Robust, efficient computational methods for axially symmetric optical aspheres

G.W. Forbes*

QED Technologies Inc., 1040 University Ave., Rochester NY 14607, USA
**forbes@qedmrf.com*

Abstract: Whether in design or the various stages of fabrication and testing, an effective representation of an asphere's shape is critical. Some algorithms are given for implementing tailored polynomials that are ideally suited to these needs. With minimal coding, these results allow a recently introduced orthogonal polynomial basis to be employed to arbitrary orders. Interestingly, these robust and efficient methods are enabled by the introduction of an auxiliary polynomial basis.

OCIS codes: (220.1250) Aspherics; (220.4830) Optical systems design; (220.4840) Optical testing; (220.4610) Optical fabrication; (000.4430) Numerical approximation and analysis.

References and links

1. G. W. Forbes, "Shape specification for axially symmetric optical surfaces," Opt. Express **15**, 5218–5226, (2007) *http://www.opticsinfobase.org/oe/abstract.cfm?URI=oe-15-8-5218*.
2. G. W. Forbes, and C. P. Brophy, "Designing cost-effective systems that incorporate high-precision aspheric optics", SPIE Optifab (2009) TD06–25 (1). Available at *http://www.qedmrf.com*.
3. C. du Jeu, "Criterion to appreciate difficulties of aspherical polishing," Proc. SPIE **5494**, 113–121 (2004), doi:10.1117/12.551420.
4. G. W. Forbes, "Robust and fast computation for the polynomials of optics," Opt. Express **18**, 13851–13862, (2010) *http://www.opticsinfobase.org/oe/abstract.cfm?URI=oe-18-13-13851*.
5. E. W. Weisstein, "Jacobi Polynomial" from MathWorld—A Wolfram Web Resource. *http://mathworld.wolfram.com/JacobiPolynomial.html*, see esp. Equations (10–14).
6. M. Abramowitz, and I. Stegun, *Handbook of Mathematical Functions* (Dover, 1978), Chap. 22.
7. 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.
8. F. J. Smith, "An Algorithm for Summing Orthogonal Polynomial Series and their Derivatives with Applications to Curve-Fitting and Interpolation," Math. Comput. **19**, 33–36 (1965).
9. W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, 1992) Section 2.9.
10. A useful overview is given at *http://en.wikipedia.org/wiki/Discrete_cosine_transform*.
11. G. W. Forbes, "Can you make/measure this asphere for me?", Frontiers in Optics, OSA Technical Digest (2009), *http://www.opticsinfobase.org/abstract.cfm?URI=FiO-2009-FThH1*.

1. Introduction

When either designing, fabricating, or testing optical aspheres, there are advantages to characterizing the surface's shape in terms of orthogonal bases [1]. The coefficients in such a description can be interpreted as a spectrum-like decomposition of the shape. These individual coefficients have immediate interpretations, and they permit a part to be assessed at a glance. For $\{Q_m^{\text{bfs}}(x)\}$ of [1], in particular, the coefficients relate directly to the asphere's deviation from the best-fit sphere measured along the surface normal. The fringe density in a conventional full-aperture interferogram of such a part is proportional to the rate of change of this normal departure, referred to here as the *normal slope*. The weighted mean square value of this normal slope is directly linked to the sum of the squares of the associated coefficients when using $\{Q_m^{\text{bfs}}(x)\}$. In fact, this property served as the defining relation for the basis and facilitates design constraints when the goal is mild aspheres that are easily tested. It is important to appreciate, however, that this basis is also well suited to the characterization of arbitrarily strong aspheres. As a first step in this direction, a related design constraint was introduced for aspheres that are to be tested by using stitched interferometry [2]. More

#129305 - $15.00 USD Received 1 Jun 2010; revised 12 Jul 2010; accepted 12 Jul 2010; published 1 Sep 2010

generally, a part's manufacturability is related to the variation in the local principal curvatures of the surface, see e.g [3]. Because the normal slope couples directly to changes in the principal curvatures, this basis is well suited to a variety of manufacturability-driven constraints.

Robust and efficient computational methods are vital during the design and modeling of systems that incorporate aspheric surfaces. A trend to more complex aspheres has led to the need for more terms in the polynomial used to characterize shape. The catastrophic failure associated with round-off in these cases was recently highlighted [4] and it was shown that, when using orthogonal polynomials, recurrence relations play a vital role in avoiding this obstacle. Methods were presented in [4] for robustly computing not only the polynomials themselves, but also any linear combination of them and of their derivatives. These methods are all based on the three-term recurrence relation satisfied by standard orthogonal polynomials. It turns out, however, that $\{Q_m^{\text{bfs}}(x)\}$ does not satisfy a three-term recurrence. Instead, it is shown in the following sections that there is a non-standard recurrence relation that can be used to derive effective algorithms for working with this basis. In fact, the most effective algorithms are found to operate with an auxiliary set of polynomials in tandem.

2. Unconventional recurrence relations

If $\rho_{\max}$ denotes one half of the part's clear aperture, a rotationally symmetric asphere's shape is specified here in cylindrical polar coordinates as

$$z \;=\; \frac{c\,\rho^2}{1+\sqrt{1-c^2\rho^2}} \;+\; \frac{u^2(1-u^2)}{\sqrt{1-c^2\rho^2}}\sum_{m=0}^{M} a_m Q_m(u^2), \tag{2.1}$$

where c is the curvature of the best-fit sphere, u is the normalized radial coordinate defined by $u := (\rho/\rho_{\max})$, and $Q_m(x)$ is a polynomial of order m. The coefficients written as a_m serve to characterize the asphere's departure from its best-fit sphere. Note that I typically drop the superscript on $Q_m^{\text{bfs}}(x)$ now because this is the only one of the bases from [1] that is used in this work. This basis is constructed so that

$$(\tfrac{2}{\pi})\int_0^1 S_m(u)\,S_n(u)\tfrac{1}{\sqrt{1-u^2}}\,\mathrm{d}u \;=\; \delta_{mn}, \tag{2.2}$$

where δ_{mn} is Kronecker's delta and the normal slope of the basis members is defined by

$$S_m(u) := \tfrac{\mathrm{d}}{\mathrm{d}u}[u^2(1-u^2)Q_m(u^2)]. \tag{2.3}$$

The normalization factor in Eq. (2.2) follows from

$$\int_0^1 \tfrac{1}{\sqrt{1-u^2}}\,\mathrm{d}u \;=\; \pi/2. \tag{2.4}$$

Notice that the additive aspheric term in Eq. (2.1) corresponds to deviation measured along the axis. To first order, this is converted to departure along the surface normal by multiplying it by a cosine factor. This conversion is why the square root in the denominator in Eq. (2.1) is absent from Eq. (2.3). A sample of these basis members is plotted in Fig. 1 together with their normal slopes. The clear sine-like character of $S_m(u)$ is intuitively consistent with the fact that the mean of the squared normal slope is just the sum of the squares of the a_m coefficients, see Eq. (15) of [1].

#129305 - $15.00 USD Received 1 Jun 2010; revised 12 Jul 2010; accepted 12 Jul 2010; published 1 Sep 2010

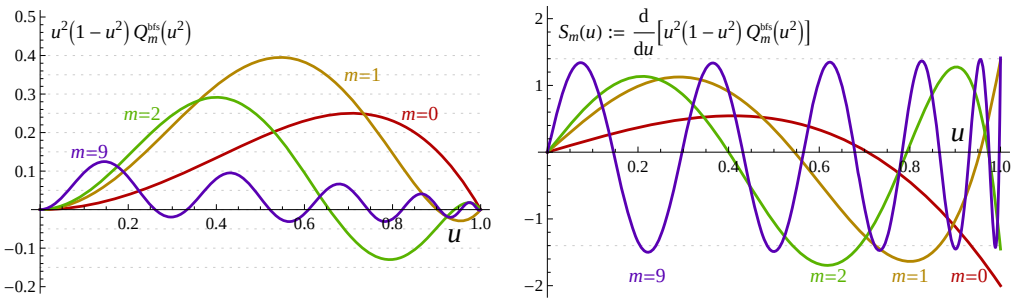

Fig. 1. A sample of the slope-orthogonal polynomials used in this work. By construction, they are not only orthogonal in slope, but normalized so that their mean square slope is unity.

It is shown in Appendix A that $\{Q_m^{bfs}(x)\}$ is simply related to a particular family of standard polynomials. In particular, Eq. (A.12) can be written more explicitly as

$$P_m(x) \quad = \quad f_m\,Q_m(x) + g_{m-1}\,Q_{m-1}(x) + h_{m-2}\,Q_{m-2}(x), \qquad (2.5)$$

where $P_m(x)$ is a scaled Jacobi polynomial. An efficient process for evaluating the constants f_m, g_m, and h_m is given in Eqs. (A.14-16) of Appendix A. It follows from Eq. (A.4) together with 22.7.1 and 22.4.1 of [6], that these auxiliary polynomials can be generated robustly by a standard three-term recurrence relation of particularly simple form:

$$P_{m+1}(x) \quad = \quad (2-4x)\,P_m(x) - P_{m-1}(x). \qquad (2.6)$$

Equation (2.6) is initiated with $P_0(x) = 2$ and $P_1(x) = 6 - 8x$. A sample of these polynomials is plotted in Fig. 2. As can be seen from Eq. (A.6), the envelope drawn in Fig. 2 is given by $2u(1-u^2)$, which peaks at $u = 3^{-1/2} \approx 0.58$ where it takes the value $3^{-1/2}\,4/3 \approx 0.77$.

Upon eliminating the P's from Eq. (2.6) by using Eq. (2.5), an unconventional five-term recurrence relation emerges for $\{Q_m^{bfs}(x)\}$. It turns out that, rather than doing this explicitly, it is more effective for several reasons to operate with $\{P_m(x)\}$ and $\{Q_m^{bfs}(x)\}$ in tandem. First, notice that $\{Q_m^{bfs}(x)\}$ can be generated robustly to arbitrary orders upon rewriting Eq. (2.5) as an unconventional three-term recurrence relation:

$$Q_{m+1}(x) \quad = \quad [P_{m+1}(x) - g_m\,Q_m(x) - h_{m-1}\,Q_{m-1}(x)]/\,f_{m+1}, \qquad (2.7)$$

with $Q_0(x) = 1$ and $Q_1(x) = 19^{-1/2}(13 - 16x)$. More importantly, as shown in the next section, the change of basis matrix determined in Appendix A can be used to great effect for a number of critical operations when working with aspheres in these terms.

#129305 - $15.00 USD Received 1 Jun 2010; revised 12 Jul 2010; accepted 12 Jul 2010; published 1 Sep 2010

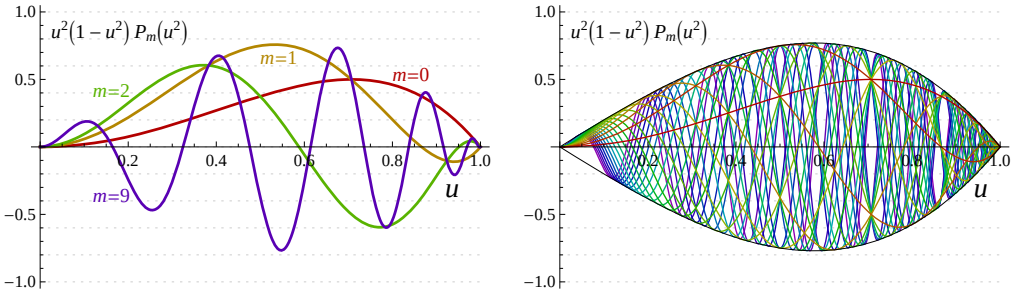

Fig. 2. A sample of the scaled Jacobi polynomials used in this work. The plot at left is for comparison with Fig. 1, and the one on the right holds up to $m = 20$ to reveal their envelope.

3. Linear combinations and their derivatives

The key piece of interest in Eq. (2.1) can be written as $S(u^2)$, where S is defined by

$$S(x) \;:=\; \sum_{m=0}^{M} a_m\, Q_m(x). \tag{3.1}$$

If the list of these basis elements is denoted by $\mathbf{q}(x)$, then $S(x) = \mathbf{a}\cdot\mathbf{q}(x)$. Similarly, write the auxiliary basis, i.e. $\{P_m(x)\}$, as $\mathbf{p}(x)$, and define a second function, say $T(x) := \mathbf{b}\cdot\mathbf{p}(x)$. Given that $\mathbf{p}(x) = \mathrm{L}\,\mathbf{q}(x)$ where L is the lower-triangular band matrix of Eq. (2.5) or Eq. (A.12), it follows that $T(x) = \mathbf{b}\cdot[\mathrm{L}\,\mathbf{q}(x)] = (\mathrm{L}^{\mathrm{T}}\,\mathbf{b})\cdot\mathbf{q}(x)$. It is now evident that $S(x) \equiv T(x)$ when $\mathbf{a} = \mathrm{L}^{\mathrm{T}}\,\mathbf{b}$. As a consequence, we can change basis from $\{P_m(x)\}$ to $\{Q_m^{\mathrm{bfs}}(x)\}$ simply by using

$$a_m \;=\; f_m\, b_m + g_m\, b_{m+1} + h_m\, b_{m+2}, \tag{3.2}$$

for $m = 0, 1, 2, \ldots M-2$ and then

$$a_{M-1} \;=\; f_{M-1}\, b_{M-1} + g_{M-1}\, b_M \qquad \text{and} \qquad a_M \;=\; f_M\, b_M. \tag{3.3a,b}$$

The inverse of this transformation is also easily realized by reversing these steps: start with

$$b_M \;=\; a_M / f_M \qquad \text{and} \qquad b_{M-1} \;=\; (a_{M-1} - g_{M-1}\, b_M)/ f_{M-1}, \tag{3.4a,b}$$

and, for $m = M-2$ down to 0, use

$$b_m \;=\; (a_m - g_m\, b_{m+1} - h_m\, b_{m+2})/ f_m. \tag{3.5}$$

The recurrence relation in Eq. (3.5) is clearly just a re-arranged version of Eq. (3.2). In this way, with about $5M$ arithmetic operations, it is possible to interchange between these two bases. [Such a process typically involves $O(M^2)$ operations, but the band matrix accelerates it in this case.]

An advantage of swapping bases is that, as shown in [4], Clenshaw's method [7] allows

$$S(x) \;\equiv\; \sum_{m=0}^{M} b_m\, P_m(x), \tag{3.6}$$

to be computed robustly by exploiting the recurrence relation in Eq. (2.6): Simply start with

$$\alpha_M \;=\; b_M \qquad \text{and} \qquad \alpha_{M-1} \;=\; b_{M-1} + (2-4x)\,\alpha_M, \tag{3.7a,b}$$

and work down with

#129305 - $15.00 USD Received 1 Jun 2010; revised 12 Jul 2010; accepted 12 Jul 2010; published 1 Sep 2010

$$\alpha_m \;=\; b_m + (2-4x)\;\alpha_{m+1} - \alpha_{m+2}\,, \tag{3.8}$$

from $m = M-2$ to 0. The desired end result is then given —see Eq. (A.9) of [4] — as

$$\begin{aligned} S(x) &= [\alpha_0 - (2-4x)\;\alpha_1]P_0(x) + \alpha_1 P_1(x) \\ &= \;2(\alpha_0 + \alpha_1). \end{aligned} \tag{3.9}$$

The final step follows from the expressions for the lowest two basis members that were given after Eq. (2.6). That is, with the roughly $5M$ arithmetic operations needed to evaluate the α's via Eqs. (3.7) and (3.8), $S(x)$ can be computed without evaluating a single member of either basis.

It follows from Eqs. (3.7) and (3.8) that α_m is a polynomial of order $M-m$ in x. As discussed in [4], Smith [8] demonstrated that any derivative of $S(x)$ of Eq. (3.6) can be evaluated with similar ease. If the j'th derivative is written as $S^{(j)}(x)$, then start with

$$\alpha^{(j)}_{M-j+1} = 0 \qquad \text{and} \qquad \alpha^{(j)}_{M-j} = -4\,j\;\alpha^{(j-1)}_{M-j+1}\,, \tag{3.10a,b}$$

and work down from $m = M-2$ to 0 by using

$$\alpha^{(j)}_m = (2-4x)\;\alpha^{(j)}_{m+1} - \alpha^{(j)}_{m+2} - 4\,j\,\alpha^{(j-1)}_{m+1}\,. \tag{3.11}$$

The desired end result is then

$$S^{(j)}(x) = 2\,(\alpha^{(j)}_0 + \alpha^{(j)}_1)\,. \tag{3.12}$$

Note that this process assumes that the $(j-1)$'th derivative has been evaluated before going after the j'th, and that $\alpha^{(0)}_m$ is equal to α_m. It follows from Eq. (2.1) that, with $\phi := (1-c^2\rho^2)^{1/2}$, the surface's sag and its first two derivatives are given in these terms by

$$z(\rho) = \frac{c\,\rho^2}{1+\varphi} \;+\; \frac{u^2\,(1-u^2)}{\varphi}\,S(u^2), \tag{3.13}$$

$$z'(\rho) = \frac{c\,\rho}{\varphi} \;+\; \frac{u[1+\varphi^2 - u^2(1+3\varphi^2)]}{\rho_{\max}\,\varphi^3}\,S(u^2) \;+\; \frac{2u^3\,(1-u^2)}{\rho_{\max}\,\varphi}\,S'(u^2), \tag{3.14}$$

$$\begin{aligned} z''(\rho) = \frac{c}{\varphi^3} \;+\; & \frac{3-\varphi^2-3u^2(1+\varphi^2+2\varphi^4)}{\rho_{\max}^{\,2}\,\varphi^5}\,S(u^2) \\ & +\; \frac{2u^2[2+3\varphi^2-u^2(2+7\varphi^2)]}{\rho_{\max}^{\,2}\,\varphi^3}\,S'(u^2) \;+\; \frac{4u^4\,(1-u^2)}{\rho_{\max}^{\,2}\,\varphi}\,S''(u^2). \end{aligned} \tag{3.15}$$

In typical applications, the sag of a specific part and its derivatives will be evaluated for multiple values of ρ. In such cases, $\{\,a_m\,\}$ is converted to $\{\,b_m\,\}$ only once, and these serve as the input —along with various ρ values— to the methods just described. For example, to plot a specific member of $\{Q^{\text{bfs}}_m(x)\}$ or its derivatives, rather than use a process like that discussed at Eq. (2.7), it is better to set all of $\{\,a_m\,\}$ to zero except the one of interest and convert this list to the associated $\{\,b_m\,\}$ via Eqs. (3.4) and (3.5). The desired results then follow efficiently from Eqs. (3.7) to (3.9) and (3.10) to (3.12). It is also worth noting that the part's axial curvature follows upon taking $\rho = 0$ (hence $u = 0$ and $\phi = 1$) in Eq. (3.15):

$$c_{\text{axial}} = z''(0) = c + \frac{2}{\rho_{\max}^{\,2}}\,S(0) \;=\; c \;+\; \frac{4}{\rho_{\max}^{\,2}}\sum_{m=0}^{M}(2m+1)\,b_m\,. \tag{3.16}$$

#129305 - $15.00 USD Received 1 Jun 2010; revised 12 Jul 2010; accepted 12 Jul 2010; published 1 Sep 2010

This final expression follows from the fact that $P_m(0) = 2(2m+1)$, which can be derived inductively from Eq. (2.6). By either solving Eq. (3.16) for c, or leaving it in terms of c_{axial}, it is possible to consider the independent curvature variable to be either the axial curvature or the best-fit curvature when using this representation. Taking c_{axial} to be the more fundamental entity can be helpful for paraxial analysis and constraints in the design process.

4. Fitting to this orthogonal basis

Consider an aspheric surface specified by a general sag function of the form

$$z = f(\rho), \tag{4.1}$$

where $f(\rho)$ is symmetric and $f(0) = 0$. To express this in the form given in Eq. (2.1), the first step is to determine the curvature of the best-fit sphere by using the requirement that it meets the asphere at both $\rho = 0$ and $\rho = \rho_{\max}$:

$$c = \frac{2f(\rho_{\max})}{\rho_{\max}^2 + f(\rho_{\max})^2}. \tag{4.2}$$

The task is then reduced to choosing $\{a_m\}$ so that

$$f(\rho) \approx \frac{c\rho^2}{1+\sqrt{1-c^2\rho^2}} + \frac{u^2(1-u^2)}{\sqrt{1-c^2\rho^2}} \sum_{m=0}^{M} a_m Q_m(u^2). \tag{4.3}$$

Since $\{a_m\}$ enters this relation linearly, a standard linear least-squares process can determine the solution via an $(M+1)\times(M+1)$ matrix inversion. This is adequate for most purposes. It is impressive, however, that swapping to the auxiliary basis —where a and Q in Eq. (4.3) are replaced by b and P — opens an option for an elegant and more efficient solution.

After re-arranging Eq. (4.3) and inserting $\rho = u\rho_{\max}$, the task in terms of the auxiliary basis is to determine $\{b_m\}$ so that

$$\sum_{m=0}^{M} b_m P_m(u^2) \approx \frac{\sqrt{1-(uc\rho_{\max})^2}}{u^2(1-u^2)} \left\{ f(u\rho_{\max}) - \frac{c(u\rho_{\max})^2}{1+\sqrt{1-(uc\rho_{\max})^2}} \right\} =: F(u). \tag{4.4}$$

The last relation serves to define $F(u)$ which, provided $f'(0) = 0$, remains well defined even at the endpoints of $0 \le u \le 1$: the term inside the braces has zeros that remove the singularities that are potentially caused by the denominator outside the braces. It follows from Eq. (B.7) of Appendix B that a particular solution for these coefficients is just

Drop this factor of 2:

$$b_m \approx \frac{(-1)^m}{2N} \sum_{j=0}^{N-1} F_j \cos\left[\tfrac{\pi}{N}(m+\tfrac{1}{2})(j+\tfrac{1}{2})\right], \tag{4.5}$$

where

$$F_j := \cos\left(\tfrac{\pi}{2N}(j+\tfrac{1}{2})\right) F\left[\cos\left(\tfrac{\pi}{2N}(j+\tfrac{1}{2})\right)\right]. \tag{4.6}$$

As discussed in Appendix B, Eq. (4.5) is a standard discrete cosine transform (namely DCT-IV) and, when more than just a few terms are required in the fit, it may be better implemented via FFT-like options. In our context, it is unusual for the spectrum to spill significantly beyond just tens of terms. For most purposes, therefore, troubles with aliasing are insignificant if we use Eq. (4.5) with N about 16 or 32. Remember that, as shown in Fig. 2, the maximum contribution to the normal departure from the best-fit sphere associated with

#129305 - $15.00 USD Received 1 Jun 2010; revised 12 Jul 2010; accepted 12 Jul 2010; published 1 Sep 2010

any b_m is about $0.8b_m$. It is therefore straightforward to truncate the list of these coefficients so that the magnitude of the dropped terms is less than whatever cut-off is desired. As a final step, Eqs. (3.2) and (3.3) yield $\{a_m\}$.

For the purpose of demonstration, consider a parabola specified by $c_{\text{axial}} = 20\text{mm}$ and $\rho_{\max} = 20\text{mm}$. It follows from Eq. (4.2) that the best-fit curvature is $c = 25\text{mm}$. A scaled lens drawing and a plot of the associated sag-based departure from best-fit sphere are presented in Fig. 3. The fit coefficients that result from Eq. (4.5) with $N = 32$ for this case are plotted in Fig. 4. As it does here, the magnitude of the coefficients generally decays rapidly with order. Notice that these coefficients hit the floor of numerical round-off for double precision at about $m = 20$. The order at which this happens depends on the characteristics of the shape. For nm-level accuracy, since the maximum contribution from each term is about $0.8b_m$, this list of coefficients can be truncated at about the level of the solid gray line in the plot.

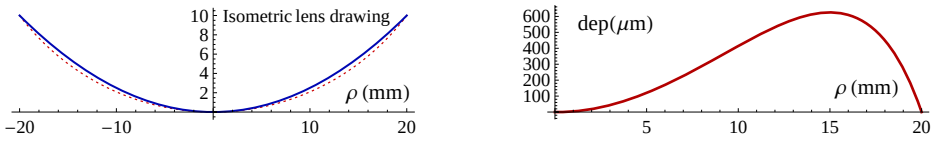

Fig. 3. A cross-section of the parabola used for demonstration (in blue) is shown with its best-fit sphere (dotted red). The aspheric departure (viz. sag difference) is plotted at right.

To facilitate code verification, the eight coefficients above the solid gray line in Fig. 4 are found to be

$$\begin{aligned}\mathbf{b} = \{&1009010.04959,\ 2770.64974485,\ -4739.30847163,\\ &1172.09704743,\ -257.270488293,\ 55.4172061289,\\ &-11.966650385,\ 2.604636\,675\,85\}\ \text{nm}.\end{aligned} \tag{4.7}$$

If the last term in this subset is also dropped, the fit error can therefore be expected to resemble the $m = 7$ plot of Fig. 2 with a scale factor of 2.6nm. This is perfectly consistent with the computed fit error plotted at right in Fig. 4. After converting the seven retained coefficients to $\{a_m\}$ according to Eqs. (3.2) and (3.3), the specification for this surface is found to be

$$\mathbf{a} = \{2019004,\ 7143,\ -13944,\ 4190,\ -1095,\ 283,\ -68\}\ \text{nm}. \tag{4.8}$$

It was these rounded coefficients that were used to compute the fit error plotted in Fig. 4. Interestingly, for any $N \geq 16$, just the last few digits vary in only the smaller members of the list in Eq. (4.7). Even with $N = 8$ in this case, the aliasing effects are so small that nm-level accuracy is achieved in the resulting fit.

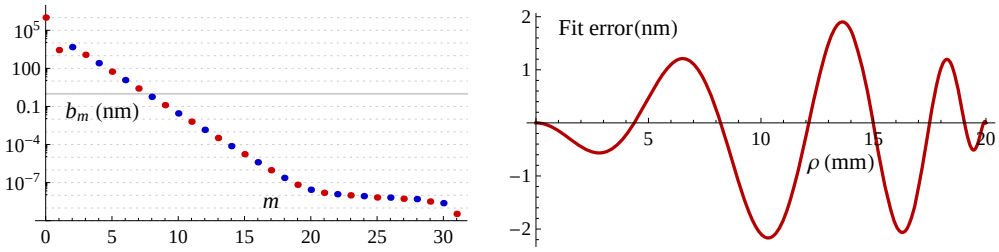

Fig. 4. Log plot of $\{b_m\}$ in nm units. The dots are red/blue for positive/negative values. The 1nm reference level is drawn as a solid gray line. The plot at right is the error in the fit that results if all but the last red dot above the 1nm line are retained (i.e. m = 0 to 6 are kept).

#129305 - $15.00 USD Received 1 Jun 2010; revised 12 Jul 2010; accepted 12 Jul 2010; published 1 Sep 2010

5. Annular apertures

For obstructed mirror systems, the aspheric departure of each surface is defined (and measured) over only the appropriate annulus, say $\varepsilon\rho_{\max} < \rho < \rho_{\max}$ where $0 < \varepsilon < 1$. With a minor compromise, the ideas described above can be applied to such cases with minimal change. The best-fit sphere is now chosen to pass through both the inner and the outer edge of the annulus, and Eq. (2.1) is replaced by

$$z = \frac{c\rho^2}{1+\sqrt{1-c^2\rho^2}} + \frac{(u^2-\varepsilon^2)(1-u^2)}{(1-\varepsilon^2)\sqrt{(1+\varepsilon)(1-c^2\rho^2)}}\sum_{m=0}^{M} a_m Q_m\left(\tfrac{u^2-\varepsilon^2}{1-\varepsilon^2}\right), \qquad (5.1)$$

This turns out to be orthonormal in slope if we choose the inner product for the mean square slope in this case to be

$$\left[\tfrac{2}{(1-\varepsilon)\pi}\right]\int_{\varepsilon}^{1} S_m(u)\,S_n(u)\,\tfrac{1}{u}\sqrt{\tfrac{u^2-\varepsilon^2}{1-u^2}}\,\mathrm{d}u, \qquad (5.2)$$

where Eq. (5.1) means that the normal slope is now defined by

$$S_m(u) := \tfrac{1}{(1-\varepsilon^2)\sqrt{1+\varepsilon}}\frac{\mathrm{d}}{\mathrm{d}u}\left[(u^2-\varepsilon^2)(1-u^2)Q_m\left(\tfrac{u^2-\varepsilon^2}{1-\varepsilon^2}\right)\right]. \qquad (5.3)$$

The normalization factor in Eq. (5.2) follows from

$$\int_{\varepsilon}^{1} \tfrac{1}{u}\sqrt{\tfrac{u^2-\varepsilon^2}{1-u^2}}\,\mathrm{d}u = (1-\varepsilon)\pi/2. \qquad (5.4)$$

Orthonormality can be confirmed by changing variables to $x := (u^2-\varepsilon^2)/(1-\varepsilon^2)$ in order to see that Eqs. (5.2) and (5.3) reduce to precisely the inner product in Eq. (A.5) of Appendix A.

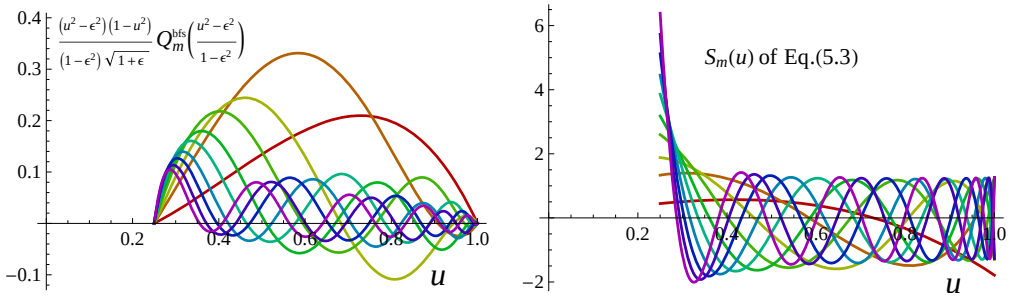

Fig. 5. The first ten basis members and their slopes for an obstructed aperture with ε = 0.25.

The compromise here is that the weight function in Eq. (5.2) goes to zero at the inner edge of the annulus, i.e. at $u = \varepsilon$. As a result, as is evident in Fig. 5, the slope maps are not as sine-like as those in Fig. 1, but tend to grow at this edge. Ideally, a new set of polynomials would be defined by replacing the weight in Eq. (5.2) by a weight that diverges similarly at each edge, namely $u\,[(u^2-\varepsilon^2)(1-u^2)]^{-1/2}$. In my opinion, this complication is unjustified, and it is better to re-use the results derived above for the unobstructed case with minor adaptation. The key is that, at its heart, Eq. (5.1) is built around a sum like that in Eq. (3.1), hence also Eq. (3.6). On account of the factor $2/x^{1/2}$ in Eq. (A.6), the envelope in the analogue of Fig. 2 is now $2(1-u^2)[(u^2-\varepsilon^2)/(1-\varepsilon)]^{1/2}/(1+\varepsilon)$, with a peak value of $\frac{4}{3}(1-\varepsilon)[(1+\varepsilon)/3]^{1/2}$. Of course, the DCT is now applied to the resulting analog of Eq. (4.4). In the limit of small ε, all these results reduce precisely to those for the unobstructed case.

#129305 - $15.00 USD Received 1 Jun 2010; revised 12 Jul 2010; accepted 12 Jul 2010; published 1 Sep 2010

6. Concluding remarks

Although two bases are used in this work, the end user need only ever consider one of them; the other can be hidden within the software as an internal computational aid. Because manufacturability is more closely coupled to the normal slope than to sag, I regard $\{Q_m^{\text{bfs}}(x)\}$ as the principal basis, i.e. I prefer to think in terms of **a** rather than **b**. If the illuminated region plus whatever margin is required (for fabrication) differs significantly from the disc of radius $\rho_{\max}$, however, the constraints based on **a** will lose validity. It can therefore be helpful during optimization to be able to adjust an asphere's clear aperture while preserving its shape. Fortunately, the method described in Section 4 allows this adjustment to be performed efficiently. The elegance and power of all the methods described above mean that such processes may well spread into the domain of characterizing mid-spatial frequencies on optical surfaces: with Eqs. (4.4)-(4.6), it is just as easy now to use hundreds of terms to fit a measured part. High orders can also be valuable in simulations during design for tolerancing, etc. Although it is beyond our current scope, it is natural for these sorts of applications —as well as for the growing field of freeform optics— to include non-rotationally symmetric terms by using a Fourier-series-based process that is similar to that used in the construction of the Zernike polynomials.

The methods reported in this work lead to concise code that offers a remarkably efficient manner to work with aspheres in terms of $\{Q_m^{\text{bfs}}(x)\}$. These methods are robust to arbitrary orders and facilitate characterizations like Eq. (4.8) that have been found to hold typically three times fewer digits than the traditional representation [11]. Their spectrum-like properties also allow the manufacturing challenge to be estimated at a glance. For the example in Sec. 4, for instance, the rapid decay in the coefficients reveals that the shape is dominated by the lowest order term, i.e. the red curve of Fig. 1. What's more, its peak value is therefore roughly 0.25 times the first coefficient of the list in Eq. (4.8). Aspheric departure plots —in this case, the solid red curve in Fig. 3— can typically be anticipated in this way with no computation. What's more, the number of terms that are essential becomes immediately self-evident. Perhaps most important of all, these new processes avoid the catastrophic round-off failure that has plagued this field. Whether in the context of design, fabrication, or testing, the benefits of an orthogonal basis are compelling, and these new recurrence-based methods mean that any additional operational costs are insignificant.

Appendix A: Recurrence involving a Jacobi polynomial

The integral in Eq. (2.2) can be re-written by changing the variable of integration from u to $x := u^2$ and using Eq. (2.3) to see that this condition is equivalent to

$$\begin{aligned}(\tfrac{2}{\pi})\int_0^1 \Big\{&[(1-2x)Q_m(x)+x(1-x)Q_m'(x)]\\ &\times[(1-2x)Q_n(x)+x(1-x)Q_n'(x)]\Big\}\sqrt{\tfrac{4x}{1-x}}\,\mathrm{d}x \;=\; \delta_{mn},\end{aligned} \tag{A.1}$$

where primes denote derivatives. If $\phi_n(x)$ is any polynomial of order n, the Jacobi polynomials satisfy [5]

$$\int_0^1 \varphi_n(x)\,P_m^{(\alpha,\beta)}(2x-1)\,(1-x)^{\alpha}x^{\beta}\mathrm{d}x \;=\; 0, \quad \text{for } m>n, \tag{A.2}$$

and

$$\tfrac{\mathrm{d}}{\mathrm{d}x}[P_n^{(\alpha,\beta)}(x)] \;=\; \tfrac{1}{2}(n+\alpha+\beta+1)\,P_{n-1}^{(\alpha+1,\beta+1)}(x). \tag{A.3}$$

#129305 - $15.00 USD Received 1 Jun 2010; revised 12 Jul 2010; accepted 12 Jul 2010; published 1 Sep 2010

By using Eqs. (A.2) and (A.3) together with the standard recurrence relation given in Eq. (12) of [5], or at 22.7.1 of [6], it can be seen that, if $Q_m(x)$ is replaced by $P_m^{(-1/2,1/2)}(2x-1)$ throughout Eq. (A.1), the expression on its left-hand side vanishes whenever $|m-n|>2$. That is, once expanded out, the integral of each term is found to be zero, so the associated matrix is quindiagonal: it has the main diagonal and two bands of non-zero elements on either side. This reveals a useful link between this particular family of Jacobi polynomials and $Q_m^{\text{bfs}}(x)$. It turns out to be convenient to express what follows in terms of

$$P_m(x) := (-1)^m \tfrac{2\,(2m)!!}{(2m-1)!!} P_m^{(-1/2,1/2)}(2x-1), \tag{A.4}$$

where $n!! = (n-2)!!\,n$, and $0!! = (-1)!! = 1$, so $7!! = 7\cdot 5\cdot 3\cdot 1$ etc.

As suggested by Eqs. (2.2) and (2.3) and the change of variables used for Eq. (A.1), the natural inner-product for this context can be defined as

$$\big\langle f, g \big\rangle := \ (\tfrac{2}{\pi}) \int_0^1 \tfrac{\mathrm{d}}{\mathrm{d}x}[x(1-x)\,f(x)]\,\tfrac{\mathrm{d}}{\mathrm{d}x}[x(1-x)\,g(x)]\sqrt{\tfrac{4x}{1-x}}\,\mathrm{d}x. \tag{A.5}$$

In this way, Eq. (2.2) can be expressed as $<Q_m, Q_n> = \delta_{mn}$. More interestingly, it was established in the previous paragraph that $<P_m, P_n>$ is quindiagonal. This can be seen more directly upon using an explicit expression for $P_m(x)$ that follows from a link to the Chebyshev polynomials of the first kind, see 22.3.15 and 22.5.29 of [6], namely

$$P_m(x) \ = \ (-1)^m 2\cos[\tfrac{2m+1}{2}\arccos(2x-1)]\big/\sqrt{x}\,, \tag{A.6}$$

Upon changing the variable of integration by using $x = \cos^2\theta$, it follows from Eqs. (A.5) and (A.6) that

$$\big\langle P_m,\ P_n \big\rangle \ = \ (\tfrac{2}{\pi}) \int_0^{\pi/2} T_m(\theta) T_n(\theta)\,\mathrm{d}\theta \ = \ (\tfrac{1}{\pi}) \int_{-\pi/2}^{\pi/2} T_m(\theta) T_n(\theta)\,\mathrm{d}\theta, \tag{A.7}$$

where

$$\begin{aligned} T_m(\theta) &:= (-1)^m \tfrac{2}{\sin\theta}\tfrac{\mathrm{d}}{\mathrm{d}\theta}\{\cos\theta \sin^2\theta \cos[(2m+1)\theta]\} \\ &= (-1)^m\{(m+2)\cos[(2m+3)\theta] + \cos[(2m+1)\theta] - (m-1)\cos[(2m-1)\theta]\}. \end{aligned} \tag{A.8}$$

The cosine function's familiar orthogonality now makes it clear from Eqs. (A.7) and (A.8) that $<P_m, P_n>$ is quindiagonal. What's more, it becomes straightforward to evaluate its elements. The three independent non-zero bands in this symmetric matrix are next handled separately.

First, with $H_m := <P_{m+2}, P_m>$, it follows that the only non-zero contribution in Eq. (A.7) is of the form

$$\begin{aligned} H_m \ &= \ -(\tfrac{1}{\pi})(m+2)(m+1) \int_{-\pi/2}^{\pi/2} \cos^2[(2m+3)\theta]\,\mathrm{d}\theta \\ &= \ -(m+2)(m+1)/2. \end{aligned} \tag{A.9}$$

Next, for $G_m := <P_{m+1}, P_m>$, it follows similarly (although now with two non-zero terms) that

$$\begin{aligned} G_m \ &= \ -(\tfrac{1}{\pi})\Big\{(m+2) \int_{-\pi/2}^{\pi/2} \cos^2[(2m+3)\theta]\,\mathrm{d}\theta - m \int_{-\pi/2}^{\pi/2} \cos^2[(2m+1)\theta]\,\mathrm{d}\theta\Big\} \\ &= \ -1, \end{aligned} \tag{A.10}$$

#129305 - $15.00 USD Received 1 Jun 2010; revised 12 Jul 2010; accepted 12 Jul 2010; published 1 Sep 2010

and finally, for $F_m := < P_m, P_m >$, it is found that the three non-zero terms give

$$\begin{aligned} F_m &= (\tfrac{1}{\pi})[(m+2)^2 + 1 + (m-1)^2]\tfrac{\pi}{2} \\ &= (m^2 + m + 3), \qquad \text{for} \quad m > 0. \end{aligned} \tag{A.11}$$

When $m = 0$, however, the $\cos[(2m+1)\theta]$ and $\cos[(2m-1)\theta]$ terms are no longer orthogonal. Instead, they add directly and it is readily seen that $F_0 = 4$. As shown next, these matrix elements play a fundamental role in the relation between $\{Q_m^{\text{bfs}}(x)\}$ and $\{P_m(x)\}$.

Any two polynomial bases are related by a change-of-basis matrix through a relation of the form

$$P_i(x) = \sum_j L_{ij}\, Q_j(x). \tag{A.12}$$

The fact that any n'th order polynomial can obviously be expressed as a combination of $Q_m(x)$ for $m = 0,1,\ldots n$ means that this matrix is lower-triangular, i.e. $L_{ij} = 0$ for $j > i$. By using $< Q_m, Q_n > = \delta_{mn}$, it now follows that

$$\begin{aligned} \langle P_m, P_n \rangle &= \left\langle \sum_j L_{mj}\, Q_j(x), \sum_k L_{nk}\, Q_k(x) \right\rangle \\ &= \sum_j \sum_k L_{mj}\, L_{nk} \left\langle Q_j(x), Q_k(x) \right\rangle = \sum_j L_{mj}\, L_{nj}\,. \end{aligned} \tag{A.13}$$

That is, the quindiagonal $< P_m, P_n >$ is just the product of the matrix with elements L_{jk} and its transpose, say $\mathrm{L}\,\mathrm{L}^{\mathrm{T}}$. It follows that L is also a band matrix: the only non-zero elements are along the diagonal and the two bands below the diagonal.

The elements of the change-of-basis matrix, i.e. L, can now be evaluated by using the standard process of Cholesky decomposition [9]. That is, if the non-zero elements of the change-of-basis matrix are denoted by $f_m := L_{mm}$, $g_m := L_{m+1\,m}$, and $h_m := L_{m+2\,m}$, these elements can be found by starting with $f_0 = 2$, $f_1 = 19^{1/2}/2$, $g_0 = -1/2$ and working up from $m = 2$ by applying

$$h_{m-2} = H_{m-2} / f_{m-2} = -m(m-1)/(2 f_{m-2}), \tag{A.14}$$

$$g_{m-1} = (G_{m-1} - g_{m-2} h_{m-2}) / f_{m-1} = -(1 + g_{m-2} h_{m-2}) / f_{m-1}, \tag{A.15}$$

$$f_m = \sqrt{F_m - g_{m-1}^2 - h_{m-2}^2} = \sqrt{m(m+1) + 3 - g_{m-1}^2 - h_{m-2}^2}, \tag{A.16}$$

in this order. The first equality in each of Eqs. (A14-16) is just for explanatory purposes; the second expression is all that is required. For reference, the initial sub-block of this change-of-basis matrix is

#129305 - $15.00 USD Received 1 Jun 2010; revised 12 Jul 2010; accepted 12 Jul 2010; published 1 Sep 2010

$$\begin{bmatrix} f_0 & 0 & 0 & 0 & 0 & 0 \\ g_0 & f_1 & 0 & 0 & 0 & 0 \\ h_0 & g_1 & f_2 & 0 & 0 & 0 \\ 0 & h_1 & g_2 & f_3 & 0 & 0 \\ 0 & 0 & h_2 & g_3 & f_4 & 0 \\ 0 & 0 & 0 & h_3 & g_4 & f_5 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ \frac{-1}{2} & \sqrt{\frac{19}{4}} & 0 & 0 & 0 & 0 \\ \frac{-1}{2} & \frac{-5}{2\sqrt{19}} & 4\sqrt{\frac{10}{19}} & 0 & 0 & 0 \\ 0 & \frac{-6}{\sqrt{19}} & \frac{-17}{2\sqrt{190}} & \frac{1}{2}\sqrt{\frac{509}{10}} & 0 & 0 \\ 0 & 0 & \frac{-3}{2}\sqrt{\frac{19}{10}} & \frac{-91}{2\sqrt{5090}} & 6\sqrt{\frac{259}{509}} & 0 \\ 0 & 0 & 0 & -20\sqrt{\frac{10}{509}} & \frac{-473}{2\sqrt{131831}} & \frac{1}{2}\sqrt{\frac{25607}{259}} \end{bmatrix}. \quad \text{(A.17)}$$

Equations (A.14) to (A.16) yield the m'th row of this matrix from the earlier rows, and this process is robust to arbitrary orders. It turns out that, for large m, $f_m \approx -h_m \approx 2^{-1/2} m$ and $g_m \approx -2^{-1/2}$.

Appendix B: Determining a fit in terms of the auxiliary Jacobi polynomial

The auxiliary polynomials used here are defined by Eq. (A.4) and satisfy

$$\tfrac{1}{2\pi}\int_0^1 P_m(x) P_m(x) \sqrt{\tfrac{x}{1-x}}\,\mathrm{d}x = \delta_{mn}. \quad \text{(B.1)}$$

For any given $g(x)$, it follows that the fit that minimizes the mean square error defined by

$$E^2 := \tfrac{1}{2\pi}\int_0^1 \{g(x) - \sum_m b_m P_m(x)\}^2 \sqrt{\tfrac{x}{1-x}}\,\mathrm{d}x, \quad \text{(B.2)}$$

is given simply by

$$b_m = \tfrac{1}{2\pi}\int_0^1 g(x) P_m(x) \sqrt{\tfrac{x}{1-x}}\,\mathrm{d}x. \quad \text{(B.3)}$$

Upon changing the variable of integration to $x = \cos^2\theta$, Eqs. (A.6) and (B.3) lead to

$$b_m = \tfrac{(-1)^m}{\pi}\int_{-\pi/2}^{\pi/2} g(\cos^2\theta)\cos\theta\cos[(2m+1)\theta]\,\mathrm{d}\theta. \quad \text{(B.4)}$$

That is, each of these coefficients is given uniquely and explicitly by a simple integral.

There are various options for efficiently evaluating the integral in Eq. (B.4). One follows upon using elementary trigonometric identities and rescaling the variable of integration:

$$\begin{aligned} b_m &= \tfrac{(-1)^m}{2\pi}\int_{-\pi/2}^{\pi/2} g(\cos^2\theta)\{\cos[2(m+1)\theta] + \cos[2m\theta]\}\,\mathrm{d}\theta \\ &= \tfrac{(-1)^m}{4\pi}\int_{2\pi} g[(1+\cos\psi)/2]\{\cos[(m+1)\psi] + \cos[m\psi]\}\,\mathrm{d}\psi. \end{aligned} \quad \text{(B.5)}$$

That is, b_m can be found by averaging adjacent Fourier coefficients from the Fourier series of the periodic and even function $g[(1+\cos\psi)/2]$. Further, because cosines are orthogonal over summation as well as integration, this opens striking new options. In particular, when only a finite number, say N, of these Fourier coefficients are non-zero, Eq. (B.5) can be computed exactly by using the discrete cosine transform (DCT) involving a sum over N points. Of course, the DCT can be accelerated by using the FFT in a variety of ways, see Sec 12.3 of [9].

#129305 - $15.00 USD Received 1 Jun 2010; revised 12 Jul 2010; accepted 12 Jul 2010; published 1 Sep 2010

As a minor variation, it is also possible to discretise the integral in Eq. (B.4) by sampling with a spacing of $\pi/(2N)$ in one of two obvious ways:

$$b_m \approx \frac{(-1)^m}{2N}\sum_{j=0}^{N-1}\left\{\frac{1}{1+\delta_{j0}}g[\cos^2(\tfrac{\pi}{2N}j)]\cos(\tfrac{\pi}{2N}j)\right\}\cos[\tfrac{\pi}{2N}(2m+1)j], \qquad \text{(B.6)}$$

Drop these factors of 2:

$$b_m \approx \frac{(-1)^m}{2N}\sum_{j=0}^{N-1}\left\{g[\cos^2\left(\tfrac{\pi}{4N}(2j+1)\right)]\cos\left(\tfrac{\pi}{4N}(2j+1)\right)\right\}\cos[\tfrac{\pi}{4N}(2j+1)(2m+1)]. \qquad \text{(B.7)}$$

These are both among the standard variants of the DCT and, because it has become a popular tool in image and signal processing, highly optimized code libraries are available to evaluate them. In particular, Eqs. (B.6) and (B.7) are instances of DCT-III and DCT-IV, respectively [10]. Like the DCTs derived from Eq. (B.5), these results are exact for band-limited functions; in other cases, the results are corrupted by the aliasing of the higher harmonics down to lower frequencies due to the sampling, see Sec 12.1 of [9]. In practice, the spectrum typically falls to insignificant levels for modest numbers of terms. That is, the functions are so close to being band limited in most cases that the aliasing is of no significance even for moderate N.

#129305 - $15.00 USD Received 1 Jun 2010; revised 12 Jul 2010; accepted 12 Jul 2010; published 1 Sep 2010