ANZIAM J. 50(2009), 435–444
doi:10.1017/S1446181109000169

ORTHOGONAL FUNCTIONS AND ZERNIKE POLYNOMIALS—A RANDOM VARIABLE INTERPRETATION

C. S. WITHERS[1]

(Received 23 January, 2009; revised 8 April, 2009)

Abstract

There are advantages in viewing orthogonal functions as functions generated by a random variable from a basis set of functions. Let Y be a random variable distributed uniformly on [0, 1]. We give two ways of generating the Zernike radial polynomials with parameter l, $\{Z^l_{l+2n}(x),\ n \geq 0\}$. The first is using the standard basis $\{x^n, n \geq 0\}$ and the random variable $Y^{1/(l+1)}$. The second is using the nonstandard basis $\{x^{l+2n}, n \geq 0\}$ and the random variable $Y^{1/2}$. Zernike polynomials are important in the removal of lens aberrations, in characterizing video images with a small number of numbers, and in automatic aircraft identification.

2000 *Mathematics subject classification*: primary 42A16; secondary 45B05.
Keywords and phrases: Zernike polynomials, orthogonal functions.

1. Introduction and summary

Orthogonal functions arise naturally as eigenfunctions of Fredholm integral equations and Green's functions of differential equations. Among their many uses, they give a generalized Fourier expansion of a function. Section 2 gives a brief presentation of the theory of orthogonal functions and Fourier series in terms of random variables. The idea of expressing orthogonal function theory in terms of random variables was introduced in [13]. This also shows some of the advantages of this point of view.

Section 3 gives a Gram–Schmidt procedure for obtaining real functions (not necessarily polynomials) that are orthogonal with respect to any real random variable X in the sense that

$$E f_m(X) f_n(X) = 0 \quad \text{for } m \neq n, \tag{1.1}$$

starting from any given set of real functions $\{g_n(x), n \geq 0\}$. We shall say that functions $\{f_n(x), n \geq 0\}$ satisfying (1.1) are *X-orthogonal.* The *basis functions*

[1]Applied Mathematics Group, Industrial Research Ltd, Box 31-310, Lower Hutt, New Zealand; e-mail: c.withers@irl.cri.nz.

$\{g_n(x)\}$ generally have a simple form. For example for orthogonal *polynomials*, one takes $g_n(x) = x^n$ for $n \geq 0$, *or* $g_0(x) = 1$, $g_n(x) = (x - E\ X)^n$ for $n \geq 1$. We call these the *noncentral* and *central* basis functions, respectively. They generate the same X-orthogonal polynomials. However, using the central basis gives $f_n(x)$ in simpler form in terms of the central moments of X, rather than in terms of its noncentral moments.

This leads to several distinct Gram–Schmidt procedures for obtaining the Zernike radial polynomials.

Section 4 gives a new derivation of the explicit formulas for the Zernike radial polynomials. These formulas were derived in [2] from the Jacobi polynomials. However, our derivation is given in terms of random variables. An explicit expression for these polynomials, as well as their geometric generating function, is given in [3].

When observations are made on an annulus, say $\beta \leq r \leq 1$, rather than on the circle $r \leq 1$, some of the results on Zernike polynomials have been extended by [8]. In optics, Zernike polynomials are used in the analysis of interferogram fringes and for minimizing aberrations when grinding lenses—see, for example, [5].

A number of papers have given methods for obtaining the Fourier coefficients of an arbitrary function on the circle using the Zernike polynomials. Teague [7] calls these coefficients the Zernike moments, and gives their relation to the ordinary moments of the function. He shows that under rotation by an angle θ, the Fourier coefficients are the same except for a factor $e^{il\theta}$, where the index l is given in Section 4. He also gives their leading invariants. His applications include characterizing video images with a small number of numbers, say 20, automatic aircraft identification, and automatic identification of biological patterns. Prata and Rusch [6] compare the matrix inversion method and the integration method for obtaining the Fourier coefficients (which they call the Zernike polynomials expansion coefficients), giving an integration algorithm based on their new expressions for the integral of a product of a power and a Zernike polynomial. They also give the interesting recurrence formulas (5) and (6) for the Zernike radial polynomials. For other references, see [9] and Wikipedia.

2. Fourier series

Let μ be a finite measure on a set Ω. Typically $\Omega \subset R^q$ for some q, but much more general domains are possible.

Let $L_2(p, \mu)$ be the space of functions $g : \Omega \to C^p$ such that $\int |g|^2\, d\mu < \infty$ where $\int |g|^2\, d\mu = \int_\Omega |g(x)|^2\, d\mu(x)$. Let $\{f_n,\ n = 0, 1, \ldots\}$ be a complete set of functions in $L_2(p, \mu)$. This holds, for example, if

$$\int f_n g\, d\mu = 0, \quad n = 0, 1, \ldots \quad \Rightarrow \quad g = 0 \text{ almost everywhere } \mu$$

or if Parseval's identity holds. Then any function $g \in L_2(p, \mu)$ can be written as a

https://doi.org/10.1017/S1446181109000169 Published online by Cambridge University Press

linear combination of $\{f_n\}$,

$$g(x) = \sum_{n=0}^{\infty} c_n(g) f_n(x), \tag{2.1}$$

where convergence is in $L_2(p, \mu)$,

$$\int |g - g_N|^2 \, d\mu \to 0 \text{ as } N \to \infty \quad \text{where } g_N(x) = \sum_{n=0}^{N} c_n(g) f_n(x).$$

Convergence holds more strongly under extra conditions. For example [4, Theorem 9.1] gives conditions for pointwise convergence of g_N to g. Now suppose that $\{f_n\}$ are orthonormal with respect to μ:

$$\int f_m^* f_n \, d\mu = \int_\Omega f_m(x)^* f_n(x) \, d\mu(x) = \delta_{mn}, \tag{2.2}$$

where f^* is the transpose of the complex conjugate of f and $\delta_{mn} = 1$ or 0 for $m = n$ or $m \neq n$. Then the Fourier coefficient $c_n(g)$ in (2.1) is given simply by

$$c_n(g) = \int f_n^* g \, d\mu,$$

and Parseval's identity states that

$$\int |g|^2 \, d\mu \equiv \sum_{n=0}^{\infty} |c_n(g)|^2.$$

If we require only that $\{f_n\}$ be orthogonal rather than orthonormal, (2.1) holds with

$$c_n(g) = \frac{\int f_n^* g \, d\mu}{\int |f_n|^2 \, d\mu},$$

and Parseval's identity becomes

$$\int |g|^2 \, d\mu \equiv \sum_{n=0}^{\infty} \frac{|\int f_n^* g \, d\mu|^2}{\int |f_n|^2 \, d\mu}. \tag{2.3}$$

Orthonormal functions arise naturally as eigenfunctions of symmetric functions $K(x, y)$ on Ω^2 satisfying

$$\iint K(x, y)^2 \, d\mu(x) \, d\mu(y) < \infty;$$

see, for example, Withers [10–12].

2.1. Orthogonal functions in terms of a random variable The standardized measure

$$d\mu_0(x) = \frac{d\mu(x)}{\mu(\Omega)}$$

has total measure 1, so it is a probability measure. So if $\Omega \subset R^q$, then one can construct a random variable X with domain Ω such that

$$P_X(x) = \mu_0((-\infty, x]) \quad \text{where we shall write} \quad P_X(x) = \text{Probability}(X \leq x).$$

Set $f_{n0}(x) = f_n(x)\mu(\Omega)^{1/2}$, $n \geq 0$. If $\{f_n\}$ are orthonormal with respect to μ, then $\{f_{n0}\}$ are X-*orthonormal*:

$$Ef_{m0}(X)^* f_{n0}(X) = \int f_{m0}^* f_{n0}\, d\mu_0 = \int f_m^* f_n\, d\mu = \delta_{mn}.$$

So without loss of generality we shall assume that $\mu(\Omega) = 1$, so that (2.2) and (2.3) can be written

$$Ef_m(X)^* f_n(X) = \delta_{mn}, \quad E|g(X)|^2 = \sum_{n=0}^{\infty} \frac{|Ef_n(X)^* g(X)|^2}{|Ef_n(X)^2|}.$$

If $f_0(x) = 1$, the orthogonality condition $\int f_m^* f_n\, d\mu = 0$ for $m \neq n$ can now be written

$$Ef_n(X) = 0 \quad \text{for } n > 0,$$
$$\text{cov}(f_m(X), f_n(X)) = 0 \quad \text{for } m \neq n,$$

where

$$\text{cov}(X, Y) = E(X - EX)(Y - EY);$$

that is, $\{f_n(X)\}$ is an infinite set of *uncorrelated* random variables with means all zero, except for $f_0(X) = 1$. This gives a very nice interpretation of orthogonality in terms of the random variable X.

In the rest of this paper we shall assume that $p = q = 1$ and that the functions we are dealing with are real.

3. Gram–Schmidt procedures

Let X be a random variable in a set Ω, typically the real line. Let $\{g_n(x), n \geq 0\}$ be a set of real functions defined on Ω, satisfying $Eg_n(X)^2 < \infty$.

X-orthogonal functions, $\{f_n(x), n \geq 0\}$, are often generated from a simpler set of basis functions $\{g_n, n = 0, 1, \ldots\}$ by a Gram–Schmidt orthogonalization procedure. These are sequential methods starting with $f_0(x) = g_0(x)$. For orthogonal *polynomials*, the usual choice is the noncentral basis $g_n(x) = x^n$. To make these polynomials X-orthonormal, one merely scales them by setting

$$f_{n0}(x) = \frac{f_n(x)}{h_n^{1/2}} \quad \text{where } h_n = Ef_n(X)^2 \text{ so that } Ef_{n0}(X)^2 = 1.$$

https://doi.org/10.1017/S1446181109000169 Published online by Cambridge University Press

We shall assume that all the bivariate moments

$$M_{kn} = E g_k(X) g_n(X)$$

are finite. This condition is needed so that the coefficients of $\{f_n(x), n \geq 0\}$ are finite.

Gram–Schmidt orthogonalization can be achieved in number of ways. Here we modify the third way given in [13]. We want to construct orthogonal f_n of the form

$$f_n(x) = \sum_{j=0}^{n} a_{nj} g_j(x) \quad \text{where } a_{nn} = 1.$$

This is equivalent to the form

$$f_n(x) = g_n(x) + \sum_{j=0}^{n-1} b_{nj} f_j(x).$$

So for $0 \leq j < n$, $0 = E f_j(X) f_n(X) = A_{jn} + b_{nj} h_j$ where $A_{jn} = E f_j(X) g_n(X)$, giving $b_{nj} = -A_{jn}/h_j$. So

$$f_0(x) = g_0(x), \qquad f_n(x) = g_n(x) - \sum_{j=0}^{n-1} \frac{f_j(x) A_{jn}}{h_j}, \qquad n \geq 1$$

where A_{jn} is given sequentially in terms of the moments by

$$A_{0n} = M_{0n}, \qquad A_{kn} = M_{kn} - \sum_{j=0}^{k-1} \frac{A_{jk} A_{jn}}{h_j}, \qquad 1 \leq k < n. \tag{3.1}$$

On substitution one obtains

$$h_n = A_{nn} = M_{nn} - \sum_{j=0}^{n-1} \frac{A_{jn}^2}{A_{jj}},$$

so that (3.1) also holds for $k = n$. For example $h_0 = A_{00} = M_{00}$,

$$A_{1n} = M_{1n} - \frac{M_{01} M_{0n}}{M_{00}}, \qquad n \geq 1,$$
$$A_{2n} = M_{2n} - \frac{M_{02} M_{0n}}{M_{00}} - \frac{A_{12} A_{1n}}{A_{11}}, \qquad n \geq 2,$$

and explicit formulas for $\{f_n,\ 0 \leq n \leq 3\}$ are $f_0(x) = g_0(x)$,

$$f_1(x) = g_1(x) - \frac{f_0(x) M_{01}}{M_{00}}, \qquad f_2(x) = g_2(x) - \frac{f_0(x) M_{02}}{M_{00}} - \frac{f_1(x) A_{12}}{A_{11}},$$
$$f_3(x) = g_3(x) - \frac{f_0(x) M_{03}}{M_{00}} - \frac{f_1(x) A_{13}}{A_{11}} - \frac{f_2(x) A_{23}}{A_{22}},$$

https://doi.org/10.1017/S1446181109000169 Published online by Cambridge University Press

where

$$A_{11} = M_{11} - \frac{M_{01}^2}{M_{00}}, \quad A_{12} = M_{12} - \frac{M_{01}M_{02}}{M_{00}}, \quad A_{13} = M_{13} - \frac{M_{01}M_{03}}{M_{00}},$$
$$A_{22} = M_{22} - \frac{M_{02}^2}{M_{00}}, \quad A_{23} = M_{23} - \frac{M_{02}M_{03}}{M_{00}} - \frac{A_{12}A_{13}}{A_{11}}.$$

The moments needed for f_n are $\{M_{kj},\ 0 \le k < n,\ k \le j \le n\}$.

EXAMPLE 1. Suppose that

$$g_n(x) = x^{l+2n} \quad \text{on } \Omega = [0, 1]$$

for some constant l. Then

$$M_{kn} = m_{2l+2k+2n} \quad \text{where } m_j = EX^j.$$

If $P_X(x) = x^2$ on $\Omega = [0, 1]$, that is, if $X = U^{1/2}$ where U is distributed uniformly on $[0, 1]$, then $m_j = 2/(j + 2)$.

We shall see in Section 4 that the corresponding X-orthogonal $f_n(x)$ are just the Zernike radial polynomials scaled so that their leading coefficient is 1.

A disadvantage of the method of Example 1 is its use of a nonstandard basis.

EXAMPLE 2. Consider the usual noncentral basis $g_n(x) = x^n$. Then

$$M_{kn} = m_{k+n} \quad \text{where } m_j = EX^j.$$

So the first $2n - 1$ moments of X, $\{m_j = EX^j,\ \ 1 \le j < 2n\}$, are needed to compute $f_n(x)$:

$$f_0(x) = 1, \quad f_n(x) = x^n - \sum_{j=0}^{n-1} \frac{f_j(x)A_{jn}}{A_{jj}}, \quad n \ge 1,$$

where A_{jn} are given sequentially by

$$A_{0n} = m_n, \quad A_{jn} = m_{n+j} - \sum_{k=0}^{j-1} \frac{A_{kj}A_{kn}}{A_{kk}}, \quad 0 < j < n.$$

If one uses instead the central basis, one obtains f_n in terms of the central moments $\nu_j = E(X - EX)^j$ and central versions of A_{jn}. By [13] the first five X-orthogonal polynomials are given in terms of $y = x - EX$ by

$$f_1(x) = y, \quad f_2(x) = y^2 - \nu_2 - \frac{y\nu_3}{\nu_2}, \quad f_3(x) = y^3 - \nu_3 - \frac{y\nu_4}{\nu_2} - \frac{f_2(x)A_{23}}{A_{22}},$$

https://doi.org/10.1017/S1446181109000169 Published online by Cambridge University Press

where

$$A_{22} = \nu_4 - \nu_2^2 - \frac{\nu_3^2}{\nu_2}, \quad A_{23} = \nu_5 - \nu_2\nu_3 - \frac{\nu_3\nu_4}{\nu_2},$$
$$f_4(x) = y^4 - \nu_4 - \frac{y\nu_5}{\nu_2} - \frac{f_2(x)A_{24}}{A_{22}} - \frac{f_3(x)A_{34}}{A_{33}},$$

in which

$$A_{24} = \nu_6 - \nu_2\nu_4 - \frac{\nu_3\nu_5}{\nu_2}, \quad A_{33} = \nu_6 - \nu_3^2 - \frac{\nu_4^2}{\nu_2} - \frac{A_{23}^2}{A_{22}},$$
$$A_{34} = \nu_7 - \nu_3\nu_4 - \frac{\nu_4\nu_5}{\nu_2} - \frac{A_{23}A_{24}}{A_{22}},$$

and

$$f_5(x) = y^5 - \frac{\nu_5 - y\nu_6}{\nu_2} - \frac{f_2(x)A_{25}}{A_{22}} - \frac{f_3(x)A_{35}}{A_{33}} - \frac{f_4(x)A_{45}}{A_{44}},$$

where

$$A_{25} = \nu_7 - \nu_2\nu_5 - \frac{\nu_3\nu_6}{\nu_2}, \quad A_{35} = \nu_8 - \nu_3\nu_5 - \frac{\nu_4\nu_6}{\nu_2} - \frac{A_{23}A_{25}}{A_{22}},$$
$$A_{44} = \nu_8 - \nu_4^2 - \frac{\nu_5^2}{\nu_2} - \frac{A_{24}^2}{A_{22}} - \frac{A_{34}^2}{A_{33}},$$
$$A_{45} = \nu_9 - \nu_4\nu_5 - \frac{\nu_5\nu_6}{\nu_2} - \frac{A_{24}A_{25}}{A_{22}} - \frac{A_{34}A_{35}}{A_{33}}.$$

Now suppose that $P_X(x) = x^{l+1}$ on $\Omega = [0, 1]$, where $l + 1 > 0$ so that $m_j = (l + 1)/(l + 1 + j)$. We shall see in the next section that it is precisely this example (with $l = 0, 1, \ldots$), that generates the Zernike radial polynomials—but this time after a transformation.

NOTE 3.1. If Y is a one-to-one transformation of X, say $Y = h(X) : \Omega \to \Omega'$ and $y = h(x)$, then $P_Y(y) = P_X(x)$, that is, $P(Y \le y) = P(X \le x)$. For example, if X has a density $p_X(x)$ with respect to Lebesgue measure, then so does Y and it is given by

$$p_Y(y) = p_X(x)\frac{dx}{dy}, \quad y \in \Omega'.$$

The Y-orthogonal functions are $f_{nY}(y) = f_n(x)$ since $E\, f_{mY}(Y) f_{nY}(Y) = E f_m(X) f_n(X)$. If $\{f_n\}$ are generated by $\{g_n\}$, then $\{f_{nY}\}$ are generated by $g_{nh}(y) = g_n(x)$.

NOTE 3.2. Suppose that $a(x)$ is a function satisfying $Ea(X)^2 = 1$. Then $f_{nZ}(x) = f_n(x)/a(x)$ are the Z-orthogonal functions of a random variable Z with density

$$dP_Z(x) = a(x)^2\, dP_X(x)$$

https://doi.org/10.1017/S1446181109000169 Published online by Cambridge University Press

since

$$Ef_m(X)f_n(X) = Ea(X)^2 f_{mZ}(X) f_{nZ}(X) = Ef_{mZ}(Z) f_{nZ}(Z).$$

If $\{f_n\}$ are generated by orthogonalizing $\{g_n\}$, then $\{f_{nZ}\}$ can be generated by orthogonalizing $\{\tilde{g}_{na}(x) = g_n(x)/a(x)\}$.

4. Zernike polynomials

The Zernike circular polynomials $Z_n^l(x, \theta)$ are orthogonal over the circle with unit radius:

$$\int_0^1 \int_0^{2\pi} Z_n^l(x, \theta)\, Z_{n'}^{l'}(x, \theta)\, dP_X(x)\, dP_\Theta(\theta) = \frac{\delta_{nn'}\delta_{ll'}}{n+1} \quad \text{for } n \geq |l|,$$

where $P_X(x) = x^2$ on $[0, 1]$, $P_\Theta(\theta) = \theta/(2\pi)$ on $[0, 2\pi]$.

This can be written

$$EZ_n^l(X, \Theta)\, Z_{n'}^{l'}(X, \Theta) = \frac{\delta_{nn'}\delta_{ll'}}{n+1} \quad \text{for } n \geq |l|,$$

where (X, Θ) are independent random variables with distributions $P_X(x)$, $P_\Theta(\theta)$. So Θ has the uniform distribution on $[0, 2\pi]$ and $X = Y^{1/2}$ where Y has the uniform distribution on $\Omega = [0, 1]$:

$$x^2 = P(X \leq x) = P(Y \leq y) = y \quad \text{for } y = x^2 = h(x), \text{ say}.$$

Because X and Θ are independent, $Z_n^l(x)$ factorizes into the product of two functions, each orthogonal with respect to one of these random variables:

$$Z_n^l(x) = R_n^{|l|}(x) e^{il\theta}, \quad l = \ldots, -1, 0, 1, \ldots,$$

where $e^{il\theta}$ are orthogonal with respect to Θ, and $R_n^l(x)$, known as the *Zernike radial polynomials*, are orthogonal with respect to X:

$$Ee^{-il\Theta} e^{il'\Theta} = \delta_{ll'}, \quad ER_n^l(X) R_{n'}^l(X) = \frac{\delta_{nn'}}{n+1}.$$

As noted in [3, Equation (2.8)],

$$f_n(x) = R_{l+2n}^l(x), \quad n \geq 0,\ l \geq 0,$$

are generated by the basis $g_n(x) = x^{l+2n}$, not the basis $g_n(x) = x^n$, and so are proportional to the polynomials generated in Example 1.

We now show how to use Notes 3.1 and 3.2 to put $f_n(x)$ in the context of the previous section. By Note 3.1, $f_n(x)$ is generated by

$$g_n(x) = cx^{l+2n} = cy^{l/2+n} = g_{nh}(y) \quad \text{for } x = y^{1/2},$$

where $c \neq 0$ is any constant. Set $a(y) = cy^{l/2}$ and choose c so that

$$1 = Ea(Y)^2 = \frac{c^2}{l+1} \quad \text{that is,} \quad c = (l+1)^{1/2}.$$

Then by Note 3.2, $\{\tilde{g}_{na}(y) = g_{nh}(y)/a(y) = y^n\}$ generate polynomials $\{f_{nZ}(y) = f_{nY}(y)/a(y)\}$ on $[0, 1]$ that are orthogonal with respect to Z where $dP_Z(y) = a(y)^2\, dP_Y(y)$, that is, $P_Z(y) = y^{l+1}$:

$$Ef_{mZ}(Z)f_{nZ}(Z) = Ef_{mY}(Y)f_{nY}(Y) = Ef_m(X)f_n(X) = \frac{\delta_{mn}}{n+l+1}.$$

But we know the orthogonal polynomials generated by $\{y^n\}$ and the beta random variable $B_{\beta\alpha}$ with density $y^\beta(1-y)^\alpha/B(\beta+1, \alpha+1)$, since $A = 2B_{\beta\alpha} - 1$ has density

$$p_A(x) = \frac{(1-x)^\alpha(1+x)^\beta}{2^{\alpha+\beta+1}B(\beta+1, \alpha+1)}$$

on $[-1, 1]$, with orthogonal polynomials the Jacobi polynomials if standardized so that $f_n(1) = \binom{n+\alpha}{n}$; see [1] and [13, Example 4.1]. Equivalently, the Jacobi polynomials on $[0, 1]$ are generated by $\{y^{n+\beta}(1-y)^\alpha,\ n \geq 0\}$. So taking $\alpha = l$, $\beta = 0$, by the uniqueness up to constant multiplier of the orthogonal polynomials generated by a given random variable, we conclude that $f_{nZ}(y)$ is a constant multiple of the *shifted Jacobi polynomial* with $\alpha = l$, $\beta = 0$, that is, the polynomial generated by B_{0l}. So from the formula for the Jacobi polynomial, one obtains

$$f_n(x) = R^l_{l+2n}(x) = \sum_{s=0}^{n} (-1)^s (2n+l-s)! \frac{x^{l+2n-2s}}{s!(l+n-s)!(n-s)!}, \qquad n \geq 0,$$

that is,

$$n!f_n(x) = n!R^l_{l+2n}(x) = \sum_{s=0}^{n} (-1)^s \binom{n}{s} [l+2n-s]_s x^{l+2n-2s}, \tag{4.1}$$

where

$$[a]_s = a(a-1)\cdots(a-s+1) = \frac{a!}{(a-s)!}.$$

(Some references give this as the *definition* of the Zernike radial polynomials.) This is a version of the derivation in [2], but it is given here in terms of random variables. Also our (4.1) is simpler than the previous formula. Note that

$$f_{nZ}(y) = (l+1)^{-1/2} f'_{nZ}(y), \quad \text{where } f'_{nZ}(y) = \sum_{s=0}^{n} (-1)^s \binom{n}{s} [l+2n-s]_s y^{n-s},$$

has leading coefficient 1. So f'_{nZ} is just the polynomial generated by Example 1.

https://doi.org/10.1017/S1446181109000169 Published online by Cambridge University Press

Acknowledgements

I wish to thank the referees for carefully reading the manuscript and suggesting improvements. I also wish to thank Warwick Kissling for his references and encouragement, and sharing his practical applications of Zernike polynomials in the field of optics.

References

[1] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions*, Volume 55 of *Applied Mathematics Series* (US Department of Commerce, National Bureau of Standards, Washington, DC 1964).

[2] A. B. Bhatia and E. Wolf, "The Zernike circle polynomials occurring in diffraction theory", *Proc. Phys. Soc. B* **65** (1952) 909–910.

[3] A. B. Bhatia and E. Wolf, "On the circle polynomials of Zernike and related orthogonal sets", *Proc. Cambridge Philos. Soc.* **50** (1954) 40–48.

[4] Y. L. Geronimus, *Polynomials orthogonal on a circle and interval* (Pergamon Press, Oxford, 1960).

[5] D. Malacara and S. L. DeVore, "Interfertogram evaluation and wavefront fitting", in: *Optical shop testing* (ed. D. Malacara), (Wiley, New York, 1992) 455–499.

[6] A. Prata and W. V. T. Rusch, "Algorithm for computation of Zernike polynomials expansion coefficients", *Appl. Optim.* **28** (1989) 749–754.

[7] M. R. Teague, "Image analysis via the general theory of moments", *J. Opt. Soc. Amer.* **70** (1980) 920–930.

[8] J. Y. Wang and D. E. Silva, "Wave-front interpretation with Zernike polynomials", *Appl. Optics* **19** (1980) 1510–1518.

[9] E. Weisstein, "Zernike polynomial", http://mathworld.wolfgram/ZernikePolynomial.html (1999).

[10] C. S. Withers, "Mercer's theorem and Fredholm resolvents", *Bull. Aust. Math. Soc.* **11** (1974) 373–380.

[11] C. S. Withers, "Fredholm theory for arbitrary measure spaces", *Bull. Aust. Math. Soc.* **12** (1975) 283–292.

[12] C. S. Withers, "Fredholm equations have uniformly convergent solutions", *J. Math. Anal. Appl.* **64** (1978) 602–609.

[13] C. S. Withers and S. Nadarajah, "Orthogonal polynomials via random variables", to appear in *Utilitas Mathematica*.