\mu_l)m)^{-1}, \quad m \geq 1,$$

and $\beta(\mu)$ has been defined in (12).

PROOF. A replica of the proof of Theorem 6A.

It is now natural to ask what are all the admissible functions. This is answered in the following lemma.

LEMMA 7. *All admissible functions are necessarily of the form*

(i) $g(x, \mu) = (x+C)^{\delta(\mu)}\varepsilon(\mu)$; or

(ii) $g(x, \mu) = e^{\delta(\mu)x}\varepsilon(\mu)$,

where $\delta(\mu)$, $\varepsilon(\mu)$ are nonvanishing functions of μ and C is a constant, arbitrary subject to g being C^1 and integrable in x . Moreover, subject to integrability, all functions given by (i) and (ii) are admissible.

PROOF. By (10) condition (I) is equivalent to

$$(13) \quad \frac{g(x, \mu)}{\partial g(x, \mu)/\partial x} = \gamma(\mu) + x\beta(\mu).$$

We have already used this in the proof of Theorem 6A. Hence, g satisfies the ordinary differential equation

$$\frac{\partial}{\partial x} g(x, \mu) = \frac{1}{\gamma(\mu) + x\beta(\mu)} g(x, \mu),$$

whose solutions are

$$(14) \quad g(x, \mu) = (\gamma(\mu) + x\beta(\mu))^{1/\beta(\mu)} \sigma(\mu)$$

if $\beta(\mu) \neq 0$ and

$$(15) \quad g(x, \mu) = e^{x/\gamma(\mu)} \sigma(\mu)$$

otherwise. Here σ is an arbitrary nonvanishing function of $\mu \in \Omega$.

Given any $f \in C^2(a, b)$, (13) implies that

$$\begin{aligned} T_\mu T_\nu f(x) = & f(x) + ((\gamma(\mu) + \gamma(\nu)) + x(\beta(\mu) + \beta(\nu)))f'(x) \\ & + (\gamma(\mu) + x\beta(\mu))(\gamma(\nu) + x\beta(\nu))f''(x) \\ & + (\gamma(\mu) + x\beta(\mu))\beta(\nu)f'(x), \quad \mu, \nu \in \Omega. \end{aligned}$$

By (II) the above expression must be symmetric in μ and ν . This is equivalent to

$$(16) \quad \beta(\mu)\gamma(\nu) = \beta(\nu)\gamma(\mu), \quad \mu, \nu \in \Omega.$$

Let us suppose that for some $\mu, \nu \in \Omega$, $\beta(\mu) = 0$, $\beta(\nu) \neq 0$. Hence, by (16), $\gamma(\mu) = 0$. This is impossible since g is monotonic in x . Hence either $\beta(\mu) \neq 0$ for every $\mu \in \Omega$ or $\beta \equiv 0$ in Ω .

In the first case we have $\gamma(\mu)/\beta(\mu) = \gamma(\nu)/\beta(\nu) \equiv C$. Case (i) now follows by substituting this value of γ into (14) and setting $\varepsilon(\mu) := |\beta(\mu)|^{1/\beta(\mu)} \sigma(\mu)$, $\delta(\mu) := 1/\beta(\mu)$.

In the second case, $\beta \equiv 0$, (15) yields (ii), upon the substitution $\varepsilon(\mu) := \sigma(\mu)$, $\delta(\mu) := 1/\gamma(\mu)$; note that, by monotonicity, if $\beta \equiv 0$ then $\gamma(\mu) \neq 0$ for all $\mu \in \Omega$.

Finally, straightforward manipulation verifies admissibility of (i) and (ii), subject to integrability. \square

We can now proceed to identify all distributions φ that give, within the framework of Theorems 6A and 6B, rise to admissible functions g . First, however, we need to introduce partial ordering into the pairs $\{\varphi, \Omega\}$ of distributions and parameter sets. So far, there has been to some extent, some redundancy implicit in our theory; e.g., the pairs $\{\varphi(x, \mu), (0, \infty)\}$ and $\{\mu^2 \varphi(x, e^\mu), (-\infty, \infty)\}$ yield isomorphic and virtually identical BOPS, whereas the pair $\{\varphi(x, \mu), (0, 1)\}$ is, in a sense, “subordinate” to the pair $\{\varphi(x, \mu), (0, 2)\} \dots$

More formally, given $\{\varphi_1, \Omega_1\}$ and $\{\varphi_2, \Omega_2\}$, we say that $\{\varphi_1, \Omega_1\}$ is *subordinate* to $\{\varphi_2, \Omega_2\}$, denoting this relation by $\{\varphi_1, \Omega_1\} \preceq \{\varphi_2, \Omega_2\}$, if there exists a positive function $\beta(\mu)$, $\mu \in \Omega_1$, such that for every $\mu \in \Omega_1$, there exists $\nu \in \Omega_2$ which satisfies the equation

$$\varphi_1(x, \mu) = \beta(\mu)\varphi_2(x, \nu), \quad x \in (a, b).$$

Intuitively, φ_2 will generate a BOPS which is “richer” than the BOPS which is generated by φ_1 and includes that system as a restricted case. Subordination imposes partial order on all pairs of the form $\{\varphi, \Omega\}$, where φ is regular.

We say that $\{\varphi_1, \Omega_1\}$ and $\{\varphi_2, \Omega_2\}$ are *equivalent* if both $\{\varphi_1, \Omega_1\} \preceq \{\varphi_2, \Omega_2\}$ and $\{\varphi_2, \Omega_2\} \succeq \{\varphi_1, \Omega_1\}$. This relation factorizes pairs $\{\varphi, \Omega\}$ into equivalence classes.

Finally a pair $\{\varphi, \Omega\}$, where φ is regular, is called *maximal* if it is not subordinate to any nonequivalent pair $\{\varphi_1, \Omega_1\}$, φ_1 regular, over the interval (a, b) .

In our characterization of all regular φ that yield themselves to the conditions of Theorems 6A and 6B it is sufficient to restrict attention to maximal pairs. Moreover, as is the standard practice with orthogonal polynomials, we specialize our results to the finite interval $(0, 1)$ and the singly infinite interval $(0, \infty)$. (Note that the interval $(-\infty, \infty)$ is outside the scope of these theorems. The choice of $(0, 1)$, rather than $(-1, 1)$, is dictated by an intention to present more elegant formulae.)

THEOREM 8. *All the nonequivalent maximal sets that lead to admissible functions are exactly*

- (a) $\varphi(x, \mu) = x^\mu/\mu$, $(a, b) = (0, 1)$; $\Omega = (0, \infty)$; and
- (b) $\varphi(x, \mu) = -e^{-x/\mu}$, $(a, b) = (0, \infty)$; $\Omega = (0, \infty)$.

PROOF. Follows at once by differentiating the expressions in the statement of Lemma 7. \square

The BOPS that correspond to the cases (a) and (b) will be presented in detail in §6, in Examples A and E respectively. In line with the standard terminology for orthogonal polynomials, we call them the *classical* biorthogonal polynomials.

5. Biorthogonality and orthogonality. In the present section we investigate the possibility of a biorthogonal system being orthogonal with respect to some distribution on the real line. If such a situation occurs then the standard theory of orthogonal polynomials provides a wealth of information on the underlying BOPS—a three-term recurrence relation, a Christoffel-Darboux identity, etc.

Perhaps unsurprisingly, it turns out that this state of affairs is unlikely: in general, biorthogonal polynomials are not orthogonal polynomials under disguise! In the sequel we provide both necessary and sufficient conditions for the identification of biorthogonality with orthogonality.

LEMMA 9. *Let φ be regular and $\{p_m\}$ be the underlying BOPS. Further, let us suppose that $\{\mu_l\} \in \Omega^D$ and a distribution ψ exist such that $\{p_m(\cdot; \mu_1, \dots, \mu_m)\}_{m=0}^\infty$ is orthogonal with respect to ψ ;*

$$\int_{-\infty}^{\infty} x^k p_m(x; \mu_1, \dots, \mu_m) d\psi(x) = 0, \quad 0 \leq k \leq m-1, \quad m \geq 1.$$

Then each $p_m(\cdot; \mu_1, \dots, \mu_m)$, $m \geq 1$, satisfies the orthogonality conditions

$$(17) \quad \int_a^b x^k p_m(x; \mu_1, \dots, \mu_m) d\varphi(x, \mu_l) = 0, \quad 0 \leq k \leq m-l, \quad 1 \leq l \leq m.$$

In particular, it is the monic orthogonal polynomial with respect to $d\varphi(\cdot, \mu_1)$.

PROOF. Orthogonality with respect to ψ and monicity imply that the set $\{p_m(\cdot, \mu_1, \dots, \mu_m)\}$ obeys a three-term recurrence relation

$$(18) \quad \begin{aligned} p_{m+1}(x; \mu_1, \dots, \mu_{m+1}) &= (c_m + x)p_m(x; \mu_1, \dots, \mu_m) \\ &\quad - d_m p_{m-1}(x, \mu_1, \dots, \mu_{m-1}), \quad m \geq 1, \end{aligned}$$

where c_m and $d_m > 0$, $m \geq 1$, are the elements of the corresponding Jacobi matrix.

Our proof proceeds by induction: Biorthogonality implies that (17) is true for $k = 0$, $1 \leq l \leq m$. Let us assume that (17) is valid for all $0 \leq k \leq k_l$, $1 \leq l \leq m-1$,

where $k_l \leq m - l - 1$. We multiply (18) by $x^{k_l} d\varphi(x, \mu_l)$ and integrate from a to b . This yields

$$\int_a^b x^{k_l+1} p_m(x; \mu_1, \dots, \mu_m) d\varphi(x, \mu_l) = 0.$$

The proof now follows. \square

To rephrase the statement of the last lemma, all we need to do to check for the coincidence of biorthogonality and orthogonality is to compare the BOPS with the monic orthogonal system with respect to $\varphi(x, \mu_1)$. Moreover, since the solution of (18) is unique, Lemma 9 also identifies ψ as $\varphi(\cdot, \mu_1)$.

LEMMA 10. *Let φ be regular and assume that $\{\mu_l\} \in \Omega^D$ exists such that, for some real constants c_0, c_1, \dots*

$$(19) \quad I_{m-k}(\mu_k) = c_{m-1}, \quad 1 \leq k \leq m, \quad m \geq 1,$$

and

$$(20) \quad D_m(\mu_1, \dots, \mu_m) > 0, \quad m \geq 1.$$

Then the set $\{p_m(\cdot; \mu_1, \dots, \mu_m)\}$ is orthogonal with respect to some distribution ψ .

PROOF. Since (19) is true, $D_m(\mu_1, \dots, \mu_m)$ is a Hankel determinant. Thus, by the positivity (20), the Hamburger moment problem for the sequence $\{c_k\}_{k=0}^\infty$ is soluble [Akhiezer, 1965]: a distribution ψ exists such that

$$c_k = \int_{-\infty}^{\infty} x^k d\psi(x), \quad k \geq 0.$$

A comparison of (3) with the standard determinantal representation of the monic polynomials which are orthogonal with respect to ψ [Chichara, 1978] leads readily to a proof of the lemma. \square

A closer examination of (19) reveals the structure that it imposes on the distribution φ . Operating within the conditions of the last lemma, we set

$$\varphi_l(x) := \varphi(x, \mu_l), \quad x \in (a, b), \quad l \geq 1.$$

We denote the k th moment of φ_l by $d_k^{(l)}$, $k \geq 0$. In other words, the sequence $\{d_k^{(l)}\}_{k=0}^\infty$ is a Hamburger sequence and corresponds to the distribution φ_l , $l \geq 1$.

Further, note that (19) is satisfied for $d\varphi_l(x) = x^{l-1} d\varphi_1(x)$, $l \geq 1$. This, together with

$$d_k^{(l)} = c_{k+l-1}, \quad k \geq 0, \quad l \geq 1,$$

and the uniqueness of the solution of the determinate moment problem imply that $d\varphi(x, \mu_l) = x^{l-1} d\varphi(x, \mu_1)$, $l \geq 1$, providing the link between Lemmas 9 and 10. It also implies that, besides the most obvious generalization of orthogonality into biorthogonality by setting $d\varphi(x, \mu) := x^\mu d\varphi(x, \mu)$, no coincidence is possible.

6. Examples of biorthogonal polynomials. Throughout the present section we will denote by $q_m(x)$ the polynomial $\prod_{k=1}^m (x - \mu_k)$ and by $\omega(x, \mu)$ the Radon-Nikodym derivative of $\varphi(x, \mu)$,

$$\omega(x, \mu) dx = d\varphi(x, \mu).$$

Besides presenting examples that illustrate the results of this paper, we also attempt an initial classification of several interesting families of BOPS.

If a regular φ is of the form

$$d\varphi(x, \mu) = x^\mu d\psi(x), \quad \mu \in \Omega, \quad x \in (a, b),$$

where ψ is independent of μ , we say that the corresponding BOPS belongs to *family I*:

EXAMPLE A. $\varphi(x, \mu) = x^\mu/\mu$, $\mu \in (0, \infty)$, $x \in (0, 1)$.

This BOPS, which is classical in the sense of §4, has been already introduced in [Iserles & Nørsett, 1987a]. The explicit form is

$$p_m(x) = \frac{1}{q_m(-m)} \sum_{k=0}^m (-1)^{m-k} \binom{m}{k} q_m(-k) x^k, \quad m \geq 1.$$

This can be ascertained, for example, from the Rodrigues-type representation (cf. §4)

$$(21) \quad p_m(x) = \frac{1}{q_m(-m)} T_{\mu_1} T_{\mu_2} \cdots T_{\mu_m} \{(1-x)^m\}, \quad m \geq 1,$$

where

$$T_\mu f(x) = \mu f(x) + x f'(x), \quad \mu > 0, \quad f \in C^1(0, \infty).$$

Both regularity and the interpolation property are easy to verify. We see that $D_m(\mu_1, \dots, \mu_m)$ is a determinant of a Cauchy matrix, whose value can be readily determined by a formula in [Gregory & Karney, 1969],

$$D_m(\mu_1, \dots, \mu_m) = 0!1!2! \cdots (m-1)! \frac{\prod_{1 \leq k < l \leq m} (\mu_l - \mu_k)}{\prod_{k=0}^{m-1} \prod_{l=1}^m (k + \mu_l)}.$$

Thus, by (6),

$$H_m(\mu) = (-1)^m \frac{m! q_m(\mu)}{(\mu)_{m+1} q_m(-m)}.$$

Regularity follows. Likewise,

$$E_m \begin{pmatrix} x_1, \dots, x_m \\ \mu_1, \dots, \mu_m \end{pmatrix} = \mu_1 \mu_2 \cdots \mu_m F_m \begin{pmatrix} x_1, \dots, x_m \\ \mu_1 - 1, \dots, \mu_m - 1 \end{pmatrix}$$

where

$$F_m \begin{pmatrix} x_1, \dots, x_m \\ \eta_1, \dots, \eta_m \end{pmatrix} := \det \begin{bmatrix} x_1^{\eta_1} & x_1^{\eta_2} & \cdots & x_1^{\eta_m} \\ \vdots & \vdots & & \vdots \\ x_m^{\eta_1} & x_m^{\eta_2} & \cdots & x_m^{\eta_m} \end{bmatrix}$$

is nonzero for all $m \geq 1$, distinct $x_1, x_2, \dots \in (0, 1)$ and distinct $\eta_1, \eta_2, \dots \in (-1, \infty)$ [Karlin & Studden, 1966]. Hence the interpolation property. Thus, each p_m has m distinct zeros in $(0, 1)$. It has been proved in [Iserles & Nørsett, 1987a] that, subject to $\mu_1 < \mu_2 < \cdots < \mu_m$, the zeros of $p_{m-1}(\cdot; \mu_1, \dots, \mu_{m-1})$ and $p_m(\cdot; \mu_1, \dots, \mu_m)$ interlace.

We now present two differential recurrence relations that are obeyed by the present BOPS. The first follows by differentiating (21). Since, in the present case,

$$\frac{d}{dx} T_\mu = T_{1+\mu} \frac{d}{dx}$$

it is true that

$$\begin{aligned}\frac{d}{dx}p_m(x; \mu_1, \dots, \mu_m) &= \frac{1}{q_m(-m)} \frac{d}{dx} T_{\mu_1} \cdots T_{\mu_m} (1-x)^m \\ &= -\frac{m}{q_m(-m)} T_{1+\mu_1} \cdots T_{1+\mu_m} (1-x)^{m-1} \\ &= \frac{m}{m+\mu_m} T_{1+\mu_m} p_{m-1}(x; 1+\mu_1, \dots, 1+\mu_{m-1}).\end{aligned}$$

Hence, it is true that

$$\begin{aligned}(m+\mu_m) \frac{d}{dx} p_m(x; \mu_1, \dots, \mu_m) &= m(1+\mu_m) p_{m-1}(x; 1+\mu_1, \dots, 1+\mu_{m-1}) \\ &\quad + mx \frac{d}{dx} p_{m-1}(x; 1+\mu_1, \dots, 1+\mu_{m-1}).\end{aligned}$$

To derive the second recurrence relation we multiply the m -degree polynomial

$$\begin{aligned}f_m(x) &:= \frac{(-1)^{m-1}}{q_m(-m+1)} p_m(x) + \frac{\mu_m(-1)^m}{q_m(-m)} p_{m-1}(x) - \frac{(-1)^m}{q_m(-m)} \frac{p_{m-1}(1)}{p_m(1)} x p'_m(x) \\ &\quad + \frac{(-1)^m}{q_m(-m)} x p'_{m-1}(x)\end{aligned}$$

by $d\varphi(x, \mu_l) = dx^{\mu_l}/\mu_l$ for l between 1 and m and integrate in x from 0 to 1. It follows easily from integration by parts and biorthogonality that

$$\int_0^1 f_m(x) dx^{\mu_l} = 0, \quad 1 \leq l \leq m.$$

Hence f_m is a multiple of p_m . Moreover, $f_m(0) = 0$, whilst $p_m(0) \neq 0$, since all the zeros of p_m reside in $(0, 1)$. Consequently $f_m \equiv 0$ and

$$\begin{aligned}(22) \quad &\frac{(-1)^{m-1}}{q_{m-1}(-m+1)} p_m(x) + \frac{\mu_m(-1)^m}{q_m(-m)} p_{m-1}(x) \\ &= \frac{(-1)^m}{q_m(-m)} \frac{p_{m-1}(1)}{p_m(1)} x p'_m(x) - \frac{(-1)^m}{q_m(-m)} x p'_{m-1}(x).\end{aligned}$$

To simplify (22) we need to evaluate

$$p_m(1) = \frac{1}{q_m(-m)} \sum_{k=0}^m (-1)^{m-k} \binom{m}{k} q_m(-k).$$

LEMMA 11. Let $r(x) = \sum_{j=0}^m r_j x^j$ be an arbitrary m th degree polynomial. Then

$$\sum_{k=0}^m (-1)^k \binom{m}{k} r(k) = (-1)^m m! r_m.$$

PROOF. We can write r in a unique way as

$$r(x) = \sum_{j=0}^m \frac{1}{j!} \tilde{r}_j (-x)_j,$$

where $\tilde{r}_m = (-1)^m m! r_m$. Hence

$$\begin{aligned} \sum_{k=0}^m (-1)^k \binom{m}{k} r(k) &= \sum_{k=0}^m (-1)^k \binom{m}{k} \sum_{j=0}^m \frac{1}{j!} \tilde{r}_j (-k)_j \\ &= \sum_{j=0}^m (-1)^j \binom{m}{j} \tilde{r}_j \sum_{k=j}^m (-1)^k \binom{m-j}{k-j} = \tilde{r}_m. \quad \square \end{aligned}$$

Setting $r(x) = q_m(-x)$ leads to $p_m(1) = (-1)^m m! / q_m(-m)$ and (22) simplifies to

$$\begin{aligned} p_m(x) - \mu_m \frac{q_{m-1}(-m+1)}{q_m(-m)} p_{m-1}(x) \\ = \frac{1}{m} x p'_m(x) + \frac{q_{m-1}(-m+1)}{q_m(-m)} x p'_{m-1}(x). \end{aligned}$$

The two recurrence relations can be combined into

$$\begin{aligned} q_m(-m) \{ m p_m(x; \mu_1, \dots, \mu_m) - x \frac{d}{dx} p_m(x; \mu_1, \dots, \mu_m) \} \\ = q_{m-1}(-m+1) (m + \mu_m - 1) \frac{d}{dx} p_m(x; \mu_1 - 1, \dots, \mu_m - 1). \end{aligned}$$

EXAMPLE B. $\omega(x, \mu) = e^{-\mu} x^\mu$, $\Omega = (0, \infty)$, $(a, b) = (0, \infty)$.

We have

$$I_k(\mu) = \Gamma(k + \mu + 1) = \Gamma(1 + \mu)(1 + \mu)_k, \quad k \geq 0,$$

therefore

$$D_m(\mu_1, \dots, \mu_m) = \prod_{l=1}^m \Gamma(1 + \mu_l) d_m(\mu_1, \dots, \mu_m),$$

where

$$d_m(\mu_1, \dots, \mu_m) := \det \begin{bmatrix} (1 + \mu_1)_0 & (1 + \mu_1)_1 & \cdots & (1 + \mu_1)_{m-1} \\ \vdots & \vdots & & \vdots \\ (1 + \mu_m)_0 & (1 + \mu_m)_1 & \cdots & (1 + \mu_m)_{m-1} \end{bmatrix}, \quad m \geq 1.$$

LEMMA 12. Let

$$(23) \quad e_m(\mu_1, \dots, \mu_m) := \det \begin{bmatrix} r_0(\mu_1) & r_1(\mu_1) & \cdots & r_{m-1}(\mu_1) \\ \vdots & \vdots & & \vdots \\ r_0(\mu_m) & r_1(\mu_m) & \cdots & r_{m-1}(\mu_m) \end{bmatrix},$$

where each r_k is a monic k th degree polynomial. Then

$$e_m(\mu_1, \dots, \mu_m) = \prod_{1 \leq k < l \leq m} (\mu_l - \mu_k).$$

PROOF. The matrix in (23) can be brought to a Vandermonde form by elementary column operations. \square

We now have

$$D_m(\mu_1, \dots, \mu_m) = \prod_{l=1}^m \Gamma(1 + \mu_l) \prod_{1 \leq k < l \leq m} (\mu_l - \mu_k),$$

hence regularity and

$$H_m(\mu) = \Gamma(1 + \mu)q_m(\mu).$$

The interpolation property is obvious, being valid for all BOPS in family I.

We will now derive the explicit form of p_m : Let $p_m(x) = \sum_{k=0}^m p_{m,k}x^k$. By Lemma 12 and (3)

$$\begin{aligned} q_m(x) &= \prod_{l=1}^m (x - \mu_l) \\ &= \frac{1}{\prod_{1 \leq k < l \leq m} (\mu_l - \mu_k)} \det \begin{bmatrix} (1 + \mu_1)_0 & \cdots & (1 + \mu_1)_m \\ \vdots & & \vdots \\ (1 + \mu_m)_0 & \cdots & (1 + \mu_m)_m \\ (1 + x)_0 & \cdots & (1 + x)_m \end{bmatrix} \\ &= \sum_{k=0}^m p_{m,k}(1 + x)_k. \end{aligned}$$

Thus, the $p_{m,k}$'s are the unique coefficients in the expansion of q_m in the $(m + 1)$ -dimensional basis $\{(1 + x)_0, (1 + x)_1, \dots, (1 + x)_m\}$. By applying the backward difference operator Δ_- on both sides we find

$$p_{m,k} = \frac{1}{k!} \sum_{l=0}^k (-1)^l \binom{k}{l} q_m(-l - 1) = \frac{1}{k!} \Delta_-^k q_m(-1), \quad 0 \leq k \leq m.$$

We now continue in the classification of BOPS. A BOPS is said to belong to *family II* if the underlying distribution φ is of the form

$$d\varphi(x, \mu) = \mu^x d\psi(x), \quad \Omega \subseteq (0, \infty),$$

where ψ is independent of μ . Again, the interpolation property is valid throughout this family, since the determinant E_m is of the form F_m (cf. Example A).

EXAMPLE C. $d\varphi(x, \mu) = \mu^x d\psi_\lambda(x)$, $\Omega = (0, \lambda)$, $(a, b) = (-\frac{1}{2}, \infty)$, where $\lambda > 0$ is fixed and ψ_λ is a step function with jumps of $((\lambda)_k/k!) \lambda^k$ at $k = 0, 1, \dots$

Let

$$\begin{aligned} I_k^*(\mu) &:= (-1)^k \int_a^b (-x)_k d\varphi(x, \mu), \quad \mu \in \Omega, \quad k \geq 0, \\ D_m^*(\mu_1, \dots, \mu_m) &:= \det \begin{bmatrix} I_0^*(\mu_1) & \cdots & I_{m-1}^*(\mu_1) \\ \vdots & & \vdots \\ I_0^*(\mu_m) & \cdots & I_{m-1}^*(\mu_m) \end{bmatrix}, \quad m \geq 1. \end{aligned}$$

Since $(-1)^k(-x)_k \in \pi_k[x]$ and is monic for all $k \geq 0$, we can use Lemma 12 to obtain at once

$$D_m^*(\mu_1, \dots, \mu_m) = D_m(\mu_1, \dots, \mu_m), \quad m \geq 1.$$

Likewise, the representation

$$p_m(x; \mu_1, \dots, \mu_m) = \frac{1}{D_m(\mu_1, \dots, \mu_m)} \det \begin{bmatrix} I_0^*(\mu_1) & \cdots & I_m^*(\mu_1) \\ \vdots & & \vdots \\ I_0^*(\mu_m) & \cdots & I_m^*(\mu_m) \\ (-1)^0(-x)_0 & \cdots & (-1)^m(-x)_m \end{bmatrix},$$

$m \geq 1,$

is an immediate consequence of (3) and Lemma 12.

In the present case

$$\begin{aligned} I_k^*(\mu) &= \sum_{l=k}^{\infty} \frac{(\lambda)_l}{(l-k)!} \left(\frac{\mu}{\lambda}\right)^l = (\lambda)_k \left(\frac{\mu}{\lambda}\right)^k {}_1F_0 \left[\begin{matrix} \lambda+k; \\ -; \end{matrix} \frac{\mu}{\lambda} \right] \\ &= \frac{(\lambda)_k}{(1-\mu/\lambda)^\lambda} \left(\frac{\mu}{\lambda-\mu}\right)^k, \quad k \geq 0. \end{aligned}$$

Therefore we get

$$\begin{aligned} D_m(\mu_1, \dots, \mu_m) &= D_m^*(\mu_1, \dots, \mu_m) \\ &= \frac{\prod_{l=0}^{m-1} (\lambda)_l}{\prod_{l=1}^m (1-\mu_l/\lambda)^\lambda} \text{VDM} \left(\frac{\mu_1}{\lambda-\mu_1}, \dots, \frac{\mu_m}{\lambda-\mu_m} \right) \\ &= \frac{\lambda^{\frac{1}{2}(m-1)m} \prod_{l=0}^{m-1} (\lambda)_l}{\prod_{l=1}^m (1-\mu_l/\lambda)^\lambda \prod_{l=1}^m (\lambda-\mu_l)^{m-1}} \prod_{1 \leq k \leq l \leq m} (\mu_l - \mu_k), \end{aligned}$$

where VDM denotes a Vandermonde determinant. Hence regularity and

$$(24) \quad H_m(\mu) = (\lambda)_m \left(\frac{\lambda}{\lambda-\mu}\right)^{m+\lambda} \frac{q_m(\mu)}{q_m(\lambda)}.$$

We will now prove that

$$p_m(x) = \frac{1}{q_m(\lambda)} \sum_{k=0}^m (-1)^k (-x)_k (\lambda+x)_{m-k} \lambda^k q_{m,k},$$

where $q_m(x) = \sum_{k=0}^m q_{m,k} x^k$. By regularity, it is enough to show that for the above definition of p_m

$$\int_{-1/2}^{\infty} p_m(x) \mu^x d\psi_\lambda(x) \equiv H_m(\mu),$$

the generator having been given by (24). But

$$\begin{aligned} \int_{-1/2}^{\infty} p_m(x) \mu^x d\psi_\lambda(x) &= \frac{1}{q_m(\lambda)} \sum_{l=0}^{\infty} \frac{(\lambda)_l}{l!} \left(\frac{\mu}{\lambda}\right)^l \sum_{k=0}^m (-1)^k (-l)_k (\lambda+l)_{m-k} \lambda^k q_{m,k} \\ &= \frac{1}{q_m(\lambda)} \sum_{k=0}^m \left\{ \sum_{l=k}^{\infty} \frac{(\lambda)_{m-k+l}}{(l-k)!} \left(\frac{\mu}{\lambda}\right)^l \right\} \lambda^k q_{m,k} \\ &= \frac{(\lambda)_m}{q_m(\lambda)} \sum_{k=0}^m q_{m,k} \mu^k {}_1F_0 \left[\begin{matrix} \lambda+m; \\ -; \end{matrix} \frac{\mu}{\lambda} \right] \\ &= (\lambda)_m \left(\frac{\lambda}{\lambda-\mu}\right)^{m+\lambda} \frac{q_m(\mu)}{q_m(\lambda)}, \end{aligned}$$

as required.

EXAMPLE D. $d\varphi(x, \mu) = \mu^x d\psi(x)$, $\Omega = (0, \infty)$, $(a, b) = (-\frac{1}{2}, \infty)$, where ψ is a step function with jumps of $1/k!$ at $k = 0, 1, \dots$.

The present BOPS is a limiting case of the BOPS from Example C when $\lambda \rightarrow \infty$ in a similar sense as the Charlier polynomials are a limiting case of the Meixner polynomials of the first kind [Chichara, 1978].

Hence, by letting λ tend to ∞ ,

$$D_m(\mu_1, \dots, \mu_m) = \exp \left(\sum_{l=1}^m \mu_l \right) \prod_{1 \leq k < l \leq m} (\mu_l - \mu_k),$$

$$H_m(\mu) = e^\mu q_m(\mu),$$

the distribution is regular and

$$p_m(x) = \sum_{k=0}^m (-1)^k (-x)_k q_{m,k}.$$

Our third and last family is different in character: we say that a BOPS belongs to *family III* if each $I_k(\mu)$ is a k th degree polynomial in μ for all $k \geq 0$.

EXAMPLE E. $\varphi(x, \mu) = \psi(x/\mu)$, $\Omega = (0, \infty)$, $(a, b) = (0, \infty)$. Here ψ is an arbitrary distribution. It follows at once that

$$I_k(\mu) = \int_0^\infty x^k d\psi \left(\frac{x}{\mu} \right) = \mu^k \int_0^\infty x^k d\psi(x) = c_k \mu^k, \quad k \geq 0,$$

where $c_k > 0$ is the k th moment of ψ . Note that $\{c_k\}_{k=0}^\infty$ is a Stieltjes sequence [Akhiezer, 1965].

Since $D_m(\mu_1, \dots, \mu_m)$ can be easily brought to a Vandermonde form, it follows at once that

$$D_m(\mu_1, \dots, \mu_m) = \prod_{k=0}^{m-1} c_k \prod_{1 \leq k < l \leq m} (\mu_l - \mu_k), \quad H_m(\mu) = c_m q_m(\mu).$$

Regularity is obvious. Moreover,

$$p_m(x) = c_m \sum_{k=0}^m \frac{q_{m,k}}{c_k} x^k.$$

This is true since for every $1 \leq l \leq m$

$$\int_0^\infty p_m(x) d\psi \left(\frac{x}{\mu} \right) = c_m \sum_{k=0}^m \frac{q_{m,k}}{c_k} \mu^k \int_0^\infty x^k d\psi(x) = c_m q_m(\mu) = H_m(\mu).$$

The interpolation property is not always valid; for example, it fails if the support of ψ is finite. Fortunately, it holds for the particular example of $\psi(x) = -e^{-x}$, which corresponds to our classical BOPS (cf. §4), since then E_m is of the form F_m from Example A.

An interesting case is

$$\psi(x) = \int_0^x \tau^\beta (1-\tau)^\gamma d\tau, \quad 0 \leq x \leq 1,$$

$$\psi(x) \equiv \Gamma(\beta+1)\Gamma(\gamma+1)/\Gamma(\beta+\gamma+2) \quad \text{for } 1 \leq x.$$

Here $\beta, \gamma > -1$ are constants. The underlying biorthogonal polynomials are

$$(25) \quad p_m^{(\beta, \gamma)}(x) = \frac{(\beta+1)_m}{(\beta+\gamma+2)_m} \sum_{k=0}^m \frac{(\beta+\gamma+2)_k}{(\beta+1)_k} q_{m,k} x^k, \quad m \geq 0.$$

The interpolation property fails, since ψ is constant in $[1, \infty)$. However, if γ is an integer then we can prove that all the zeros are distinct and in $(0, \infty)$ (but not necessarily in the true support $(0, 1)$):

To emphasize that γ is an integer, we replace it by $n \geq 0$. It is an easy matter to show from (25) by direct expansion that

$$(26) \quad p_m^{(\beta, n)}(x) = \frac{x^{-\beta}}{(\beta + m + 1)_{n+1}} \frac{d^{n+1}}{dx^{n+1}} \{x^{\beta+n+1} q_m(x)\}.$$

Positivity and distinctness of the zeros of $p_m^{(\beta, n)}$ follow by $n + 1$ consecutive applications of the Rolle theorem to the function $x^{\beta+n+1} q_m(x)$. Note that (26) is not of the same "kind" as Rodrigues' formulae—the degree of differentiation is dependent on n , rather than m .

Further, we prove that the zeros of $p_m^{(\beta, n)}$ and $p_m^{(\beta, n-1)}$ interlace for all $m \geq 1$, $n \geq 1$ and $\beta > 1$: Let $\nu := xd/dx$. Given $f \in C^\infty$, it is elementary that

$$x^j \frac{d^j}{dx^j} f(x) = (-1)^j (-\nu)_j f(x), \quad j \geq 0.$$

Acting on (26) with the Leibnitz rule and using the last identity leads to

$$\begin{aligned} p_m^{(\beta, n)}(x) &= \frac{x^{-\beta}}{(\beta + m + 1)_{n+1}} \sum_{k=0}^{n+1} \binom{n+1}{k} \left\{ \frac{d^{n+1-k}}{dx^{n+1-k}} x^{\beta+n+1} \right\} q_m^{(k)}(x) \\ &= \frac{(\beta + 1)_{n+1}}{(\beta + m + 1)_{n+1}} \sum_{k=0}^{n+1} \binom{n+1}{k} \frac{1}{(\beta + 1)_k} x^k q_m^{(k)}(x) \\ &= \frac{(\beta + 1)_{n+1}}{(\beta + m + 1)_{n+1}} \sum_{k=0}^{n+1} \binom{n+1}{k} \frac{(-1)^k (-\nu)_k}{(\beta + 1)_k} q_m(x) \\ &= \frac{(\beta + 1)_{n+1}}{(\beta + m + 1)_{n+1}} {}_2F_1 \left[\begin{matrix} -n-1, & -\nu; & 1 \end{matrix} \right] q_m(x) \\ &= \frac{(\beta + 1 + \nu)_{n+1}}{(\beta + m + 1)_{n+1}} q_m(x), \end{aligned}$$

where we have used the Vandermonde theorem [Rainville, 1967] to sum up a hypergeometric function with unit argument.

Since

$$\frac{(\beta + 1 + \nu)_{n+1}}{(\beta + m + 1)_{n+1}} = \frac{\beta + n + 1 + \nu}{\beta + m + n + 1} \frac{(\beta + 1 + \nu)_n}{(\beta + m + 1)_n},$$

it now follows that

$$p_m^{(\beta, n)}(x) = \frac{\beta + n + 1 + \nu}{\beta + m + n + 1} p_m^{(\beta, n-1)}(x),$$

or, after simplification,

$$(27) \quad (\beta + m + n + 1) p_m^{(\beta, n)}(x) = (\beta + n + 1) p_m^{(\beta, n-1)}(x) + x \frac{d}{dx} p_m^{(\beta, n-1)}(x).$$

Let $\zeta_k^{(\beta, n)}$, $1 \leq k \leq m$, denote the zeros of $p_m^{(\beta, n)}$, monotonically arranged; we already know that they are positive and distinct. It follows from (27) that

$$(\beta + m + n + 1) p_m^{(\beta, n)}(\zeta_k^{(\beta, n-1)}) = \zeta_k^{(\beta, n-1)} \frac{d}{dx} p_m(\zeta_k^{(\beta, n-1)}), \quad 1 \leq k \leq m.$$

Since all $\zeta_k^{(\beta, n-1)}$'s are positive and the derivative of $p_m^{(\beta, n-1)}$ changes sign at the zeros $\zeta_k^{(\beta, n-1)}$, it follows that between $\zeta_k^{(\beta, n-1)}$ and $\zeta_{k+1}^{(\beta, n-1)}$ there is exactly one zero of $p_m^{(k, n)}$, $1 \leq k \leq m-1$. The interlace is proved.

In the absence of the interpolation property it is perfectly possible for some zeros of p_m to reside outside the support of φ . For example, let us consider the sequence $c_k = (k+1)/(2k+1)$, $k \geq 0$. It is Stieltjes determinate by a theorem in [Brezinski, 1980], hence it corresponds to a distribution ψ . The corresponding biorthogonal polynomial with $\varphi(x, \mu) = \psi(x/\mu)$, $m=3$ and $\mu_1=1$, $\mu_2=2$, $\mu_3=3$ is $x^3 - \frac{40}{7}x^2 + \frac{66}{7}x - \frac{24}{7}$, with one positive and two complex conjugate zeros.

All distributions φ of the form $\psi(x/\mu)$ share the property that $I_k(\mu) = c_k \mu^k$, $k \geq 0$. In fact, they are also the only distributions that possess this property:

LEMMA 13. *Let $1 \in \Omega \subseteq (0, \infty)$ and without loss of generality $(a, b) = \mathbf{R}$. Suppose that $I_k(\mu) = c_k \mu^k$, $k \geq 0$, where $\{c_k\}$ is a given sequence of real numbers. Then $\{c_k\}$ is a Hamburger moment sequence and if it is determinate then $\varphi(x, \mu) = \psi(x/\mu)$, where ψ is the unique distribution with moments c_0, c_1, \dots .*

PROOF. Since $I_k(\mu)$ is the k th moment of a distribution it follows that $\{c_k \mu^k\}$ is a Hamburger moment sequence for all $\mu \in \Omega$, in particular for $\mu=1$. We expand

$$F(x, \mu) := \int_{-\infty}^{\infty} \frac{d\varphi(\tau, \mu)}{1 - \tau x}$$

in formal powers of x . This yields

$$F(x, \mu) = \sum_{k=0}^{\infty} c_k (\mu x)^k = F(\mu x, 1).$$

But

$$F(\mu x, 1) = \int_{-\infty}^{\infty} \frac{d\varphi(\tau, 1)}{1 - \mu \tau x} = \int_{-\infty}^{\infty} \frac{d\varphi(\tau/\mu, 1)}{1 - \tau x}.$$

Note that since the solution of the determinate Hamburger moment problem is unique [Akhiezer, 1965], it follows that

$$\varphi(x, \mu) = \varphi(x/\mu, 1) := \psi(x/\mu)$$

for all $x \in \mathbf{R}$, $\mu \in \Omega$. \square

EXAMPLE F. $\varphi(x, \mu) = \psi(x - \mu)$, $\Omega = (-\infty, \infty)$, $(a, b) = (-\infty, \infty)$.

Here ψ is an arbitrary distribution with moments c_0, c_1, \dots . Without loss of generality $c_0 = 1$. Regularity follows at once from Lemma 12:

$$I_k(\mu) = \sum_{l=0}^k \binom{k}{l} c_l \mu^{k-l} \in \pi_k[\mu], \quad k \geq 0,$$

$$D_m(\mu_1, \dots, \mu_m) = \prod_{1 \leq k < l \leq m} (\mu_l - \mu_k),$$

$$H_m(\mu) = q_m(\mu).$$

Some BOPS of this form satisfy the interpolation property, although this need be checked in every specific instance. E.g., let

$$\psi(x) = \int_0^x e^{-t^2} dt, \quad x \in \mathbf{R}.$$

Then $\omega(x, \mu) = e^{-(x-\mu)^2}$ and

$$E_m \begin{pmatrix} x_1, \dots, x_m \\ \mu_1, \dots, \mu_m \end{pmatrix} = \exp \left(- \sum_{k=1}^m x_k^2 - \sum_{l=1}^m \mu_l^2 \right) F_m \begin{pmatrix} e^{2x_1}, \dots, e^{2x_m} \\ \mu_1, \dots, \mu_m \end{pmatrix} \neq 0$$

for all distinct $x_1, \dots, x_m; \mu_1, \dots, \mu_m \in \mathbf{R}$, $m \geq 1$.

Many interesting distributions φ fall outside the scope of our three families. The present classification of BOPS is, at best, tentative. However, in general one has to exercise vigilance, since many fairly “innocent” distributions lack regularity.

EXAMPLE G. $d\varphi(x, \mu) = dx/(x^2 + \mu^2)$, $\Omega = (0, \infty)$, $(a, b) = (-1, 1)$.

It is straightforward that $I_{2k+1}(\mu) \equiv 0$ for all $k \geq 0$ and $\mu \in \Omega$. However, the interpolation property is, nonetheless, valid (if irrelevant): $\omega(x, \mu) = 1/(x^2 + \mu^2)$ leads to a Cauchy matrix E_m , which is nonsingular for all distinct $x_1, \dots, x_m; \mu_1, \dots, \mu_m$, $m \geq 1$.

To sum up our examples, they are neither an exhaustive list of BOPS—since this is, clearly, an impossible task—nor a list of all interesting instances of biorthogonality: it is much too early to anticipate what might be the applications of this construct. Nonetheless, they help to highlight the theory of this paper.

ACKNOWLEDGMENTS. The authors greatly benefited from stimulating discussion and correspondence with numerous individuals. R. Askey, M. G. de Bruin, S. W. Ellacott, W. B. Gragg, P. E. Koch, M. J. D. Powell, E. B. Saff, L. N. Trefethen and Y.-X. Yuan deserve particular mention and gratitude.

REFERENCES

- N. I. Akhiezer [1965], *The classical moment problem*, Hafner, New York.
- C. Brezinski [1980], *Padé-type approximation and general orthogonal polynomials*, Birkhäuser, Basel.
- T. S. Chichara [1978], *An introduction to orthogonal polynomials*, Gordon and Breach, New York.
- C. W. Cryer [1970], *Rodrigues' formula and the classical orthogonal polynomials*, Boll. Un. Mat. Ital. **3**, 1–11.
- R. T. Gregory and D. L. Karney [1969], *A collection of matrices for testing computational algorithms*, Wiley-Interscience, New York.
- P. R. Halmos [1950], *Measure theory*, Van Nostrand, Toronto.
- A. Iserles and S. P. Nørsett [1987a], *Two-step methods and bi-orthogonality*, Math. Comp. **49**, 543–552.
- , [1987b], *Bi-orthogonality and zeros of transformed polynomials*, J. Comput. Appl. Math. **19**, 39–45.
- A. Iserles and E. B. Saff [1987], *Biorthogonality in rational approximation*, J. Comput. Appl. Math. **19**, 47–54.
- S. Karlin and W. J. Studden [1966], *Tchebycheff systems: with applications in analysis and statistics*, Wiley-Interscience, New York.
- E. M. Nikišin [1982], *On simultaneous Padé approximants*, Math. USSR Sbornik **41**, 409–425.
- P. I. Pastro [1984], *Orthogonal polynomials and some q-beta integrals of Ramanujan*, Univ. of Wisconsin-Madison, Tech. Rep.
- M. J. D. Powell [1981], *Approximation theory and methods*, Cambridge Univ. Press, Cambridge.
- E. D. Rainville [1967], *Special functions*, Macmillan, New York.
- G. Szegő [1982], *Collected papers*, Birkhäuser, Boston, Mass.

DEPARTMENT OF APPLIED MATHEMATICS AND THEORETICAL PHYSICS, UNIVERSITY OF CAMBRIDGE, CAMBRIDGE CB3 9EW, ENGLAND

INSTITUTT FOR NUMERISK MATEMATIKK, NORGES TEKNISKE HØGSKOLE, 7034 TRONDHEIM-NTH, NORWAY