Robust and fast computation for the polynomials of optics

G. W. Forbes*

QED Technologies Inc., 1040 University Ave., Rochester NY [illegible] USA
**forbes@qedmrf.com*

Abstract: Mathematical methods that are poorly known in the field of optics are adapted and shown to have striking significance. Orthogonal polynomials are common tools in physics and optics, but problems are encountered when they are used to higher orders. Applications to arbitrarily high orders are shown to be enabled by remarkably simple and robust algorithms that are derived from well known recurrence relations. Such methods are demonstrated for a couple of familiar optical applications where, just as in other areas, there is a clear trend to higher orders.

OCIS codes: (000.3860) Mathematical methods in physics; (220.0220) Optical design and fabrication; (260.1960) Diffraction theory; (220.1250) Aspherics.

References and links

1. M. Born, and E. Wolf, *Principles of Optics* (Cambridge, 1999), see Sec. 9.2 and Appendix VII.
2. A. E. Siegman, *Lasers* (University Science Books, 1986), Chaps. 16–17.
3. M. Abramowitz, and I. Stegun, *Handbook of Mathematical Functions* (Dover, 1978), Chap. 22.
4. A. B. Bhatia, E. Wolf, and M. Born, “On the circle polynomials of Zernike and related orthogonal sets,” Proc. Camb. Philos. Soc. **50**(01), 40–48 (1954).
5. D. R. Myrick, “A generalization of the radial polynomials of F. Zernike,” J. Soc. Ind. Appl. Math. **14**(3), 476–492 (1966).
6. E. C. Kintner, “On the mathematical properties of the Zernike polynomials,” Opt. Acta (Lond.) **23**, 679–680 (1976).
7. R. Barakat, “Optimum balanced wave-front aberrations for radially symmetric amplitude distributions: generalizations of Zernike polynomials,” J. Opt. Soc. Am. **70**(6), 739–742 (1980).
8. C.-W. Chong, P. Raveendran, and R. Mukundan, “A comparative analysis of algorithms for fast computation of Zernike moments,” Pattern Recognit. **36**(3), 731–742 (2003).
9. W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, 1992) Section 5.5.
10. C. W. Clenshaw, “A Note on the Summation of Chebyshev Series,” Math. Tables Other Aids Comput. **9**, 118–120 (1955), http://www.jstor.org/stable/2002068.
11. F. J. Smith, “An Algorithm for Summing Orthogonal Polynomial Series and their Derivatives with Applications to Curve-Fitting and Interpolation,” Math. Comput. **19**(89), 33–36 (1965).
12. H. E. Salzer, “A Recurrence Scheme for Converting from One Orthogonal Expansion into Another,” Commun. ACM **16**(11), 705–707 (1973).
13. E. H. Doha, “On the coefficients of differentiated expansions and derivatives of Jacobi polynomials,” J. Phys. Math. Gen. **35**(15), 3467–3478 (2002).
14. R. Barrio, and J. M. Peña, “Numerical evaluation of the p’th derivative of Jacobi series,” Appl. Numer. Math. **43**(4), 335–357 (2002).
15. B. Y. Ting, and Y. L. Luke, “Conversion of Polynomials between Different Polynomial Bases,” IMA J. Numer. Anal. **1**(2), 229–234 (1981).
16. K. A. Goldberg, and K. Geary, “Wave-front measurement errors from restricted concentric subdomains,” J. Opt. Soc. Am. A **18**(9), 2146–2152 (2001).
17. J. Schwiegerling, “Scaling Zernike expansion coefficients to different pupil sizes,” J. Opt. Soc. Am. A **19**(10), 1937–1945 (2002).
18. C. E. Campbell, “Matrix method to find a new set of Zernike coefficients from an original set when the aperture radius is changed,” J. Opt. Soc. Am. A **20**(2), 209–217 (2003).
19. G. M. Dai, “Scaling Zernike expansion coefficients to smaller pupil sizes: a simpler formula,” J. Opt. Soc. Am. A **23**(3), 539–543 (2006).
20. H. Shu, L. Luo, G. Han, and J.-L. Coatrieux, “General method to derive the relationship between two sets of Zernike coefficients corresponding to different aperture sizes,” J. Opt. Soc. Am. A **23**(8), 1960–1966 (2006).
21. A. J. E. M. Janssen, and P. Dirksen, “Concise formula for the Zernike coefficients of scaled pupils,” J. Microlith. Microfab Microsyst. **5**(3), 030501–3 (2006).

#128403 - $15.00 USD Received 12 May 2010; revised 2 Jun 2010; accepted 3 Jun 2010; published 11 Jun 2010

22. G. W. Forbes, “Shape specification for axially symmetric optical surfaces,” Opt. Express **15**(8), 5218–5226 (2007), http://www.opticsinfobase.org/oe/abstract.cfm?URI=oe-15-8-5218.

1. Introduction

Many of the special functions used in optics satisfy three-term recurrence relations. These include trigonometric functions, Bessel functions (including modified Bessels and Hankels), and various orthogonal polynomials. Although most of the ideas discussed in what follows apply for applications of any in this class of functions, I concentrate on orthogonal polynomials in particular. Among the better known orthogonal polynomials in optics are the Zernike polynomials [1] (for characterizing a function defined on a disc, or modified for cases with non-uniform weighting, or for annular domains) and the Laguerre and Hermite polynomials [2] (for characterizing profiles of transversely localized beams). Traditionally, such polynomials were only ever needed to modest orders, say up to ten or so terms. (Note that, in the case of Zernikes, there are distinct polynomials for each azimuthal order, so the total number of terms is often more than thirty, but these hold a more modest number of terms for each azimuthal order.) This has meant that explicit expressions have been widely used for the polynomials, and they have generally been perfectly adequate.

The inclusion of higher-order terms in such decompositions is increasingly being used to boost capabilities. Two considerations limit progress in this direction, however. First, the explicit expressions for the polynomials lead to numerical round-off problems that become catastrophic when the number of terms reaches around 20 —a lesson sometimes learned the hard way. Second, computational efficiency (i.e. arithmetic operation count) remains a vital consideration in applications such as optical design: despite the staggering growth of computer power, multi-dimensional global optimization is so challenging that its results will always benefit from efficiency gains. The methods discussed below lead to simple code that is remarkably efficient while also avoiding round-off failures.

Any set of orthogonal polynomials, say $\{P_m\}$ where $P_m(x)$ is a polynomial of order m, is defined so that

$$\langle P_m, P_n \rangle = 0, \qquad \text{when}\, m \neq n, \tag{1.1}$$

where $<f, g>$ is typically an integral over the domain of interest of the product of $f(x)$ and $g(x)$. When the basis is normalized by $<P_n, P_n> = 1$, the coefficients in an expansion of the form

$$S(x) = \sum_m s_m\, P_m(x) \tag{1.2}$$

are then determined individually by noticing that $<S, P_n> = \sum_m < s_m P_m, P_n> = s_n$, i.e.

$$s_m = \langle S, P_m \rangle. \tag{1.3}$$

These coefficients are evidently like a spectrum, where the analogue of Parseval’s theorem follows similarly from Eqs. (1.1) and (1.2):

$$\langle S, S \rangle = \sum_m s_m^{\,2}. \tag{1.4}$$

This is an intuitive conservation law that couples the total energy in the spectrum to the energy in the original function.

A result that is of central importance in this work follows upon considering the expansion of the function defined by $S(x) := x\,P_n(x)$. On account of Eq. (1.1), $<\phi, P_m>$ vanishes when ϕ is *any* polynomial of order less than m. It follows that $s_m = <x\,P_n, P_m> = <P_n, x\,P_m>$ vanishes unless $|m-n| \leq 1$. This observation can be expressed as

#128403 - $15.00 USD Received 12 May 2010; revised 2 Jun 2010; accepted 3 Jun 2010; published 11 Jun 2010

$$x\,P_n(x) \;=\; q\,P_{n+1}(x) + r\,P_n(x) + s\,P_{n-1}(x), \tag{1.5}$$

where q, r, and s, are constants and $q \neq 0$. Equation (1.5) can be re-arranged to give the standard form for the three-term recurrence relation that underpins the following sections:

$$P_{n+1}(x) \;=\; (a_n + b_n\,x)\,P_n(x) - c_n\,P_{n-1}(x). \tag{1.6}$$

Such a relation follows regardless of the normalization. In fact, in place of $< P_n, P_n > = 1$, normalization by peak value is the usual convention for the applications considered in this work. Simple closed forms for a_n, b_n, and c_n are known for all the most commonly used orthogonal polynomials, e.g. see Sec. 22.7 of [3].

Since the well known orthogonal polynomials in optics are among the standard set, the existence of the associated recurrence relations, generating functions, Christoffel-Darboux identities, etc. has often been noted and occasionally rediscovered. In the case of Zernikes, see [4–7] as examples. Some of the associated recurrence relations have been adopted in the field of image processing, as reviewed in [8]. Even so, in most applications of Zernikes, significant effort —and space— is devoted to working with the explicit expressions for the polynomials of moderate orders. *It has not been generally appreciated that, in practice, this is a road to grief.* With r as the radial coordinate on a unit disc, each element of a Zernike decomposition involves $\cos m\theta$ or $\sin m\theta$ times $r^m\,Z_n^m(r^2)$. These polynomials for m'th azimuthal order are related to the Jacobi polynomials by $Z_n^m(x) = P_n^{(0,m)}(2x-1)$, e.g. see Eq. (3.7) of [4]. The domain of interest is $0 \le x \le 1$. The explicit expression that is the focus of many applications in such work is

$$Z_n^m(x) \;=\; \sum_{j=0}^{n} (-1)^{n-j} \binom{n}{j} \binom{m+n+j}{n} x^j \;=\; \sum_{j=0}^{n} \frac{(-1)^{n-j}(m+n+j)!}{j!\,(m+j)!\,(n-j)!}\, x^j\,. \tag{1.7}$$

(This is given by, for example, 22.3.3 and 22.5.2 of [3]). From among the rotationally symmetric subset, as an example, it follows that

$$\begin{aligned} Z_{10}^0(x) \;=\; & 1 - 110x + 2{,}970x^2 - 34{,}320x^3 + 210{,}210x^4 - 756{,}756x^5 \\ & + 1{,}681{,}680x^6 - 2{,}333{,}760x^7 + 1{,}969{,}110x^8 - 923{,}780x^9 + 184{,}756x^{10}. \end{aligned} \tag{1.8}$$

Since $|Z_n^m(1)| = 1$, six or seven significant digits are evidently lost through heavy cancellation when evaluating Eq. (1.8) for any $x \approx 1$. More generally, the magnitude of the coefficients of $Z_n^m(x)$ is found to be about $10^{0.7n}$. This means that, even when using double precision arithmetic, it is impossible to guarantee any accurate decimal places at all in $Z_n^m(1.0)$ once n exceeds about 20. Accuracy can be seen to suffer long before that point of total catastrophe and, as demonstrated in Fig. 1, this problem is avoided by using Eq. (1.6). The details of this particular recurrence relation are given in Section 4, and the stability of such relations is discussed in, for example [9].

Another benefit offered by Eq. (1.6) can be appreciated when evaluating a function expressed as in Eq. (1.2) with the sum truncated at, say, $m = M$. Even when nested according to the Horner scheme, the evaluation of $P_m(x)$ requires about $2m$ arithmetic operations. In this way, assuming all the coefficients are pre-computed, the evaluation of $S(x)$ therefore requires about M^2 arithmetic operations. When Eq. (1.6) is employed, on the other hand, each of the basis elements is then determined with just five arithmetic operations, so only $5M$ are required for the entire set. In either case, forming the linear combination that constitutes $S(x)$ takes another $2M$ operations, but the main point to be made is that an "$O(M^2)$ process" is converted to one of $O(M)$. This comes on top of the vital accuracy gain

#128403 - $15.00 USD Received 12 May 2010; revised 2 Jun 2010; accepted 3 Jun 2010; published 11 Jun 2010

demonstrated in Fig. 1. More importantly, however, the focus of this paper is on three additional steps that were developed decades ago by Clenshaw [10], Smith [11], and Salzer [12], respectively, to extract additional benefits from Eq. (1.6). Their elegant, but little used, ideas are given unified derivations below and are adapted to the context of optics.

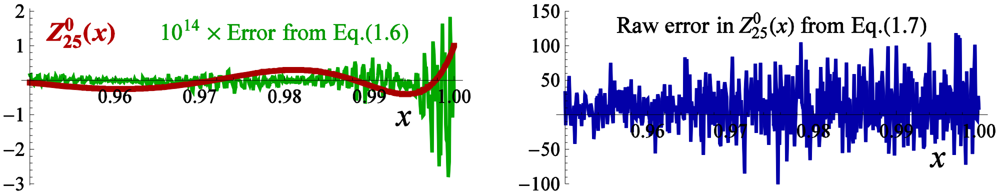

Fig. 1. The red curve at left is a plot of a high-order rotationally symmetric Zernike over just $0.95 < x < 1$; the green curve is 10^{14} times the error when this function is evaluated by using Eq. (1.6), so there are typically 14 significant digits in this double-precision result. The curve on the right is the catastrophic error when Eq. (1.7) is used at double precision. Although exact to within round-off, the explicit polynomial gives values with the wrong order of magnitude and chaotic sign. This failure grows exponentially with the polynomial's order.

2. Linear combinations of orthogonal polynomials and their derivatives

An obvious application of Eq. (1.6) is for functions defined as a truncated sum of the type in Eq. (1.2), say

$$S(x) := \sum_{m=0}^{M} s_m P_m(x). \tag{2.1}$$

The required basis members can first be evaluated sequentially by using Eq. (1.6), and the end result then formed as a linear combination with the coefficients s_m. Clenshaw [10] presented a related recurrence relation that allows these two steps to be combined. As proven in Appendix A, by starting with $\alpha_M = s_M$ and $\alpha_{M-1} = s_{M-1} + (a_{M-1} + b_{M-1}\,x)s_M$, and progressively working downward from $n = M-2$ by using

$$\alpha_n \;=\; s_n + (a_n + b_n\,x)\alpha_{n+1} - c_{n+1}\alpha_{n+2}\,, \tag{2.2}$$

the desired result is given by just $S \;=\; \alpha_0$. This process has all the robustness of Eq. (1.6), and it is even more efficient at evaluating Eq. (2.1). Each cycle of Eq. (2.2) involves three multiplications and three additions, and this must be done $M-1$ times to determine the sum of interest, namely α_0. In this way, when a_n, b_n, and c_n are pre-computed, Eq. (2.1) can be evaluated robustly with about $6M$ operations via a remarkably simple bit of code to perform nothing more than Eq. (2.2). Notice that no single basis member is evaluated along the way; *this is an elegant short-circuit from the spectrum to the function value.* When nested as a Horner scheme, evaluating a regular polynomial of order M requires only $2M$ operations. A factor of three in operation count is evidently the minimal cost that must be paid to avoid catastrophic numerical instability. There are significant side-benefits, however, and a sample is commented on in one context in Section 5. Many of them follow from the fact that the rms value of a function is just the square root of the sum of the squared coefficients, as in Eq. (1.4).

Clenshaw's process turns out to open the way to an equally effective scheme for computing derivatives. Although Eq. (1.6) can be differentiated to find a recurrence relation for $P_n'(x) = \frac{\mathrm{d}}{\mathrm{d}x}P_n(x)$, a more direct determination of $S'(x)$ can be performed with a simple variant of Eq. (2.2). In fact, the same thing works for higher derivatives. If the j'th derivative

#128403 - $15.00 USD Received 12 May 2010; revised 2 Jun 2010; accepted 3 Jun 2010; published 11 Jun 2010

is written as $S^{(j)}(x)$, it is proven in Appendix B that if we start with $\alpha^{(j)}_{M-j+1}=0$ and $\alpha^{(j)}_{M-j} = j\,b_{M-j}\,\alpha^{(j-1)}_{M-j+1}$, then

$$\alpha^{(j)}_n = j\,b_n\,\alpha^{(j-1)}_{n+1} + (a_n + b_n\,x)\alpha^{(j)}_{n+1} - c_{n+1}\,\alpha^{(j)}_{n+2} \tag{2.3}$$

can be used by working down from $n = M - j - 1$ to find the desired result: $S^{(j)} = \alpha^{(j)}_0$ for any $j > 0$. Since Eqs. (2.2) and (2.3) have the same structure, it is effectively the same block of highly optimized code that can be employed to implement them both. That the derivatives follow so readily is a more significant benefit than the minor efficiency gain in going from the $7M$ operations discussed before Fig. 1 to the $6M$ operations for Clenshaw.

Something equivalent to Eqs. (2.3) was presented by Smith [11], but with minimal derivation. The argument in his paper seems to proceed via differentiation of Eq. (1.6) instead of Eq. (2.2), as done in Appendix B. He presents two intermediate expressions that appear to complicate any interpretation of the process, and inconsistently applies a notational convention for higher derivatives. [For example, Smith presents something akin to Eq. (2.3), but without the factor of j that appears immediately after the equals sign.] The confusion this may have caused could explain the fact that his beautiful result seems to have been ignored in much of the related literature, even in numerical classics like [9]. Similarly, Clenshaw's result has also been poorly appreciated in physics and optics and, again, it may be no coincidence that I have been unable to find a clear derivation of Eq. (2.2) in the literature. This is what motivated me to develop Appendices A and B. It is also interesting to note that Doha [13] and Barrio [14] both offer some interesting analyses and alternative results, although they are specific to Jacobi polynomials: Doha re-expresses $S^{(j)}$ as a linear combination (with constant coefficients) of the original basis elements, i.e. in terms of P_m instead of $P^{(j)}_m$, while Barrio gives $S^{(j)}$ without requiring all lower derivatives. The fact that the derivatives of Jacobi polynomials are once again Jacobi polynomials, opens the door for their special results. Thankfully, the simple process described above applies more generally to any polynomials that satisfy Eq. (1.6) and Eq. (2.3) is all that is required in our context to determine derivatives.

3. Change of basis

When a function is expressed in terms of the basis $\{P^{\text{I}}_m\}$ as

$$S = \sum_{m=0}^{M} s^{\text{I}}_m P^{\text{I}}_m\,, \tag{3.1}$$

consider determining its representation in terms of a second basis, say

$$S = \sum_{m=0}^{M} s^{\text{II}}_m P^{\text{II}}_m\,. \tag{3.2}$$

That is, the goal is to determine s^{II}_m from s^{I}_m where each of the polynomial bases satisfies a three-term recurrence relation like Eq. (1.6), say

$$P^{\text{I}}_{n+1} = (a^{\text{I}}_n + b^{\text{I}}_n x)P^{\text{I}}_n - c^{\text{I}}_n P^{\text{I}}_{n-1}\,, \qquad P^{\text{II}}_{n+1} = (a^{\text{II}}_n + b^{\text{II}}_n x)P^{\text{II}}_n - c^{\text{II}}_n P^{\text{II}}_{n-1}\,. \tag{3.3a,b}$$

By using Eqs. (3.3), Ting and Luke [15] derived a five-term recurrence relation for determining the elements of an upper-triangular change-of-basis matrix that multiplies s^{I}_m to give s^{II}_m. Just as the methods discussed in the previous section absorbed s_m into a single-pass recurrence, Salzer [12] presented an analogous, apparently little-known process to change bases. That is, Salzer's five-term recurrence absorbs s^{I}_m, and its output is the s^{II}_m elements

#128403 - $15.00 USD Received 12 May 2010; revised 2 Jun 2010; accepted 3 Jun 2010; published 11 Jun 2010

themselves. (Salzer predated Ting and Luke, but the latter's comments suggest that they misunderstood his method.) Sample applications are presented in Sections 4 and 5 where this method greatly simplifies past and current developments in optics.

Salzer's process is derived in Appendix C in terms of what can be regarded as an $(M+1)\times(M+1)$ matrix with elements α_k^n [not to be confused with the differentiated polynomial written as $\alpha_n^{(j)}$ in Section 2]. It is convenient to introduce some auxiliary matrices in terms of the constants from the recurrence relations of Eqs. (3.3), namely

$$f_k^n := b_n^{\mathrm{I}} / b_{k-1}^{\mathrm{II}}, \qquad g_k^n := a_n^{\mathrm{I}} - b_n^{\mathrm{I}} a_k^{\mathrm{II}} / b_k^{\mathrm{II}}, \qquad h_k^n := b_n^{\mathrm{I}} c_{k+1}^{\mathrm{II}} / b_{k+1}^{\mathrm{II}}. \quad \text{(3.4a,b,c)}$$

In some applications, it may be appropriate for these to be pre-computed and stored. It turns out that α_k^n is triangular with $(M+1)(M+2)/2$ non-zero elements: $\alpha_k^n = 0$ unless $0 \le k \le M-n$. This property can be used to simplify the penultimate equation of Appendix C, namely Eq. (C.6). At $k=0$, one term drops out leaving

$$\alpha_0^n = s_n^{\mathrm{I}} + g_0^n \alpha_0^{n+1} + h_0^n \alpha_1^{n+1} - c_{n+1}^{\mathrm{I}} \alpha_0^{n+2}. \tag{3.5}$$

A different term drops for $0 < k < M-n-1$, so that

$$\alpha_k^n = f_k^n \alpha_{k-1}^{n+1} + g_k^n \alpha_k^{n+1} + h_k^n \alpha_{k+1}^{n+1} - c_{n+1}^{\mathrm{I}} \alpha_k^{n+2}. \tag{3.6}$$

For $k = M-n-1$, three terms vanish leaving only

$$\alpha_{M-n-1}^n = f_{M-n-1}^n \alpha_{M-n-2}^{n+1} + g_{M-n-1}^n \alpha_{M-n-1}^{n+1}, \tag{3.7}$$

and only one term remains when $k = M-n$, namely

$$\alpha_{M-n}^n = f_{M-n}^n \alpha_{M-n-1}^{n+1}. \tag{3.8}$$

The change of basis is realized by starting at $n=M$, where the only non-zero term is $\alpha_0^M = s_M^{\mathrm{I}}$, and with the two terms for $n = M-1$, namely $\alpha_0^{M-1} = s_{M-1}^{\mathrm{I}} + g_0^{M-1}\alpha_0^M$ and $\alpha_1^{M-1} = f_1^{M-1}\alpha_0^M$. From these; Eqs. (3.5)-(3.8) can be used to work down from $n = M-2$ to $n=0$. It is shown in the appendix that the desired result is just $s_m^{\mathrm{II}} = \alpha_m^0$.

Equations (3.5)-(3.8) allow each α_k^n to be determined in no more than seven arithmetic operations. The roughly $M^2/2$ non-zero elements can therefore be determined with about $3.5M^2$ operations. Notice that, as for the processes in Section 2, this change of basis is achieved without evaluating a single member of either basis set. More significantly, there is no need to determine integrals involving the basis functions as you might expect from Eq. (1.3). It is a powerful and efficient shortcut. Its value is enhanced by the fact that the bases need not be orthogonal polynomials; all that is required is that they satisfy a three-term recurrence. For example, the regular monomial basis $\{x^m\}$ obviously satisfies Eq. (1.6) with $b_n = 1$ and $a_n = c_n = 0$, and this is exploited in Section 5.

4. Applications for Zernike polynomials

As stated in the Introduction, the Zernike polynomials of m'th azimuthal order are related to the Jacobi polynomials by $Z_n^m(x) = P_n^{(0,m)}(2x-1)$. It follows that $Z_n^m(x)$ satisfies a recurrence relation like Eq. (1).6). With $s := m+2n$, this relation has

$$a_n = -\frac{(s+1)[(s-n)^2+n^2+s]}{(n+1)(s-n+1)s}, \quad b_n = \frac{(s+2)(s+1)}{(n+1)(s+n-1)}, \quad c_n = \frac{(s+2)(s-n)n}{(n+1)(s-n+1)s}. \quad \text{(4.1a,b,c)}$$

#128403 - $15.00 USD Received 12 May 2010; revised 2 Jun 2010; accepted 3 Jun 2010; published 11 Jun 2010

(See, for example, 22.7.1 of [3]). Any number of Zernikes can therefore be computed robustly and efficiently by using Eq. (1.6). More importantly, the same is possible for linear combinations of Zernikes, and also of their derivatives, by applying the results of Section 2. Perhaps the most significant observation relates to the problem of changing aperture size. Suppose that the modified aperture is a fraction, say $\varepsilon > 0$, of the first, where typically $\varepsilon < 1$. This process arises in a variety of contexts including ophthalmology and lithography, and it has been revisited repeatedly over the last decade, see e.g [16–20]. In all of this work, each azimuthal order can be treated separately for this change of basis. The challenge therefore is, for given values of s_n, find t_n so that $\sum_n s_n r^m Z_n^m(r^2)$ is identical (as a function of r) to $\sum_n t_n (r/\varepsilon)^m Z_n^m(r^2/\varepsilon^2)$.

The authors of [16–20] have started from Eq. (1.7) and tackled laborious algebra in order to derive explicit expressions for the matrix elements in the change-of-basis matrix. The problem is that their results take an almost identical form to Eq. (1.7) —see e.g., Eq. (19) of [19] or Eq. (29) of [20]. These matrix elements are polynomials in ε^2 and, while the expressions are exact to within round-off errors, they nevertheless fail as catastrophically as Eq. (1.7) when the number of terms grows to 20 or more. To make matters worse, this failure is most extreme for $\varepsilon \approx 1$, which is precisely where the expressions are typically used. An elegant, specialized alternative is given in [21] where the matrix elements are themselves expressed as differences between two Zernikes. Their impressive derivation relies on two integral relations involving Bessel functions. A similarly robust and efficient scheme follows simply from the more general results given in the previous section, however.

This direct application of the results of Section 3 involves finding t_n so that

$$\sum_{n=0}^{M} s_n P_n^{\mathrm{I}}(x) \equiv \sum_{n=0}^{M} (t_n/\varepsilon^m) P_n^{\mathrm{II}}(x/\varepsilon^2), \tag{4.2}$$

where $\{P_n^{\mathrm{I}}\}$ and $\{P_n^{\mathrm{II}}\}$ are both just $\{Z_n^m\}$. Both of the recurrence relations in Eqs. (3.3) are therefore given by Eqs. (4.1), but with the one distinction that $b_n^{\mathrm{II}} = b_n/\varepsilon^2$ due to the scaling of x in the second basis. The end result is then evidently just $t_n = \varepsilon^m \alpha_n^0$. In this way, Salzer's elegant general-purpose routine means that there was no need to wrestle with special functions or explicit high-order expansions. More importantly, the failure of the explicit expressions is thereby avoided. It is also interesting to notice that, for this application, f_k^n, g_k^n, and h_k^n of Eqs. (3.4) are linear functions of ε^2. This means that α_k^j is a polynomial of order $M - j$ in ε^2 and, just as in Eq. (B.3), Eqs. (3.5)-(3.8) can be differentiated with respect to ε^2 to obtain a recurrence for determining the derivative of t_n with respect to ε^2. Such an entity is used for sensitivity analysis in [21].

5. Applications for asphere shape specification

Advances in both fabrication and metrology are enabling the introduction of more complex and precise aspheric optical surfaces. This trend demands larger numbers of terms in the polynomial used in the standard characterization of surface shape. The methods described above are ideally suited to one of the orthogonal bases that were recently introduced for efficient characterization in this context [22]. If c denotes the axial curvature and κ the conic constant, the surface's sag is written as

$$z(\rho) = c\,\rho^2 \Big/ [1 + \sqrt{1-(1+\kappa)c^2\rho^2}\,] + u^4 \sum_{m=0}^{M} s_m Q_m^{\mathrm{con}}(u^2), \tag{5.1}$$

where u is just the normalized radial coordinate, i.e. $u := \rho / \rho_{\max}$ with $\rho_{\max} = \mathrm{CA}/2$. A sample of these basis members is plotted in Fig. 2. It so happens that $Q_m^{\mathrm{con}}(x) \equiv Z_m^4(x)$, so the recurrence relation for this set follows from Eqs. (4.1) with $m = 4$:

$$a_n = -\frac{(2n+5)(n^2+5n+10)}{(n+1)(n+2)(n+5)}, \quad b_n = \frac{2(n+3)(2n+5)}{(n+1)(n+5)}, \quad c_n = \frac{(n+3)(n+4)n}{(n+1)(n+2)(n+5)}. \quad (5.2\text{a,b,c})$$

The methods of Sections 2 and 3 are ideally matched to working with aspheres in these terms. In this way, the potentially chronic round-off problems are localized to the basis members themselves, and tamed there by using recurrence relations. As discussed in [22], the fact that the coefficients now function like a spectrum has a variety of benefits.

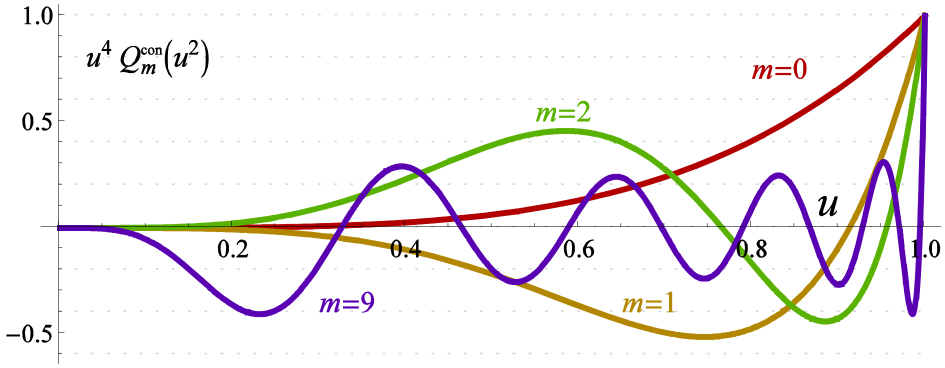

Fig. 2. A sample of the members of the basis used in Eq. (5.1).

Upon introducing $S(x) := \sum_m s_m\, Q_m^{\mathrm{con}}(x)$ and $\phi = [1-(1+\kappa)c^2\rho^2]^{1/2}$, Eq. (5.1) and its derivatives become

$$z(\rho) = c\,\rho^2/(1+\phi) + u^4 S(u^2), \quad (5.3)$$

$$z'(\rho) = c\,\rho/\phi + \frac{2u^3}{\rho_{\max}}\left[2S(u^2) + u^2 S'(u^2)\right], \quad (5.4)$$

$$z''(\rho) = c/\phi^3 + \frac{2u^2}{\rho_{\max}{}^2}\left[6\,S(u^2) + 9\,u^2 S'(u^2) + 2\,u^4 S''(u^2)\right]. \quad (5.5)$$

Such entities are fundamental to working with aspheres, and they can now be robustly evaluated to arbitrary orders with remarkable efficiency by using Eqs. (2.2) and (2.3). Scaling the aperture size is another basic process in this context. Because $Q_m^{\mathrm{con}}(x) \equiv Z_m^4(x)$, this can be achieved simply by taking $m = 4$ in the process described in Section 4. That is, the new coefficients that give exactly the same surface —but now orthogonalized over a different aperture size— can be determined robustly and efficiently via simple code.

Perhaps more interestingly, the conventional manner to describe surfaces is by expressing the additive polynomial of Eq. (5.1) as a sum of monomials, say

$$u^4 \sum_{m=0}^{M} s_m\, Q_m^{\mathrm{con}}(u^2) = \sum_{m=0}^{M} A_{2m+4}\, \rho^{2m+4} = (u\,\rho_{\max})^4 \sum_{m=0}^{M} A_{2m+4}\,(u\,\rho_{\max})^{2m}. \quad (5.6)$$

With $x := u^2$, this relation can be rewritten as

$$\sum_{m=0}^{M} s_m\, Q_m^{\mathrm{con}}(x) = \sum_{m=0}^{M} t_m\, x^m, \quad (5.7)$$

#128403 - $15.00 USD Received 12 May 2010; revised 2 Jun 2010; accepted 3 Jun 2010; published 11 Jun 2010

where $t_m := A_{2m+4}\,\rho_{\max}^{2m+4}$. It is once again straightforward to apply the process presented in Section 3 to convert backwards and forwards between these two representations: one of the recurrences is described by Eq. (5.2), while the other has a simpler form, namely

$$a_n = 0, \qquad b_n = 1, \qquad c_n = 0. \tag{5.8a,b,c}$$

The fact that two of these coefficients vanish means that terms drop out of Eqs. (3.5)-(3.8), and a slightly optimized form may be implemented for each of these inter-conversions. Of course, the degeneracy of $\{x^m\}$ means that the conventional representation is prone to round-off failure. The conversion to and from this basis is therefore recommended only for modest numbers of terms (not many more than a dozen or so). Orthogonal bases will hopefully soon be the norm in this field, and there will be little need for such conversions.

6. Concluding remarks

Four basic processes are described in this work, namely the evaluation of (i) a set of orthogonal polynomials, (ii) a linear combination of them, (iii) a linear combination of their derivatives, and (iv) changing from one such basis to another. These regularly enter optics, but much of the related work was focused on explicit expressions for the polynomials. It has not been widely appreciated that, although they are exact, these expressions fail numerically when the number of terms grows much beyond ten or so. Further, these decompositions also tend to become computationally intensive as the number of terms grows. Robust and efficient alternatives were presented here in each case. None of these alternatives for (ii)-(iv) require the evaluation of a single orthogonal polynomial; they all operate directly upon the coefficients of the expansion and lead to simple code that avoids round-off problems.

Although these methods were introduced decades ago, it appears that they were ahead of their time. They now open the door to benefits delivered by larger numbers of terms in various contexts including modeling based on Gaussian-beam-mode decompositions, various applications of Zernike polynomials, or optical design based on the methods of Section 5. The two-step induction-based process that was introduced in the Appendices clarifies how these methods can be generalized to other contexts. In my view, the work of Clenshaw, Smith, and Salzer certainly deserves to be better appreciated in optics, and more widely known in general. Many of us can then benefit significantly by having faster, more versatile code that does not suffer from catastrophic round-off failure and can proceed to arbitrary orders.

Appendix A: Directly evaluating linear combinations

The three-term recurrence relations that underpin the key results in this work can be expressed as

$$P_{n+1} = u_n P_n - v_n P_{n-1}, \tag{A.1}$$

where explicit expressions are in hand for u_n and v_n. Once P_0 and P_1 are specified, this relation can be used to generate those of higher order. Clenshaw [10] pointed out that Eq. (A.1) leads to a fast, robust way to evaluate any sum of the form

$$S := \sum_{m=0}^{M} s_m P_m. \tag{A.2}$$

The idea is to avoid the evaluation of the P_m elements by progressively eliminating them from the top end of the sum. The three-term recurrence in Eq. (A.1) means that the topmost two terms are therefore progressively modified. After eliminating P_M, $P_{M-1}, \ldots$ P_{n+2}, where $n < M$, the result can be expressed as

#128403 - $15.00 USD Received 12 May 2010; revised 2 Jun 2010; accepted 3 Jun 2010; published 11 Jun 2010

$$S = \alpha_{n+1} P_{n+1} + \beta_{n+1} P_n + \sum_{m=0}^{n-1} s_m P_m \,, \tag{A.3}$$

for some α_{n+1} and β_{n+1}. It follows trivially from Eqs. (A.1) and (A.3) that

$$\begin{aligned} S &= \alpha_{n+1}\left(u_n P_n - v_n P_{n-1}\right) + \beta_{n+1} P_n + \sum_{m=0}^{n-1} s_m P_m \\ &= \left(u_n \alpha_{n+1} + \beta_{n+1}\right) P_n + \left(s_{n-1} - v_n \alpha_{n+1}\right) P_{n-1} + \sum_{m=0}^{n-2} s_m P_m . \end{aligned} \tag{A.4}$$

Upon replacing n in Eq. (A.3) with $n-1$ and comparing the result with Eq. (A.4), it is then evident that

$$\alpha_n = u_n \alpha_{n+1} + \beta_{n+1} \,, \tag{A.5}$$

$$\beta_n = s_{n-1} - v_n \alpha_{n+1} \,. \tag{A.6}$$

These two relations hold the key to the efficient evaluation of Eq. (A.2) since taking $n = 0$ in Eq. (A.3) reveals that

$$S = \alpha_1 P_1 + \beta_1 P_0 \,. \tag{A.7}$$

By starting with $\alpha_{M+2} = \beta_{M+2} = 0$, Eqs. (A.5) and (A.6) yield α_n and β_n for all $n \le M+1$, hence S can then be determined simply with Eq. (A.7).

Things can be simplified a little further by using Eqs. (A.6) and (A.5), respectively, to eliminate the β 's from Eqs. (A.5) and (A.7):

$$\alpha_n = s_n + u_n \alpha_{n+1} - v_{n+1} \alpha_{n+2} \,, \tag{A.8}$$

$$S = (\alpha_0 - u_0 \alpha_1) P_0 + \alpha_1 P_1 \,. \tag{A.9}$$

Since Eq. (A.5) establishes that $\alpha_{M+1} = 0$, start Eq. (A.8) with $\alpha_{M+2} = \alpha_{M+1} = 0$ and work downwards to determine α_1 and, ultimately, α_0. The value of the sum in Eq. (A.2) then follows from Eq. (A.9). Keep in mind that P_0 and P_1 are specified at the outset. Although, as pointed out in [9], it is also possible to proceed from bottom up rather than top down in eliminating the P 's, this is not effective for the optical applications considered here. It is important to note, however, that [9] gives a useful treatment of applications involving trigonometric and Bessel functions as well as a sketch of stability analysis for recurrence-based processes. For many cases of interest, including those considered in Sections 4 and 5, it turns out that $v_0 = 0$ hence $P_1 = u_0 P_0$, and the conventional normalization is $P_0 \equiv 1$. As a result, Eq. (A.9) then reduces to an even more beautiful result:

$$S = \alpha_0 \,. \tag{A.10}$$

Appendix B: Directly evaluating derivatives of linear combinations

When P_m in Eq.(A.2) is a polynomial of order m in x and each s_m is independent of x, Smith [11] presented a scheme to compute the derivative of that sum with respect to x. With primes denoting derivatives, we now seek

$$S' = \sum_{m=1}^{M} s_m P'_m \,. \tag{B.1}$$

#128403 - $15.00 USD Received 12 May 2010; revised 2 Jun 2010; accepted 3 Jun 2010; published 11 Jun 2010

(Note that $P_0' = 0$, so it was dropped.) In this case, u_n of Eq. (A.1) is a linear function, say

$$u_n = a_n + b_n x, \tag{B.2}$$

and α_n of Appendix A is readily seen to be a polynomial of order $M-n$ for $0 \le n \le M$. Since v_n is independent of x, differentiating Eq. (A.8) now leads directly to

$$\alpha_n' = b_n \alpha_{n+1} + u_n \alpha_{n+1}' - v_{n+1} \alpha_{n+2}'. \tag{B.3}$$

This can be initialized with $\alpha_{M+1}' = \alpha_M' = 0$, and the result that we seek is then simply

$$S' = \alpha_0'. \tag{B.4}$$

That is —assuming that S was evaluated beforehand— S' can be evaluated just as robustly and with much the same number of arithmetic operations.

Of course, if $v_0 \ne 0$, it is the derivative of Eq. (A.9) that yields the desired end result, but such cases aren't needed in this work. What is more, even when u_n and v_n are more general functions of x, it is straightforward to use the derivative of Eq. (A.8) to get a slight generalization of Eq. (B.3). For the most common case laid out above, it is also clear that higher derivatives can be determined in much the same way: if a superscript in parentheses is used to denote higher derivatives, it follows trivially from Eq. (A.8) that

$$\alpha_n^{(j)} = j\, b_n \alpha_{n+1}^{(j-1)} + u_n \alpha_{n+1}^{(j)} - v_{n+1} \alpha_{n+2}^{(j)}. \tag{B.5}$$

This is initialized by $\alpha_{M+2-j}^{(j)} = \alpha_{M+1-j}^{(j)} = 0$, and the desired higher derivative for our cases is then simply $\alpha_0^{(j)}$:

$$S^{(j)} = \sum_{m=j}^{M} s_m P_m^{(j)} = \alpha_0^{(j)}. \tag{B.6}$$

Remarkably, nothing more complex than Eq. (B.5) is needed to robustly determine derivatives of any order.

Appendix C: Directly changing to a new basis

The inter-conversion considered in Sec. 3 can be carried out by, much as in Appendix A, progressively eliminating P_m^{I} from the top end of the sum in Eq.(3.1). In this case, however, α_n is to be expresses as a linear combination of $\{P_m^{\mathrm{II}}\}$ rather than as the nested polynomials of Appendix A. The intermediate result that replaces Eq.(A.3) is written as

$$S = \left(\sum_{k=0}^{M-n-1} \alpha_k^{n+1} P_k^{\mathrm{II}}\right) P_{n+1}^{\mathrm{I}} + \left(\sum_{k=0}^{M-n} \beta_k^{n+1} P_k^{\mathrm{II}}\right) P_n^{\mathrm{I}} + \sum_{m=0}^{n-1} s_m^{\mathrm{I}} P_m^{\mathrm{I}}, \tag{C.1}$$

where the α's and β's are now just constant coefficients. By using Eq. (3.3a), the first term on the right-hand side of Eq. (C.1) can be re-expressed as

$$\begin{aligned} \left(\sum_{k=0}^{M-n-1} \alpha_k^{n+1} P_k^{\mathrm{II}}\right) P_{n+1}^{\mathrm{I}} &= \left(\sum_{k=0}^{M-n-1} \alpha_k^{n+1} P_k^{\mathrm{II}}\right) \left[(a_n^{\mathrm{I}} + b_n^{\mathrm{I}} x) P_n^{\mathrm{I}} - c_n^{\mathrm{I}} P_{n-1}^{\mathrm{I}} \right] \\ &= \left(\sum_{k=0}^{M-n-1} \alpha_k^{n+1} x P_k^{\mathrm{II}}\right) b_n^{\mathrm{I}} P_n^{\mathrm{I}} + \left(\sum_{k=0}^{M-n-1} \alpha_k^{n+1} P_k^{\mathrm{II}}\right) \left[a_n^{\mathrm{I}} P_n^{\mathrm{I}} - c_n^{\mathrm{I}} P_{n-1}^{\mathrm{I}} \right]. \end{aligned} \tag{C.2}$$

In turn, the first term on the last line of Eq. (C.2) can be re-expressed upon using Eq. (3.3b) to see that

$$xP_k^{\mathrm{II}} = \left[P_{k+1}^{\mathrm{II}} - a_k^{\mathrm{II}} P_k^{\mathrm{II}} + c_k^{\mathrm{II}} P_{k-1}^{\mathrm{II}}\right] / \, b_k^{\mathrm{II}} . \tag{C.3}$$

Combining Eqs. (C.2) and (C.3) with Eq. (C.1) gives

$$\begin{aligned} S = {} & \left(\sum_{k=0}^{M-n-1} \alpha_k^{n+1} \left[P_{k+1}^{\mathrm{II}} - a_k^{\mathrm{II}} P_k^{\mathrm{II}} + c_k^{\mathrm{II}} P_{k-1}^{\mathrm{II}}\right] / \, b_k^{\mathrm{II}} \right) b_n^{\mathrm{I}} P_n^{\mathrm{I}} \\ & + \left(\sum_{k=0}^{M-n-1} \alpha_k^{n+1} P_k^{\mathrm{II}} \right) \left[a_n^{\mathrm{I}} P_n^{\mathrm{I}} - c_n^{\mathrm{I}} P_{n-1}^{\mathrm{I}} \right] + \left(\sum_{k=0}^{M-n} \beta_k^{n+1} P_k^{\mathrm{II}} \right) P_n^{\mathrm{I}} + \sum_{m=0}^{n-1} s_m^{\mathrm{I}} P_m^{\mathrm{I}} . \end{aligned} \tag{C.4}$$

Upon gathering the terms of Eq. (C.4) that involve P_{n-1}^{I} and, separately, those that involve P_n^{I}, the result can be equated to Eq. (C.1) with n replaced by $n-1$. The first relation that emerges is simply

$$\beta_k^n = s_{n-1}^{\mathrm{I}} \, \delta_{k0} - c_n^{\mathrm{I}} \alpha_k^{n+1} , \tag{C.5}$$

where δ_{k0} is zero for all k except $\delta_{00} = 1$. As done for the β 's in Appendix A, Eq. (C.5) makes it easy to eliminate β_k^n from the second relation that emerges in order to determine an analogue for Eq. (A.8), namely a recurrence relation for α_k^n :

$$\alpha_k^n = s_n^{\mathrm{I}} \, \delta_{k0} + \left(\frac{b_n^{\mathrm{I}}}{b_{k-1}^{\mathrm{II}}} \right) \alpha_{k-1}^{n+1} + \left(a_n^{\mathrm{I}} - \frac{b_n^{\mathrm{I}} a_k^{\mathrm{II}}}{b_k^{\mathrm{II}}} \right) \alpha_k^{n+1} + \left(\frac{b_n^{\mathrm{I}} c_{k+1}^{\mathrm{II}}}{b_{k+1}^{\mathrm{II}}} \right) \alpha_{k+1}^{n+1} - c_{n+1}^{\mathrm{I}} \alpha_k^{n+2} . \tag{C.6}$$

Notice that the factors inside the parentheses in Eq. (C.6) involve just the constants from the recurrence relations in Eqs. (3.3). For the cases of interest in the body, we always have $c_0^{\mathrm{I}} = c_0^{\mathrm{II}} = 0$ and $P_0^{\mathrm{I}} \equiv P_0^{\mathrm{II}} \equiv 1$. To appreciate the power of Eq. (C.6), notice that just as Eq. (A.3) becomes (A.10) upon taking $n = -1$, Eq. (C.1) leads to

$$S = \sum_{k=0}^{M} \alpha_k^0 P_k^{\mathrm{II}} . \tag{C.7}$$

Upon comparing Eqs. (3.2) and (C.7), it follows that the coefficients sought here are just $s_m^{\mathrm{II}} = \alpha_m^0$. It follows from Eq. (C.1) that $\alpha_k^n = 0$ whenever either $k < 0$ or $k > M - n$ so, for example, $\alpha_k^{M+2} = \alpha_k^{M+1} = 0$ for all k. With this as a given, Eq. (C.6) can be used, for fixed n, by running over $0 \le k \le M - n$. The process starts at $n = M$ and works down to $n = 0$ in order to yield the desired end result.