ORTHOGONAL POLYNOMIALS

FRANCIS BAHK

Abstract. In recent years the study of fitting orthogonal polynomials to freeform surfaces has received much attention. In this report we will detail how the Legendre and Chebyshev polynomials were derived and compare their properties. In the pursuit for polynomials orthogonal over elliptical apertures, we will show our attempted derivation.

Contents

1. Introduction

Optics is the study of light. Scientists have historically used symmetrical surfaces to study light and its properties, but new research shows that manufacturing optical components using asymmetrical surfaces — also called freeform surfaces — is promising. By introducing a freeform surface in an optical system, designers can balance the system's optical aberrations. This field is called freeform optics.

However, precisely manufacturing freeform surfaces according to the specifications of clients is difficult with standard machines. Instead, optical designers interpolate a point cloud of the desired freeform surface using polynomials, and this mathematical description can then be interpreted by the special machine that manufactures freeform surfaces.

To minimize the interpolation error, designers use orthogonal polynomials, a set of polynomials such that any two unique polynomials in the set are orthogonal to each other under some inner product. It is common to use Legendre or Chebyshev polynomials for surfaces defined over rectangular apertures.

Date: August 10, 2023.

More precisely, suppose for each point (x, y) on the Euclidean plane, we approximate $f(x, y)$ using a weighted sum of the basis functions. The interpolating polynomials are given to a state-of-the-art diamond turning tool that fabricates freeform surfaces. However, the fabrication process introduces tiny ridges in the surface, and certain kinds of these fabrication errors cause specific aberration patterns. Understanding the mathematical properties of orthogonal polynomials is therefore useful for finding relationships between fabrication error and measured aberrations.

2. Orthogonal Polynomials

Consider a continuous function $f(x)$ on the interval $[a, b]$ and suppose we have sample data $(x_i, f(x_i))_i$ with which to interpolate f. According to the Weierstrass approximation theorem, f can be uniformly approximated as closely as desired by a polynomial function. A natural idea is to use orthogonal polynomials since they have nice properties for minimax interpolation. In this section, we will re-derive and state the properties of Legendre and Chebyshev polynomials. In Section 4 we will compare their rates of convergence and accuracies.

Let $\mathcal{P}(\mathbb{R}^n)$ denote the polynomial space on $\mathbb{R}^n$. Suppose we equip $\mathcal{P}(\mathbb{R}^n)$ with the inner product $\langle f,\ g\rangle$ defined by

$$\int_a^b f(x)g(x)\omega(x)\,dx,$$

where ω is a non-negative weight function that assigns a certain importance to each portion of the domain. Then the Gram-Schmidt process gives us a method to generate an orthogonal basis for $\mathcal{P}(\mathbb{R}^n)$.

Definition 2.1. An orthogonal polynomial is a sequence of polynomials $P_n(x)$ of degree n satisfying

$$\int_a^b P_n(x)P_m(x)\omega(x)\,dx = \begin{cases} c_n \neq 0 & \text{if } m = n \\ 0 & \text{if } m \neq n \end{cases}$$

where c_n is a scalar.

The weight function is often of the form $\omega(x) = (1-x)^{\alpha}(1+x)^{\beta}$. For $\alpha = \beta = 0$, the Gram-Schmidt process on the resulting inner product space generates Legendre polynomials, and similarly for $\alpha = \beta = -\frac{1}{2}$ it generates Chebyshev polynomials. Orthogonal polynomials satisfy the recurrence relation

$$P_{n+1}(x) = (a_n(x) + b_n)P_n(x) + c_n P_{n-1}(x), \tag{2.2}$$

where $a_n(x)$, b_n, and c_n are real-valued functions.

2.1. **Chebyshev Properties.**

We will now define some properties of Chebyshev polynomials of the First Kind. Let $T_n(x)$ denote the nth Chebyshev polynomial.

Definition 2.3. Chebyshev polynomials are orthogonal with respect to $\omega(x) = (1-x^2)^{-\frac{1}{2}}$ over $[-1, 1]$ and satisfy

$$\int_{-1}^{1} T_n(x)T_m(x)\,\frac{dx}{\sqrt{1-x^2}} = \begin{cases} 0 & \text{if } n \neq m \\ \pi & \text{if } n = m = 0 \\ \frac{\pi}{2} & \text{if } n = m \neq 0. \end{cases}$$

Definition 2.4. $T_n(x) = \cos(n\cos^{-1}(x))$.

For each $x \in [-1, 1]$, there exists an angle $\theta \in [0, \pi]$ such that $x = \cos\theta$.

Definition 2.5. $T_n(x) = \cos(n\theta)$.

Theorem 2.6. *$T_n(x)$ is a polynomial of degree n.*

Proof. We will first prove the following lemma about an expansion of $\cos(n\theta)$.

Lemma 2.7.

$$\cos(n\theta) = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \binom{n}{2k} \cos^{n-2k}\theta \sin^{2k}\theta.$$

Proof. De Moivre's formula states $e^{i\theta} = \cos\theta + i\sin\theta$, so taking both sides to the power of n yields $(e^{i\theta})^n = (\cos\theta + i\sin\theta)^n$. The binomial formula states

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k. \tag{2.8}$$

Letting $x = \cos\theta$ and $y = i\sin\theta$ in the binomial formula and using the identity $\cos n\theta + i\sin n\theta = (\cos\theta + i\sin\theta)^n$ leads to the following two equations

$$\begin{cases} \cos(n\theta) + i\sin(n\theta) = (\cos\theta + i\sin\theta)^n = \sum_{k=0}^{n} \binom{n}{k} \cos^{n-k}\theta i \sin^k\theta \\ \cos(n\theta) - i\sin(n\theta) = (\cos\theta - i\sin\theta)^n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \cos^{n-k}\theta i \sin^k\theta \end{cases}$$

whose sum is

$$2\cos(n\theta) =$$
$$\left[\binom{n}{0}\cos^n\theta(i\sin\theta)^0 + \binom{n}{1}\cos^{n-1}\theta(i\sin\theta)^1 + \binom{n}{2}\cos^{n-2}\theta(i\sin\theta)^2 + \cdots\right] +$$
$$\left[\binom{n}{0}\cos^n\theta(i\sin\theta)^0 - \binom{n}{1}\cos^{n-1}\theta(i\sin\theta)^1 + \binom{n}{2}\cos^{n-2}\theta(i\sin\theta)^2 - \cdots\right].$$

As shown above, corresponding terms in each summation have the same sign when k is even, and otherwise they are opposite signs. Therefore,

$$2\cos(n\theta) = 2\left[\binom{n}{0}\cos^n\theta(i\sin\theta)^0 + \binom{n}{2}\cos^{n-2}\theta(i\sin\theta)^2 + \cdots\right],$$

which implies

$$\cos(n\theta) = \binom{n}{0}\cos^n\theta(i\sin\theta)^0 + \binom{n}{2}\cos^{n-2}\theta(i\sin\theta)^2 + \cdots.$$

In other words,

$$\cos(n\theta) = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \binom{n}{2k} \cos^{n-2k}\theta \sin^{2k}\theta \quad \text{as needed.}$$

□

It remains to show the nth Chebyshev polynomial has degree n. The identity $\sin^2\theta = 1 - \cos^2\theta$ implies $\sin^{2k}\theta = (\sin^2\theta)^k = (1-\cos^2\theta)^k$. By the binomial theorem,

$$(1-\cos^2\theta)^k = \begin{cases} 1 + \cdots + \cos^{2k}\theta & \text{if } k \text{ even} \\ 1 + \cdots - \cos^{2k}\theta & \text{if } k \text{ odd.} \end{cases}$$

This means the degree of $(1-\cos^2\theta)^k$ is $2k$. If k is even, Lemma 2.7 permits us to write

$$\cos(n\theta) = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \binom{n}{2k} (\cos^{n-2k}\theta(1+\cdots+\cos^{2k}\theta)),$$

which after expanding leads to

$$\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \binom{n}{2k} (\cos^{n-2k}\theta + \cdots + (\cos^{n-2k}\theta)(\cos^{2k}\theta))$$
$$= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \binom{n}{2k} (\cos^{n-2k}\theta + \cdots + \cos^n\theta).$$

Letting $x = \cos\theta$ in the above equation yields

$$T_n(x) = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \binom{n}{2k} (x^{n-2k} + \cdots + x^n),$$

which is a polynomial of degree n. The same argument can be applied when k is odd, since the sign of a term does not affect the power of the product between itself and another term, completing the proof. □

Theorem 2.9. *Chebyshev polynomials satisfy the recurrence relation*

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x) \quad \textit{for } n \geq 2.$$

Proof. Recall the trigonometric identities $\cos(\alpha+\beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$ and $\cos(\alpha-\beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta)$. We know $\cos(n\theta) = \cos(\theta + \theta(n-1))$, which is equivalent to $\cos\theta\cos\theta(n-1) - \sin\theta\sin\theta(n-1)$ by the first identity with $\alpha = \theta$ and $\beta = \theta(n-1)$. Therefore,

$$\cos(n\theta) = \cos\theta\cos\theta(n-1) - \sin\theta\sin\theta(n-1). \tag{2.10}$$

In addition, $\cos(n-2)\theta = \cos(\theta(n-1)-\theta)$, which by the second identity is equivalent to $\cos\theta(n-1)\cos\theta + \sin\theta(n-1)\sin\theta$. Rearranging yields

$$\sin\theta\sin\theta(n-1) = \cos(n-2)\theta - \cos\theta\cos\theta(n-1),$$

which we can substitute into (2.10) and simplify to yield

$$\cos(n\theta) = 2\cos(n-1)\theta - \cos(n-2)\theta.$$

Recurrence then follows from setting $x = \cos\theta$ as needed. □

Theorem 2.9 allows us to the recursively generate the nth Chebyshev polynomial starting $T_0 = 1$ and $T_1 = x$. The next several are

$$\{T_2, T_3, T_4, T_5\} = \{2x^2 - 1, 4x^3 - 3x, 8x^4 - 8x^2 + 1, 16x^5 - 20x^3 + 5x\}.$$

Theorem 2.11. *The leading coefficient of T_n is 2^{n-1} for $n > 0$.*

Proof. We present a proof by induction starting with the base case $n = 1$, which holds since $T_1 = x$. Let the leading coefficient of T_n be 2^{n-1}. $T_{n+1} = 2T_n - T_{n-1}$ and the inductive hypothesis implies the degree of T_{n-1} is $n-2$, so the leading coefficient of T_{n-1} does not influence that of T_n, which is 2^{n-1}. Hence $2 \cdot 2^{n-1} = 2^{1+(n-1)}$ is the leading coefficient of T_{n+1}. □

Freeform surfaces are defined in three dimensions, which motivates a definition for 2-D Chebyshev polynomials.

Definition 2.12. The jth 2-D Chebyshev polynomial $F_j(x,y)$ is equivalent to $T_n(x)T_m(y)$, where j depends on n and m.

Definition 2.13. $F_j(x,y)$ is a sequence of polynomials over $[-1,1]$ satisfying

$$\int_{-1}^{1}\int_{-1}^{1} F_j(x,y)F_{j^*}(x,y)\,\frac{dx}{\sqrt{1-x^2}}\,\frac{dy}{\sqrt{1-y^2}} = \begin{cases} 0 & \text{if } j \neq j^* \\ \pi^2 & \text{if } n=m=0 \\ \frac{\pi^2}{4} & \text{if } n=m\neq 0 \\ \frac{\pi^2}{2} & \text{otherwise.} \end{cases}$$

2-D Chebyshev polynomials are easily determined since they are the product of the 1-D versions. For example $\{F_0, F_1, F_2, F_3, F_4, F_5, F_6, F_7\}$, which is determined by $\{T_0, T_1, T_2\}$, is

$$\{1, x, y, 2x^2-1, xy, 2y^2-1, 4x^3-3x, (2x^2-1)y\}.$$

2.2. **Théorie des Mécanismes.** Chebyshev polynomials were introduced in 1854 by Chebyshev in *Théorie des mécanismes connus sous le nom de parallélogrammes.* My goals when writing an exposition of the original paper were to precisely interpret the French language, remove redundancy wherever possible, clarify motives, preserve order of ideas, and introduce new mathematics to corroborate the paper. Hence I have neither taken credit for any original mathematics below nor have I uncritically paraphrased Chebyshev.

Chebyshev used linkages to develop the theory of approximation of functions. A mechanical linkage is a set of rigid bars connected to each other at their ends via revolving hinges. For example, a human arm is a mechanical linkage.

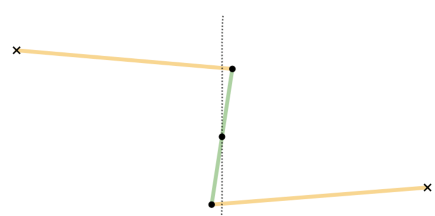

FIGURE 1. Watt's linkage.[1]

Linkages are useful because they produce certain kinds of motion. In Figure 1, let the green line be the "middle bar" and the yellow lines the "rockers." The points marked x are fixed while the rest are not. In Figure 2, Watt's linkage is defined by the points $ADMFG$, where D and F are endpoints on the middle bar and M is an arbitrary point on the middle bar. If M is located exactly in the middle of D and F and if the length of the rockers $\overline{AD}$ and $\overline{FG}$, are the same, then M moves along the vertical dotted line in Figure 1. Otherwise, M follows a path called Watt's curve.[2]

[1]See Wikipedia's Watt's linkage page for animation of Figure 1.
[2]See Wikipedia's Watt's curve page.

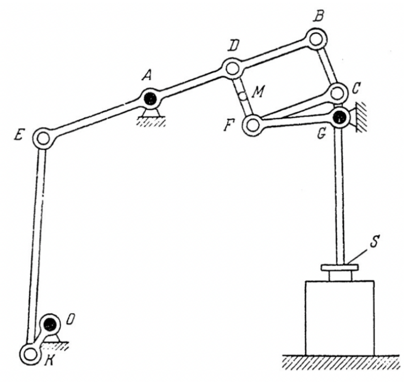

FIGURE 2. Watt's parallelogram in the steam engine.[3]

Chebyshev was interested in a specific mechanism called Watt's parallelogram, created by combining Watt's linkage with a pantograph, which is a linkage of parallelograms that produce identical movements in various parts of the linkage.

The points O, A, and G in Figure 2 are fixed and the rest are variable. Assume we are moving K in a circle about O. This circular motion rocks E and B in such a way that M and C move rectilinearly, which causes the piston S to move rectilinearly relative to a vertical line. Therefore, the steam machine is generally good at converting circular motion into rectilinear motion.

However, Chebyshev found two problems with this machine. First, if $\|\overline{AD}\| \neq \|\overline{FG}\|$, then the optimal choice for M was unknown. Second, even if M were located optimally, the linear motion at S might not be sufficiently rectilinear. To solve problems such as finding optimal proportions for Watt's mechanism, Chebyshev formulated the problem theoretically. More explicitly, Chebyshev explicitly wanted to determine the changes to make to the approximation of $f(x)$ in order to minimize the maximum approximation error in $[a-h, a+h]$ for small h.

Suppose $f(x)$ is a given function of degree n. Let $U(x)$ with arbitrary coefficients ζ_U be a polynomial of degree n which approximates f. Suppose f and U are defined on $[a-h, a+h]$ for small positive h. We want to determine ζ_U using the minimax criterion. Suppose the value of $|f(x) - U(x)|$ is considered maximum at $a-h$ and $a+h$ and we have chosen ζ_U such that the maximum of $f - U$ is minimized in $(a-h, a+h)$. Then $|f-U|$ reaches its maximum value δ_U at least $n+2$ times over $[a-h, a+h]$, which implies $|f-U| = \delta_U$ at $n+2$ points. Therefore, the equation

$$|f(x) - U(x)|^2 - \delta^2 = 0 \tag{2.14}$$

has $n+2$ roots in $[a-h, a+h]$ that also satisfy

$$(x-a+h)(x-a-h)\frac{d}{dx}(f(x) - U(x)) = 0. \tag{2.15}$$

Also let us simplify the Taylor series of f at a by letting $k_n = \frac{f^n(a)}{n!}$ and $hz = (x-a)$. The Taylor series of f at a is therefore $k_0 + k_1hz + h_2h^2z^2 + \cdots$ where $z \in [-1, 1]$. Recall we want to determine ζ_U so that we minimize δ_U. Let the difference $f - U$

[3] From https://bhavana.org.in/math-and-motion-a-look-at-chebyshevs-works-on-linkages.

be Y. We replace f with its Taylor series to yield

$$k_0 + k_1 hz + k_2 h^2 z^2 + \cdots - U = Y. \tag{2.16}$$

Suppose our degree of precision in the Taylor series is n, meaning we delete terms of order greater than n. Since U must have degree n, it is possible to reduce Y to 0 which implies

$$U = k_0 + k_1 hz + k_2 h^2 z^2 + \cdots + k_n h^n z^n \tag{2.17}$$

and

$$k_{n+1} = k_{n+2} = \cdots = k_{n+m} = 0, \tag{2.18}$$

where we have introduced m to index the zero terms. We will introduce V, an arbitrary polynomial of degree n, whose coefficients are finite when $h = 0$. Then

$$U = U_0 + Vh^{n+m+1}, \tag{2.19}$$

where $U_0 = k_0 + k_1 hz + k_2 h^2 z^2 + \cdots + k_n h^n z^n$. Therefore, we can rewrite (2.16) as

$$\begin{aligned} k_0 + k_1 hz + k_2 h^2 z^2 + \cdots - (k_0 + k_1 hz + k_2 h^2 z^2 + \cdots + k_n h^n z^n + Vh^{n+m+1}) \\ = (k_{n+m+1} h^{n+m+1} z^{n+m+1} + k_{n+m+2} h^{n+m+2} z^{n+m+2} + \cdots - Vh^{n+m+1}) \\ = (k_{n+m+1} z^{n+m+1} + k_{n+m+2} hz^{n+m+2} + \cdots - V)h^{n+m+1} = Y. \end{aligned}$$

In the above expression, suppose we discard h^{n+m+1} and consider terms not containing $h, h^2, h^3, \ldots$. Then

$$Y = k_{n+m+1} z^{n+m+1} - V,$$

where V is determined to minimize $k_{n+m+1} z^{n+m+1} - V$ for $z \in [-1, 1]$ among all other polynomials of the same degree. We wrote the equations (2.14) and (2.15) to minimize δ_U. Similarly, V is of degree n, so there exist $n+2$ roots to the equations

$$\begin{cases} (k_{n+m+1} z^{n+m+1} - V)^2 - \delta_V^2 = 0 \\ (z^2 - 1)) \frac{d}{dz}(k_{n+m+1} z^{n+m+1} - V) = 0. \end{cases}$$

Abbreviate the pair of equations above by letting $L = \delta_V$ and

$$y = k_{n+m+1} z^{n+m+1} - V. \tag{2.20}$$

Hence the following equations share $n + 2$ roots.

$$\begin{cases} y^2 - L^2 = 0 \\ (z^2 - 1) \frac{dy}{dz} = 0. \end{cases} \tag{2.21}$$

Setting the two equations above equal to each other and rearranging implies

$$\frac{y^2 - L^2}{(\frac{dy}{dz})^2} = \frac{P(z^2 - 1)}{Q^2}.$$

Proof. Let $\{z_1, z_2, \ldots, z_n, z_{n+1}, z_{n+2}\}$ be the roots of (2.21). Suppose $z_{n+1} = -1$ and $z_{n+2} = 1$ satisfy $z^2 - 1 = 0$. Call these *edge roots*. The remaining roots $\{z_1, z_2, \ldots, z_n\}$ must satisfy $\frac{dy}{dz} = 0$ so let them be *normal roots*. If z is a normal root, say z_1, then we know that $(z - z_1)$ divides the first equation in (2.21) and $\frac{dy}{dz}$ so $(z - z_1)^2$ divides $(\frac{dy}{dz})^2$. Thus the product $(z - z_1)^2 (z - z_2)^2 \cdots (z - z_n)^2$ divides the first equation and $(\frac{dy}{dz})^2$. Otherwise, z is an edge root so $(z + 1)(z - 1)$ divides the second equation.

Therefore the degree of the first equation in (2.21) is $2(m+n+1)-2n = 2m+2$ and the degree of $\frac{dy}{dz}$ is $(m+n+1-1)-n = m$. These degrees imply the degrees of P and Q are $2m$ and m as needed. According to [[1], p. 9], the differential equation holds even if one or both edge roots satisfy $\frac{dy}{dz} = 0$. □

We rearrange the differential equation to find

$$\frac{dy}{\sqrt{y^2 - L^2}} = \frac{Qdz}{\sqrt{P(z^2-1)}}.$$

The integral of the left side is

$$\frac{1}{2}\log\frac{y+\sqrt{y^2-L^2}}{y-\sqrt{y^2-L^2}},$$

so the integral of the right side must be of the form

$$\frac{1}{2}\log\frac{p+q\sqrt{R}}{p-p\sqrt{R}},$$

where p must also have degree $n+m+1$. Since V is of degree n, we know y and p cannot have terms containing $z^{n+1}, \ldots, z^{n+m-1}, z^{n+m}$.

We now solve for y based on the value of m. If $m = 0$, then Q and P become constants in the integral so

$$\frac{dy}{\sqrt{y^2 - L^2}} = \lambda\frac{dz}{\sqrt{(z^2-1)}}.$$

The above equation implies

$$y = \pm\frac{L}{2}\left[\left(z+\sqrt{z^2-1}\right)^{\lambda} + \left(z-\sqrt{z^2-1}\right)^{\lambda}\right].$$

To determine the values of L and λ, notice that y has degree $n+1$. According to (2.20), $y = k_{n+1}z^{n+1} - V$ so $\lambda = n+1$. Moreover, we know $\pm\frac{L}{2} = \pm\frac{k_{n+1}}{2^{n+1}}$ implies $L = \pm\frac{k_{n+1}}{2^n}$, so $L = \pm\frac{k_{n+1}}{2^{\lambda-1}}$. Altogether,

$$\lambda = n+1, \quad L = \pm\frac{k_{n+1}}{2^{\lambda-1}} = \pm\frac{k_{n+1}}{2^n}. \tag{2.22}$$

These values of λ and L imply

$$y = \frac{k_{n+1}}{2^{n+1}}\left[\left(z+\sqrt{z^2-1}\right)^{n+1} + \left(z-\sqrt{z^2-1}\right)^{n+1}\right] \tag{2.23}$$

and according to (2.20),

$$V = k_{n+1}z^{n+1} - \frac{k_{n+1}}{2^{n+1}}\left[\left(z+\sqrt{z^2-1}\right)^{n+1} + \left(z-\sqrt{z^2-1}\right)^{n+1}\right]$$

which can be simplified to

$$V = k_{n+1}\left[z^{n+1} - \left(\frac{z+\sqrt{z^2-1}}{2}\right)^{n+1} - \left(\frac{z-\sqrt{z^2-1}}{2}\right)^{n+1}\right]. \tag{2.24}$$

If $m = 1$, then (2.20) implies $y = k_{n+2}z^{n+2} - V$. However, V is of degree n so y cannot contain powers of z greater than $n+1$ besides $k_{n+2}z^{n+2}$. According to

(2.23), $y = k_{n+2}z^{n+2} - V$ is the polynomial that deviates least from 0 among all polynomials of degree n for $z \in [-1, 1]$ and V is

$$V = k_{n+2}\left[z^{n+2} - \left(\frac{z+\sqrt{z^2-1}}{2}\right)^{n+2} - \left(\frac{z-\sqrt{z^2-1}}{2}\right)^{n+2}\right].$$

Finally, if $m > 1$, then we can find V with $2m$ equations according to (2.20) and explained on [[1], p. 13]. Therefore, for any m, we can find V.

We learned the polynomial g of degree n that minimizes $\max|f(x) - g(x)|$ for $|x| \leq a + h$ with $f^{(n+1)}$ nonzero takes the form $g = U + Vh^{n+1}$, where $U = k_0 + \cdots + k_n h^n z^n$ and V is the polynomial of degree n that deviates least from $k_{n+1}z^{n+1} + k_{n+2}hz^{n+2} + \cdots$ compared to all other polynomials of degree n for $|z| \leq 1$. We also learned

$$V = k_{n+1}\left[z^{n+1} - \left(\frac{z+\sqrt{z^2-1}}{2}\right)^{n+1} - \left(\frac{z-\sqrt{z^2-1}}{2}\right)^{n+1}\right].$$

Thus (2.20) implies

$$\begin{aligned} y &= k_{n+1}z^{n+1} - V \\ &= k_{n+1}z^{n+1} - \left(k_{n+1}\left[z^{n+1} - \left(\frac{z+\sqrt{z^2-1}}{2}\right)^{n+1} - \left(\frac{z-\sqrt{z^2-1}}{2}\right)^{n+1}\right]\right) \\ &= k_{n+1}\left[\left(\frac{z+\sqrt{z^2-1}}{2}\right)^{n+1} + \left(\frac{z-\sqrt{z^2-1}}{2}\right)^{n+1}\right]. \end{aligned}$$

We will now show that V can be found for any desired accuracy.

Remark 2.25. For an immediate derivation of Chebyshev nodes and polynomials, see Derivation 2.3.

If V_0 is a polynomial with finite coefficients when $h = 0$, then

$$V = k_{n+1}z^{n+1} - y + V_0 h. \tag{2.26}$$

The maximum approximation error that the coefficients ζ_V minimize is

$$\begin{aligned} &k_{n+1}z^{n+1} + k_{n+2}hz^{n+2} + \cdots - V \\ &= k_{n+2}hz^{n+2} + \cdots + y - V_0 h, \end{aligned}$$

where $|z| \leq 1$. Hence

$$\left[k_{n+2}hz^{n+2} + y - V_0 h\right]^2 - L^2 = 0 \tag{2.27}$$

and

$$(z^2-1)\frac{d}{dz}\left(k_{n+2}hz^{n+2} + y - V_0 h\right) = 0 \tag{2.28}$$

have $n + 2$ common roots. We can rewrite (2.28) as

$$(z^2-1)\frac{dy}{dz} + (z^2-1)\frac{d}{dz}(k_{n+2}hz^{n+2} - V_0)h = 0$$

implying that for sufficiently small h, all its roots living in $[-1, 1]$ must satisfy $(z^2 - 1)\frac{dy}{dz}$. Also, (2.27) can be written as

$$(y^2 - L_1^2)y + 2y^2(k_{n+2}z^{n+2} - V_0)h = 0.$$

However, we showed that

$$y = k_{n+1}\left[\left(\frac{z+\sqrt{z^2-1}}{2}\right)^{n+1} + \left(\frac{z-\sqrt{z^2-1}}{2}\right)^{n+1}\right]$$

must satisfy $y^2 - L^2 = 0$, so (2.27) is equivalent to

$$(L^2 - L_1^2)y + 2L^2(k_{n+2}z^{n+2} - V_0)h = 0.$$

This equation has degree $n+2$ because the degrees of y and V_0 are $n+1$ and n respectively, which are both less than $n+2$. The new expressions for (2.27) and (2.28) both have degree $n+2$ so they are multiples of each other. Hence,

$$(L^2 - L_1^2)y + 2L^2(k_{n+2}z^{n+2} - V_0)h - C((z^2-1)\frac{dy}{dz}) = 0$$

for some constant C. Since y does not contain a term containing z^n and the degree of V_0 is at most n, the coefficient of z^{n+1} in the first term cannot equal 0 unless $L^2 - L_1^2 = 0$. Under this assumption, (2.22) implies

$$L_1 = \pm\frac{k_{n+1}}{z^n}, \tag{2.29}$$

which implies

$$V_0 = k_{n+2}z^{n+2} - \frac{C}{2hL^2}(z^2-1)\frac{dy}{dz} = 0.$$

Since the degree of $\frac{dy}{dz}$ must be $(n+1)-1 = n$, there exists a term of the form $(n+1)k_{n+1}z^{n+2}$ in the expansion of $(z^2-1)\frac{dy}{dz}$. Factoring out z^{n+2} yields

$$V_0 = z^{n+2}\left(k_{n+2} - \frac{(n+1)Ck_{n+1}}{2hL^2}\right) = 0.$$

However $z^{n+2} \neq 0$ over our entire domain, so

$$k_{n+2} - \frac{(n+1)Ck_{n+1}}{2hL^2} = 0 \implies \frac{C}{2hL^2} = \frac{k_{n+2}}{(n+1)k_{n+1}}.$$

Substituting this into the expression for V_0 implies

$$V_0 = k_{n+2}z^{n+2} - \frac{k_{n+2}}{(n+1)k_{n+1}}(z^2-1)\frac{dy}{dz} = k_{n+2}\left(z^{n+2} - \frac{(z^2-1)}{(n+1)k_{n+1}}\frac{dy}{dz}\right) = 0.$$

We can finally write according to (2.26) an expression for V, which is

$$V = k_{n+1}z^{n+1} - y + k_{n+2}\left(z^{n+2} - \frac{z^2-1}{(n+1)k_{n+1}}\frac{dy}{dz}\right)h. \tag{2.30}$$

Theorem 2.31. *Assume V is given up to (but not including) terms of order l. Then, V can be found up to (but not including) terms of order $2l$.*

Proof. Let V_1 denote the given V, which is of the form (2.30). Then,

$$V = V_1 + V_2h^l,$$

is an equation for finding the new value of V, where V_2 is of degree n. Finding V_2 is straightforward because we know it satisfies the condition that the equations

$$[k_{n+1}z^{n+1} + k_{n+2}hz^{n+2} + \cdots - V_1 - V_2h^l] - L_2^2 = 0$$

$$(z^2-1)\frac{d}{dz}(k_{n+1}z^{n+1} + k_{n+2}hz^{n+2} + \cdots - V_1 - V_2h^l) = 0$$

share $n+2$ roots in $[-1,1]$. Since V only requires precision up to terms of order $2l$, we remove terms containing $h^{2l}, h^{2l+1}, h^{2l+2}, + \cdots$ to find the system of equations

$$\begin{cases} [y_1 + Sh^l - V_2h^l]^2 - L_2^2 = 0 & \text{(2.33A)} \\ (z^2-1)\frac{d}{dz}(y_1 + Sh^l - V_2h^l) = 0 & \text{(2.33B)}, \end{cases} \tag{2.32}$$

where

$$\begin{cases} y_1 = k_{n+1}z^{n+1} + k_{n+2}hz^{n+2} + \cdots + k_{n+l}h^{l-1}z^{n+1} - V_1 \\ S = k_{n+l+1}z^{n+l+1} + k_{n+l+2}hz^{n+l+2} + \cdots + k_{n+2l}h^{l-1}z^{n+2l}, \end{cases} \tag{2.33}$$

which can be reduced to

$$\begin{cases} y_1 = \left(\sum_{p=n+1}^{n+l} k_p h^{p-n-1} z^p\right) - V_1 \\ S = \sum_{p=n+l+1}^{n+2l} k_p h^{p-n-l-1} z^p. \end{cases}$$

Since y_1 should satisfy

$$\begin{cases} y_1^2 - L_1^2 = 0 \\ (z^2-1)\frac{dy_1}{dz} = 0, \end{cases}$$

we know

$$L_2 = L_1 + \lambda h^l. \tag{2.34}$$

Also, y_1 must satisfy $(z^2-1)\frac{dy_1}{dz} = 0$ in order to satisfy (2.32B) which allows us to replace (2.32B) with the prerequisite. For simplicity, also replace $\frac{dy_1}{dz}$ with some function W of degree n that contains the same roots as $\frac{dy_1}{dz}$ the interval $[-1,1]$. Altogether we now have an equivalent form for (2.32B), which is

$$(z^2-1)W = 0.$$

The above equation has degree $n+2$ and shares all of its roots with $z^2 - 1\frac{dy_1}{dz} = 0$, so $(z^2-1)W = 0$ satisfies

$$\begin{cases} y_1^2 - L_1^2 = 0 \\ (2.32A). \end{cases}$$

Substituting (2.34) into (2.32A) implies

$$\begin{aligned} &[y_1 + Sh^l - V_2h^l]^2 - L_2^2 \\ &= [y_1 + (S - V_2)h^l]^2 - (L_1 + \lambda h^l)^2 \\ &= y_1^2 + 2y_1(S-V_2)h^l + (S-V_2)^2h^{2l} - (L_1^2 + 2L_1\lambda h^l - \lambda^2 h^{2l}) \\ &= y_1^2 + 2y_1(S-V_2)h^l + (S-V_2)^2h^{2l} - L_1^2 - 2L_1\lambda h^l + \lambda^2 h^{2l} = 0. \end{aligned}$$

Removing terms containing h^{2l} leaves us with

$$y_1^2 + 2y_1(S-V_2)h^l - L_1^2 - 2L_1\lambda h^l = 0,$$

multiplying by y_1 and substituting for L_1 yields

$$(y_1^2 - L_1^2)y_1 + 2L_1^2(S-V_2)h^l - 2\lambda L_1 h^l y_1 = 0,$$

and rearranging gives

$$V_2h^l + \frac{\lambda h^l}{L_1}y_1 - Sh^l - \frac{(y_1^2 - L_1^2)}{2L_1^2}y_1 = 0.$$

We are sure that this equation and $(z^2-1)W = 0$ share $n+2$ roots, so V_2h^l is divisible by $(z^2-1)W$. Thus,

$$\frac{V_2h^l + \frac{\lambda h^l}{L_1}y_1 - Sh^l - \frac{(y_1^2-L_1^2)y_1}{2L_1^2} = 0}{(z^2-1)W} = V_2h^l + \frac{\lambda}{L_1}R_0 - R_1 = 0,$$

where

$$R_0 = \frac{y_1h^l}{(z^2-1)W} \quad \text{and} \quad R_1 = \frac{Sh^l + \frac{(y_1^2-L_1^2)y_1}{2L_1^2}}{(z^2-1)W}.$$

This implies $V_2h^l = R_1 - \frac{\lambda}{L_1}R_0$. Dividing R_1 by R_0 yields $R_1 = r + qR_0$ where r is the remainder and q the quotient, so

$$V_2h^l = (r+qR_0) - \frac{\lambda}{L_1}R_0 = r + (q - \frac{\lambda}{L_1})R_0.$$

Letting $q = \frac{\lambda}{L_1}$ implies $V_2h^l = r$, which is the condition under which we can determine V_2h^l and V. $\lambda = L_1q$ so the value of L_2 according to (2.34) is

$$L_2 = L_1 + \lambda h^l = L_1 + (L_1q)h^l = L_1(1+qh^l).$$

We now know in (2.32A) the value of L_2, which is

$$\max|k_{n+1}z^{n+1} + k_{n+2}hz^{n+2} + k_{n+3}h^2z^{n+3} + \cdots - V|$$

for $z \in [-1,1]$. Ultimately, the equation $V = V_1 + V_2h^l$ implies that given some approximation of V, we can always find a more precise approximation of V. □

We will now solve the initial parallelogram problem. The goal of the problem is to determine the coefficients of the degree 4 Taylor series approximation of $f(x)$ at $x = a$ that minimize the difference between the approximation and $f(x)$ in the interval $[a-h, a+h]$. Assuming h is small enough and $f^5(a) \neq 0$, the desired coefficients are given by $Vh^{n+1} = Vh^5$, where V is a function of $z = \frac{x-a}{h}$ and is chosen to minimize

$$\max|k_{n+1}z^{n+1} + k_{n+2}hz^{n+2} + k_{n+3}h^2z^{n+3} + \cdots - V|.$$

According to (2.30),

$$V = k_{n+1}z^{n+1} - y + k_{n+2}\left(z^{n+2} - \frac{z^2-1}{(n+1)k_{n+1}}\frac{dy}{dz}\right)h,$$

where

$$y = k_{n+1}\left[\left(\frac{z+\sqrt{z^2-1}}{2}\right)^{n+1} - \left(\frac{z-\sqrt{z^2-1}}{2}\right)^{n+1}\right].$$

For $n = 4$,

$$V = k_5z^5 - y + k_6\left(z^6 - \frac{z^2-1}{5k_5}\frac{dy}{dz}\right)h$$

and

$$y = k_5\left[\left(\frac{z+\sqrt{z^2-1}}{2}\right)^5 - \left(\frac{z-\sqrt{z^2-1}}{2}\right)^5\right] = k_5\left(z^5 - \frac{5}{4}z^3 + \frac{5}{16}z\right).$$

Thus,

$$V = k_5z^5 - \left[k_5\left(z^5 - \frac{5}{4}z^3 + \frac{5}{16}z\right)\right] + k_6\left(z^6 - \frac{z^2-1}{5k_5}\frac{dy}{dz}\right)h$$

$$= k_5\left(\frac{5}{4}z^3 - \frac{5}{16}z\right) + k_6\left(\frac{7}{4}z^4 - \frac{13}{16}z^2 + \frac{1}{16}\right)h.$$

Finally, (2.29) implies

$$|L_1| = \left|\frac{k_{n+1}}{2^n}\right| = \left|\frac{k_5}{16}\right|.$$

With this information, we can find the exact value of V up to h^4 according to Theorem 2.31 since we know V up to h^2. Thus $l = 2$ so $V = V_1 + V_2 h^2$. Assigning to V_1 the value of V implies

$$\begin{aligned} y_1 &= k_5 z^5 + k_6 h z^6 - V_1 \\ &= k_5 z^5 + k_6 h z^6 - \left[k_5\left(\frac{5}{4}z^3 - \frac{5}{16}z\right) + k_6\left(\frac{7}{4}z^4 - \frac{13}{16}z^2 + \frac{1}{16}\right)h\right] \\ &= k_5\left(z^5 - \frac{5}{4}z^3 + \frac{5}{16}z\right) + k_6\left(z^6 - \frac{7}{4}z^4 + \frac{13}{16}z^2 - \frac{1}{16}\right)h \\ &= y + k_6\left(z^6 - \frac{7}{4}z^4 + \frac{13}{16}z^2 - \frac{1}{16}\right)h, \end{aligned}$$

and $S = k_7 z^7 + k_8 h z^8$. We now determine V_2. The degree 4 equation W that contains the roots of

$$\frac{dy_1}{dz} = k_5\left(5z^4 - \frac{15}{4}z^2 + \frac{5}{16}\right) + k_6\left(6z^5 - 7z^3 + \frac{13}{8}z\right)h = 0$$

is determined by $V_2 h^l = r$, which is the remainder of the division

$$\frac{k_6\left(6z^5 - 7z^3 + \frac{13}{8}z\right)}{k_5\left(5z^4 - \frac{15}{4}z^2 + \frac{5}{16}\right)}.$$

The remainder is $r = k_6\left(-\frac{5}{2}z^3 + \frac{5}{4}z\right)h$, so

$$W = k_5\left(5z^4 - \frac{15}{4}z^2 + \frac{5}{16}\right) + k_6\left(-\frac{5}{2}z^3 + \frac{5}{4}z\right)h = 0.$$

Next, R_0 is the remainder of the division

$$\frac{y_1 h^2}{(z^2-1)W} = \frac{\left[k_5\left(z^5 - \frac{5}{4}z^3 + \frac{5}{16}z\right) + k_6\left(z^6 - \frac{7}{4}z^4 + \frac{13}{16}z^2 - \frac{1}{16}\right)h\right]h^2}{(z^2-1)\left[k_5\left(5z^4 - \frac{15}{4}z^2 + \frac{5}{16}\right) + k_6\left(-\frac{5}{2}z^3 + \frac{5}{4}z\right)h\right]},$$

so $R_0 = k_5\left(z_5 - \frac{5}{4}z^3 + \frac{5}{16}z\right)h^2 = yh^2$. Similarly, R_1 is the remainder of

$$\frac{Sh^l + \frac{(y_1^2 - L_1^2)y_1}{2L_1^2}}{(z^2-1)W} = \frac{Sh^2 + \frac{(y_1^2 - L_1^2)y_1}{2L_1^2}}{(z^2-1)\left[k_5\left(5z^4 - \frac{15}{4}z^2 + \frac{5}{16}\right) + k_6\left(-\frac{5}{2}z^3 + \frac{5}{4}z\right)h\right]}.$$

Hence $R_1 = [a_1 z^5 - a_2 z^3 + a_3 z]h^2 + [a_4 z^4 - a_5 z^2 + a_6]h^3$, where

$$\{a_1, a_2, a_3, a_4, a_5, a_6\} = \{\frac{7k_5k_7 + k_6^2}{4k_5}, \frac{13k_5k_7 + 6k_6^2}{16k_5}, \frac{k_5k_7 + 2k_6^2}{16k_5}$$

$$\frac{36k_5^2k_8 + 2k_5k_6k_7 - k_6^3}{16k_5^2}, \frac{87k_5^2k_8 + 10k_5k_6k_7 - 5k_6^3}{64k_5^2}, \frac{7k_5^2k_8 + 2k_5k_6k_7 - k_6^3}{64k_5^2}\}.$$

Notice

$$\frac{R_1}{R_0} = \frac{[a_1 z^5 - a_2 z^3 + a_3 z]h^2 + [a_4 z^4 - a_5 z^2 + a_6]h^3}{yh^2} = q + r.$$

Hence, $r = b_1 h^3 z^4 + b_2 h^2 z^3 - b_3 h^3 z^2 - b_4 h^2 z + b_5 h^3$, where

$$\{b_1, b_2, b_3, b_4, b_5\} = \{\frac{36k_5^2 k_8 + 2k_5 k_6 k_7 - k_6^3}{16k_5^2}, \frac{22k_5 k_7 - k_6^2}{16k_5}$$
$$\frac{87k_5^2 k_8 + 10k_5 k_6 k_7 - 5k_6^3}{64k_5^2}, \frac{31k_5 k_7 - 3k_6^2}{64k_5}, \frac{7k_5^2 k_8 + 2k_5 k_6 k_7 - k_6^3}{64k_5^2}\}.$$

Since $V_2 h^2 = r$ up to h^4, we can now determine V.

$$\begin{aligned} V = V_1 + V_2 h^2 = \\ \left[k_5\left(\frac{5}{4}z^3 - \frac{5}{16}z\right) + k_6\left(\frac{7}{4}z^4 - \frac{13}{16}z^2 + \frac{1}{16}\right)h\right] \\ + \left[b_1 h^3 z^4 + b_2 h^2 z^3 - b_3 h^3 z^2 - b_4 h^2 z + b_5 h^3\right] \\ = c_1 z^4 + c_2 z^3 - c_3 z^2 - c_4 z + c_5 h + c_6 h^3, \end{aligned}$$

where

$$\{c_1, c_2, c_3, c_4, c_5, c_6\} = \{\frac{7}{4}k_6 h + b_1, \frac{5}{4}k_5 + b_2, \frac{13}{16}k_6 h + b_3, \frac{5}{16}k_5 + b_4, \frac{1}{16}k_6, b_5\}.$$

The division $\frac{R_1}{R_0}$ also yields $q = \frac{7k_5k_7 + k_6^2}{4k_5^2}$. Hence, the value of L_2 must be

$$L_2 = L_1(1 + qh^l) = \left|\frac{k_5}{16}\right|\left(1 + \frac{7k_5 k_7 + k_6^2}{4k_5^2}h^2\right).$$

Finally, the substitution $x - a = hz$ implies

$$\begin{aligned} Vh^5 = \left[c_1 z^4 + c_2 z^3 - c_3 z^2 - c_4 z + c_5 h + c_6 h^3\right] h^5 \\ = (c_1 + \cdots)\left(\frac{x-a}{h}\right)^4 + (c_2 + \cdots)\left(\frac{x-a}{h}\right)^3 - (c_3 + \cdots)\left(\frac{x-a}{h}\right)^2 \\ -(c_4 + \cdots)\left(\frac{x-a}{h}\right) + (c_5 + \cdots)h^6 + (c_6 + \cdots)h^8, \end{aligned}$$

which are the coefficients of the degree four Taylor series that minimize the maximum deviation from $f(x)$ when $x \in [a-h, a+h]$. The maximum error is

$$|L_2 h^5| = \left|\frac{k_5}{16}\right|\left(1 + \frac{7k_5 k_7 + k_6^2}{4k_5^2}h^2\right)h^5.$$

2.3. **Chebyshev Derivation.** Suppose $f(x) = k_{n+1}x^{n+1}$. Let γ and α satisfy $[\alpha - \gamma, \alpha + \gamma] \subset [a-h, a+h]$. Suppose we are looking for a polynomial u of degree n that equals f when $|x| = a + \gamma$ and deviates least from f among all other polynomials of the same degree when $|x| < a + \gamma$. If $l = \max(f(x) - u(x))$ for $|x| < a + \gamma$, then the equations

$$(f(x) - u(x))^2 - l^2 = 0, \quad \frac{d(f(x) - u(x))}{dx} = 0$$

share n roots that are located in the interval $[a-\gamma, a+\gamma]$. Moreover, the equations

$$(f(x) - u(x))^2 - l^2 = 0, \quad (x - a + h)(x - a - h)\frac{d(f(x) - u(x))}{dx} = 0 \tag{2.35}$$

share $n + 2$ roots that are located in the interval $[a-h, a+h]$. Suppose

$$y = \frac{k_{n+1}x^{n+1} - u}{h^{n+1}}, \quad x = a + hz, \quad L = \frac{l}{h^{n+1}}.$$

Then, (2.35) is equivalent to the equations

$$y^2 - L^2 = 0, \quad (z^2-1)\frac{dy}{dz} = 0,$$

which share $n+2$ roots for $|z| \le 1$. Recall

$$y = k_{n+1}\left[\left(\frac{z+\sqrt{z^2-1}}{2}\right)^{n+1} + \left(\frac{z-\sqrt{z^2-1}}{2}\right)^{n+1}\right].$$

Substituting our supposed values for y, x, and L into the above equation implies $k_{n+1}x^{n+1} - u$ equals

$$k_{n+1}\left[\left(\frac{x-a+\sqrt{(x-a)^2-h^2}}{2}\right)^{n+1} + \left(\frac{x-a-\sqrt{(x-a)^2-h^2}}{2}\right)^{n+1}\right].$$

Suppose $|x| = a+\gamma$. Then, the above equation must be equal to zero. Substituting $x - a = h\cos\theta$ implies

$$\begin{aligned}
&\left(\frac{h\cos\theta+\sqrt{(h\cos\theta)^2-h^2}}{2}\right)^{n+1} + \left(\frac{h\cos\theta-\sqrt{(h\cos\theta)^2-h^2}}{2}\right)^{n+1} \\
&= \left(\frac{h\cos\theta+hi\sin\theta}{2}\right)^{n+1} + \left(\frac{h\cos\theta-hi\sin\theta}{2}\right)^{n+1} \\
&= \left(\frac{h}{2}\right)^{n+1}\left((\cos\theta+i\sin\theta)^{n+1} + (\cos\theta-i\sin\theta)^{n+1}\right) = 0.
\end{aligned}$$

Dividing both sides by $\left(\frac{h}{2}\right)^{n+1}$ and using Euler's formula twice implies

$$\left(e^{i\theta}\right)^{n+1} + \left(e^{-i\theta}\right)^{n+1} = \left(e^{i(n+1)\theta}\right) + \left(e^{-i(n+1)\theta}\right) = 2\cos(n+1)\theta = 0.$$

Finally, dividing by 2 yields $\cos(n+1)\theta = 0$. To find the values of x for which the above equation is true, we should find all θ for which

$$(n+1)\theta = \frac{\pi}{2} + \pi n.$$

Hence

$$\theta = \frac{2\pi n + \pi}{2n+2},$$

so

$$x = h\cos\left(\frac{2\pi m+\pi}{2n+2}\right) + a$$

where m is an integer. This formula gives the abscissa for which the interpolation error is zero for given values of h, a, and n. Recall that

$$\frac{d(k_{n+1}x^{n+1}-u)}{dx} = 0$$

which reduces to

$$\frac{\sin(n+1)\theta}{\sin\theta} = 0$$

has n roots for $x - a = h\cos\theta$. The roots to the above equation are understood for only two values which are

$$a - h\cos\frac{\pi}{2n+2}, \quad a + h\cos\frac{\pi}{2n+2}.$$

Hence, the assumed the value of x implies

$$a = \alpha, h = \frac{\gamma}{\cos \frac{\pi}{2n+2}}.$$

We will now derive the Chebyshev polynomials. Let a function $y = f(x)$ be given. We need to choose the n parameters of our interpolant $Y = F(x,h)$ in such a way that y and Y intersect at n intersection points lying in $[a-h, a+h]$ which minimize the maximum interpolation error in this interval. In other words, we need to find the optimal points of intersection between y and Y. F has n parameters, so F should have degree $n-1$. When $h = 0$, F should satisfy

$$F(a,0) = f(a), \frac{d}{dx}F(a,0) = \frac{d}{dx}f(a), \ldots, \frac{d^{n-1}}{dx^{n-1}}F(a,0) = \frac{d^{n-1}}{dx^{n-1}}f(a),$$

with the nth derivative of F being 0. Therefore,

$$\frac{d^n}{dx^n}F(x,h) = \frac{d^n}{dx^n}f(x) + k \tag{2.36}$$

for some constant k. Suppose the above derivatives remain continuous for h small but nonzero. Then the Taylor series $\mathcal{T}$ of $Y - y$ at a is of the form

$$c_0 + c_1(x-a) + c_2(x-a)^2 + \ldots + c_{n-1}(x-a)^{n-1} + c_n(x-a)^n.$$

Notice

$$\frac{d^n}{dx^n}Y - y = \frac{d^n}{dx^n}F(x,h) - \frac{d^n}{dx^n}f(x)$$

$$= \underbrace{\frac{d^n}{da^n}F(a,0) - \frac{d^n}{da^n}f(a)}_{N} + \underbrace{\frac{d^n}{dx^n}F(x,h) - \frac{d^n}{dx^n}f(x) - \frac{d^n}{da^n}F(a,0) - \frac{d^n}{da^n}f(a)}_{\psi(x)}$$

which implies

$$c_n = \frac{N + \psi(a+h)}{n!} = \frac{N}{n!}.$$

The existence of u now depends on whether $\mathcal{T}$ reduces to

$$c_n\left[\left(\frac{x-a+\sqrt{(x-a)^2-h^2}}{2}\right)^n + \left(\frac{x-a-\sqrt{(x-a)^2-h^2}}{2}\right)^n\right].$$

According to our previous work, the above expression for $\mathcal{T}$ reduces to $\cos n\theta$, whose roots are

$$\theta = \pi\left(\frac{2m+1}{2n}\right)$$

for any integer m. Since the domain is $[-1,1]$, h and a are 1 and 0 respectively, so

$$x = \cos\left(\pi\frac{2m+1}{2n}\right)$$

are the abscissa of the optimal interpolation nodes, which are the Chebyshev nodes and $T_n(\cos\theta) = \cos n\theta$ then follows from Definition 2.9.

2.4. **Legendre Properties.** This section is about Legendre polynomials of the First Kind.

Definition 2.37. Legendre polynomials are a sequence of polynomials $P_n(x)$ orthogonal over $[-1,1]$ with respect to the weight function $\omega(x) = 1$. Legendre polynomials satisfy

$$\int_{-1}^{1} P_n(x)P_m(x)\,dx = \begin{cases} 0 & \text{if } n \neq m \\ 1 & \text{if } n = m. \end{cases}$$

Legendre polynomials are an orthonormal basis for any real polynomial space. The first few Legendre polynomials are

$$\{P_0, P_1, P_2, P_3\} = \{1, x, \frac{3}{2}x^2 - \frac{1}{2}, \frac{5}{2}x^3 - \frac{3}{2}x\}.$$

2.5. **Legendre Derivation.** We will explain three important preliminaries before deriving Legendre polynomials.

Definition 2.38. A generating function G defined as $G(a_n; x) = \sum_{n=0}^{\infty} a_n x^n$ encodes a sequence of numbers (a_n) as the coefficients of a power series.

Remark 2.39. In $\mathbb{R}^3$, a multipole expansion is a power series whose first few terms provide a good approximation of a function that depends on angles.

Remark 2.40. The Newtonian potential function ϕ gives the gravitational potential at some point p in a Euclidean space due to a fixed point mass a.

According to [[5] p. 528], mathematicians Euler and Lagrange were the first to derive ϕ in the rectangular coordinate system, but in 1782 French polymath Pierre-Simon Laplace independently derived ϕ in both the rectangular and polar systems. At around the same time, Laplace's compatriot Adrien-Marie Legendre showed in his memoir "Sur l'attraction des sphéroïdes homogènes" published in *Mémoires de Mathématiques et de Physique* that the Newtonian potential given by Laplace is the generating function for the Legendre polynomials. Hence, many mathematicians — including Jacobi, Dirichlet, and Heine — agreed to attach the polynomials to Legendre since he developed the multipole expansion that led to them. The following derivation uses the method of [2] and shows the Legendre polynomials are the coefficients in the expansion of ϕ.

The gravitational potential at some point p in a Euclidean space due to a fixed point mass a is described by the Newtonian potential function $\phi_p(D) = \frac{-GM_a}{D}$. Here, G is the gravitational constant, M_a is the mass of a, and D is the distance between a and p. Since $-GM_a$ is constant, let us call it C_a. Thus, $\phi_p(D) = \frac{C_a}{D}$. Let the coordinates of a be (x_a, y_a, z_a) and p be (x, y, z). Then, the distance from a to p is given by

$$\sqrt{(x - x_a)^2 + (y - y_a)^2 + (z - z_a)^2}.$$

Also, let $\vec{Oa}$ and $\vec{Op}$ be vectors from the origin to a and to p. The length of $\vec{Oa}$ is $d_a = \sqrt{(x_a)^2 + (y_a)^2 + (z_a)^2}$ and the length of $\vec{Op}$ is $d_p = \sqrt{x^2 + y^2 + z^2}$. Let θ be the angle between $\vec{Oa}$ and $\vec{Op}$. Using the Law of Cosines,

$$D = \sqrt{(d_p)^2 + (d_a)^2 - 2d_p d_a \cos\theta}.$$

Let $t = \frac{d_a}{d_p}$. Substitute $d_a = td_p$ and $x = \cos\theta$ into the above equation to find

$$D = \sqrt{(d_p)^2 + (td_p)^2 - 2d_p(td_p)(x)} = (d_p)\sqrt{1 + t^2 - 2tx}.$$

Therefore,

$$\phi_p(D) = \frac{C}{D} = \frac{C}{(d_p)\sqrt{1+t^2-2tx}} = \frac{C}{(d_p)g(x,t)}$$

where $g(x,t) = \sqrt{1+t^2-2tx}$. To express $g(x,t)$ as a Maclaurin series, which is a Taylor series expansion of a function at $x = 0$, we first calculate the successive derivatives of g. For simplification, let $\alpha = 2tx - t^2$. Then,

$$g(\alpha) = \frac{1}{\sqrt{1-\alpha}} = (1-\alpha)^{-\frac{1}{2}}.$$

The next few successive derivatives of g are

$$g' = \frac{1}{2}(1-\alpha)^{-\frac{3}{2}}, \quad g'' = \frac{3}{4}(1-\alpha)^{-\frac{5}{2}}, \quad g''' = \frac{15}{6}(1-\alpha)^{-\frac{7}{2}}.$$

The pattern implies the nth derivative of g is

$$\frac{(1-\alpha)^{-\frac{2n+1}{2}}(2n-1)!!}{2^n}$$

for non-negative n and $\alpha = 0$. Therefore,

$$g(\alpha) = \sum_{n=0}^{\infty} \frac{(2n-1)!!\alpha^n}{2^n n!}.$$

Proposition 2.41. *The double factorial* $(2k-1)!!$ *is equivalent to* $\frac{(2k)!}{2^k k!}$ *for any positive integer* k*.*

Proof. First we prove two lemmas regarding double factorials.

Lemma 2.42. $(2k)!! = 2^k k!$.

Proof.

$$(2k)!! = (2k)(2k-2)(2k-4)\cdots 4\cdot 2 = 2(k)2(k-1)2(k-2)\cdots 2(2)\cdot 2(1),$$

which is equivalent to

$$= 2^k((k)(k-1)(k-2)\cdots 2\cdot 1) = 2^k k!.$$

□

Lemma 2.43. $n! = n!!(n-1)!!$ *for any non-negative integer* n*.*

Proof. If n is odd, then

$$n!! = n(n-2)(n-4)\cdots 3\cdot 1 \text{ and } (n-1)!! = (n-1)(n-3)(n-5)\cdots 4\cdot 2.$$

Therefore,

$$n!!(n-1)!! = (n\cdot\cdots 5\cdot 3\cdot 1)((n-1)\cdots 6\cdot 4\cdot 2) = n(n-1)\cdots 3\cdot 2\cdot 1 = n!.$$

Otherwise n must be even, so the factors of $n!!$ and $(n-1)!!$ are the even numbers up to n and the odd numbers up to $n-1$, respectively. Altogether, the factors of $n!!(n-1)!!$ must be all the numbers up to n, which is $n!$. □

Lemma 2.43 implies

$$n!! = \frac{n!}{(n-1)!!} = \frac{n(n-1)(n-2)\cdots}{(n-1)(n-3)\cdots} = \frac{(n+1)}{(n+1)}\frac{n(n-1)(n-2)\cdots}{(n-1)(n-3)\cdots} = \frac{(n+1)!}{(n+1)!!}.$$

Now, let $n = 2k-1$. Then,

$$(2k-1)!! = \frac{((2k-1)+1)!}{((2k-1)+1)!!} = \frac{(2k)!}{(2k)!!} = \frac{(2k)!}{2^k k!}$$

using Lemma 2.42 in the denominator. □

Expanding the series explicitly for the first few terms yields

$$\frac{(-1)!!\alpha^0}{2^0 0!} + \frac{(1)!!\alpha^1}{2^1 1!} + \frac{(3)!!\alpha^2}{2^2 2!} + \frac{(5)!!\alpha^3}{2^3 3!} + \cdots = 1 + \frac{1}{2}\alpha + \frac{3}{8}\alpha^2 + \frac{5}{16}\alpha^3 + \cdots$$

using Proposition 2.41. We can then substitute $2xt - t^2$ for α and rearrange terms to find the Legendre polynomials.

$$g = 1 + \frac{1}{2}(2xt - t^2) + \frac{3}{8}(2xt - t^2)^2 + \frac{5}{16}(2xt - t^2)^3 + \cdots$$

$$= 1 + xt - \frac{1}{2}t^2 + \frac{3}{8}((2xt)^2 - 4xt^3 + t^4) + \frac{5}{16}((2xt)^3 + 12xt^4 + 6xt^5 + t^6) + \cdots$$

$$= (1) + (x)t + (\frac{3}{2}x^2 - \frac{1}{2})t^2 + (\frac{5}{2}x^3 - \frac{3}{2}x)t^3 + (\ldots)t^4 + (\ldots)t^5 + \cdots.$$

The expansion of g reveals that the nth Legendre polynomial P_n is the coefficient of t^n, as needed.

3. Interpolation Error

The Runge Phenomenon occurs when uniform spacing between interpolation nodes causes the interpolating polynomial to experience severe oscillation along the edges of the interval of approximation. For example, consider $(1 + 25x^2)^{-1}$ on $[-1, 1]$. The Chebyshev nodes reduce the Runge phenomenon.

Definition 3.1. The Chebyshev nodes x_k defined by

$$x_k = \cos\left(\frac{(2k-1)\pi}{2n}\right)$$

are the roots of $T_n(x)$ on $[-1, 1]$.

Suppose $f \in C^{n+1}[-1, 1]$ and p are polynomials of degree n and p interpolates f at n nodes $\{x_1, x_2, ..., x_n\}$.

Theorem 3.2. *$\|f - p\|_\infty$ is bounded by*

$$\frac{\|f^{(n+1)}\|_\infty}{(n+1)!}\|w\|_\infty$$

where $w(x) = (x - x_1)(x - x_2)\cdots(x - x_n)$.

Since f and n are given, the minimum $\|f - p\|_\infty$ corresponds to the set of interpolation nodes that minimize $\|w\|_\infty$. Let $\tilde{P}$ be the monic polynomials of degree n defined over $[-1, 1]$. Hence $w \in \tilde{P}$. Theorem 2.11 implies that dividing the nth Chebyshev polynomial by its leading coefficient which is 2^{n-1} yields a monic polynomial of degree n whose roots are exactly the Chebyshev nodes.

Theorem 3.3. *$\|f-p\|_\infty$ is bounded by*

$$\frac{1}{2^{(n-1)}(n+1)!}\|f^{(n+1)}\|_\infty. \tag{3.4}$$

Proof. Suppose $g(x) \in \tilde{P}$ and $h(x) = \frac{1}{2^{n-1}}T_n(x)$.

Lemma 3.5. *$h \in \tilde{P}$ satisfies $\|h\|_\infty \le \|p\|_\infty$ for all $p \in \tilde{P}$ over $[-1,1]$. Also, $\|h\|_\infty = \frac{1}{2^{n-1}}$ and $|h| = \|h\|_\infty$ at $n+1$ points $\{x_0, x_1, ..., x_n\}$ such that $x_k = \cos\frac{k\pi}{n}$ for k between 0 and n, inclusive.*

Proof. For a contradiction, suppose $\|g\|_\infty < \frac{1}{2^{n-1}}$. Let $f(x) = \frac{1}{2^{n-1}}T_n(x) - g(x)$. The leading coefficient of $T_n(x)$ is 2^{n-1}, which implies $f(x)$ must have degree less than n since the terms with the highest order in $g(x)$ and $T_n(x)$ disappear. Since $|T_n(x)| = 1$ at its extreme points, $\|g\|_\infty < \frac{1}{2^{n-1}} \le \frac{1}{2^{n-1}}|T_n(x)|$. This implies $|g(x)| < |\frac{1}{2^{n-1}}T_n(x)|$ because n is positive. Altogether, letting $x = \cos(\frac{2k\pi}{n})$ in Definition 2.4 implies

$$T_n(x) = \cos\left(n\cos^{-1}(\cos\frac{2k\pi}{n})\right) = \cos(2k\pi) = 1 \quad \text{for } 0 \le 2k \le n.$$

By assumption, $\|g\|_\infty < \frac{1}{2^{n-1}} \le 1$, so $\|g\|_\infty < 1$. Since $T_n(x) = 1 > |g(x)|$ and k is an integer, we know $f(x)$ must be positive at $\lfloor\frac{n}{2}\rfloor + 1$ distinct points. Similarly, whenever $x = \cos(\frac{(2k+1)\pi}{n})$, we find

$$T_n(x) = \cos\left(n\cos^{-1}\left(\cos\frac{(2k+1)\pi}{n}\right)\right) = \cos(2k\pi+\pi) = -1 \quad \text{for } 0 \le 2k+1 \le n.$$

We know $T_n(x) = -1 < |g(x)|$ so there are $\lfloor\frac{n}{2}\rfloor + 1$ unique points at which $f(x)$ must be negative. To be clear, $|f(x)| = \|f\|_\infty$ at a total of $2\lfloor\frac{n}{2}\rfloor + 2$ extrema which can be described by the set $E = \{x_0, x_1, ..., x_n\} = \{\cos\frac{0\pi}{n}, \cos\frac{1\pi}{n}, ..., \cos\frac{n\pi}{n}\}$. The method of construction of the set E implies $f(x)$ is positive at x_k when k is odd and negative when k is even. The Intermediate Value Theorem asserts that whenever $f(a) < L < f(b)$, there exists a point c in (a,b) satisfying $f(c) = L$. Adjacent elements of the set $f(E) = \{f(x_0), f(x_1), ..., f(x_n)\}$ differ in sign and polynomials are continuous. Without loss of generality, let $f(x_i) < f(x_j)$ for any x_i adjacent to x_j. Then, the Intermediate Value Theorem implies there is some point c in between x_i and x_j such that $f(x_i) < f(c) = 0 < f(x_j)$. Because f is positive $\lfloor\frac{n}{2}\rfloor + 1$ unique times and negative $\lfloor\frac{n}{2}\rfloor + 1$ times, in total f must cross the x-axis n times. In other words, f has n roots, which is impossible because our assumption showed f must have degree less than n. Therefore, it must be true that $\|g\|_\infty \ge \frac{1}{2^{n-1}}$. □

Thus the maximum of any monic polynomial g of degree n over $[-1,1]$ can be made as small as $\frac{1}{2^{n-1}}$, so Theorem 3.2 implies $\|f-p\|_\infty$ can be reduced to

$$\frac{\|f^{(n+1)}\|_\infty}{2^{(n-1)}(n+1)!},$$

completing the proof. □

4. Comparison of Chebyshev and Legendre Expansions

Suppose $f, p^* \in C[-1,1]$ are of degree n. Then the minimax interpolating polynomial p^* exists and is unique. The Legendre expansion $P_n(f)$ of degree n is defined as

$$P_n(f) = a_1P_1(x) + a_2P_2(x) + \cdots + a_nP_n(x)$$

and the Chebyshev expansion $T_n(f)$ of degree n is defined as

$$T_n(f) = a_1T_1(x) + a_2T_2(x) + \cdots + a_nT_n(x),$$

where the coefficients (a_n) for each expansion can be determined using the application of an inner product. It is well known that in the Euclidean norm, $p^* = P_n(f)$ with respect to $\omega(x) = 1$ and $p^* = T_n(f)$ with respect to $\omega(x) = (1-x^2)^{-\frac{1}{2}}$. The rate at which these series converge to f is called the rate of convergence. A larger rate of convergence implies quicker convergence on an approximation that is arbitrarily precise.

Definition 4.1. f is real analytic on (a,b) if it is infinitely differentiable and the Taylor series at $x_0 \in (a,b)$ converges pointwise to $f(x)$ for all $x \in V_\epsilon(x_0)$.

According to [[4], p. 5], the optimal rate of convergence of p^* is better by a constant factor than that of $T_n(f)$ and by a factor of $\sqrt{n}$ than that of $P_n(f)$ for analytic functions. This means the optimal rate of $T_n(f)$ is a factor $\sqrt{n}$ greater than that of $P_n(f)$. For non-analytic smooth functions, the rate of convergence of p^* is better by a constant factor than that of both $T_n(f)$ and $P_n(f)$.

Chebyshev expansions are also better minimax candidates than Legendre expansions of the same degree. According to [[4], p. 3],

$$\|f - T_n(f)\|_\infty \le \left(\frac{4}{\pi^2}\log n + 4\right)\|f - p^*\|_\infty \tag{4.2}$$

and

$$\|f - P_n(f)\|_\infty \le \left(\frac{2^{\frac{2}{3}}}{\sqrt{\pi}}\sqrt{n} + k\right)\|f - p^*\|_\infty \tag{4.3}$$

for some constant k. This implies the maximum error of Chebyshev expansions is worse by at most a logarithmic factor compared to that of p^*, while the maximum error of Legendre expansions is worse by at most a factor of $\sqrt{n}$ compared to that of p^*. Since

$$\lim_{x\to\infty} \frac{\log(x)}{f(x)} = 0$$

for any polynomial f, $\|T_n(f)\|_\infty < \|P_n(f)\|_\infty$ as n approaches ∞. Examples of analytic functions for which the superiority of Chebyshev expansions was verified include $\exp(x^5)$, $\ln(1.2+x)$, and $(1+4x^2)^{-1}$ in [4] and $\sin^{-1} x$ and e^x in [10] for the interval $[-1,1]$.

The optimal rates of convergence of $P_n(f)$ and $T_n(f)$ are a constant times less than that of p^* for non-analytic functions. Non-analytic freeform surfaces might contain discontinuities or discontinuities of the derivative where the deviations of $P_n(f)$ and $T_n(f)$ from f increase noticeably. Examples of various non-analytic functions which were tested in [4] are $(x-\frac{1}{2})_+^3$ and $|\sin(5x)|$ (piecewise analytic), $(x-\frac{1}{2})^{\frac{5}{2}}$, $|(x-\frac{4}{5})^{\frac{5}{4}}|$, and $|x|^{\frac{2}{3}}$ (contains interior singularity), and $(x+1)^{\frac{5}{2}}$, $(1-x^2)^{\frac{3}{2}}$, and $\cos^{-1}(x)$ (contains endpoint singularities).

It is clear however that Chebyshev and Legendre polynomial expansions are both more suitable than square Zernike polynomials for describing analytic freeform surfaces over rectangular apertures. Diagrams on [[7], p. 9] obtained during aberration correction show that the optical quality achieved over rectangular apertures is better with the Legendre polynomials than with square Zernike polynomials and the former yields better image quality than existing systems after identical optimization procedures were applied.

Remark 4.4. Around 30 Legendre or Chebshev terms are needed for interpolation of freeform surfaces over rectangular apertures.

Although Chebyshev expansions are a better minimax candidate than Legendre expansions, the former may perform worse than the latter at specific points. For example, [9] showed that the Henyey-Greenstein phase function, which describes the angular distribution of light scattered by small particles, is better approximated by the Legendre expansion near a 0° forward scattering angle and by the Chebyshev expansion at most other scattering angles.

5. Polynomials for Elliptic Apertures

We turn our attention to the final result of this report, an attempt to derive orthogonal polynomials for elliptical apertures. We follow the method of [6].

Elliptical coordinates are μ and ν.

Definition 5.1. For $\mu \in \mathbb{R}^+$ and $\nu \in [0, 2\pi]$,

$$x = a\cos h\mu\cos\nu \quad \text{and} \quad y = a\sin h\mu\sin\nu.$$

Suppose f is a continuous function defined over the unit ellipse. Weiertrass's approximation theorem implies we can approximate f as precisely as needed with a function of polynomials. Hence,

$$f(x, y) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty} C_{pq}x^p y^q, \tag{5.2}$$

where C_{pq} are arbitrary coefficients. Definition 5.1 implies

$$x^p y^q = (a\cosh\mu\cos\nu)^p(a\sinh\mu\sin\nu)^q.$$

Since hyperbolic functions cosh and sinh equal

$$\cosh x = \frac{e^x + e^{-x}}{2} \quad \text{and} \quad \sinh x = \frac{e^x - e^{-x}}{2},$$

we have

$$\begin{aligned} x^p y^q &= \left[(a\cos\nu)\left(\frac{e^\mu + e^{-\mu}}{2}\right)\right]^p \left[(a\sin\nu)\left(\frac{e^\mu - e^{-\mu}}{2}\right)\right]^q \\ &= a^{p+q}\cos^p\nu\sin^q\nu\left(\frac{e^\mu + e^{-\mu}}{2}\right)^p\left(\frac{e^\mu - e^{-\mu}}{2}\right)^q, \end{aligned}$$

which Euler's formula for $\cos^p\nu\sin^q\nu$ implies is equivalent to

$$a^{p+q}\left(\frac{e^{i\nu} + e^{-i\nu}}{2}\right)^p\left(\frac{e^{i\nu} - e^{-i\nu}}{2i}\right)^q\left(\frac{e^\mu + e^{-\mu}}{2}\right)^p\left(\frac{e^\mu - e^{-\mu}}{2}\right)^q.$$

Factoring out the powers of 2 gives

$$\frac{a^{p+q}}{2^{(2p+2q)}i^q}\left(e^{i\nu} + e^{-i\nu}\right)^p\left(e^{i\nu} - e^{-i\nu}\right)^q\left(e^\mu + e^{-\mu}\right)^p\left(e^\mu - e^{-\mu}\right)^q.$$

For abbreviation, let

$$\zeta = \frac{a^{p+q}}{2^{(2p+2q)} i^q}.$$

Then, (2.8) implies

$$x^p y^q = \zeta \left[\sum_{s=0}^{p} \binom{p}{s} (e^{i\nu})^{p-s} (e^{-i\nu})^s \right] \left[\sum_{t=0}^{q} \binom{q}{t} (e^{i\nu})^{q-t} (-e^{-i\nu})^t \right]$$
$$\left[\sum_{s=0}^{p} \binom{p}{s} (e^{\mu})^{p-s} (e^{-\mu})^s \right] \left[\sum_{t=0}^{q} \binom{q}{t} (e^{\mu})^{q-t} (-e^{-\mu})^t \right],$$

which is equivalent to

$$4\zeta \sum_{s=0}^{p} \sum_{t=0}^{q} \binom{p}{s} (e^{i\nu})^{p-s} (e^{-i\nu})^s (e^{\mu})^{p-s} (e^{-\mu})^s$$
$$\binom{q}{t} (e^{i\nu})^{q-t} (-e^{-i\nu})^t (e^{\mu})^{q-t} (-e^{-\mu})^t.$$

after combining identical power series and combinations. For better readability, let

$$\sigma = 4\zeta \sum_{s=0}^{p} \sum_{t=0}^{q} \binom{p}{s} \binom{q}{t},$$

which implies $x^p y^q$ is equivalent to

$$\sigma (e^{i\nu})^{p-s} (e^{i\nu})^{q-t} (e^{-i\nu})^s (-e^{-i\nu})^t (e^{\mu})^{p-s} (e^{\mu})^{q-t} (e^{-\mu})^s (-e^{-\mu})^t.$$

After combining terms with identical bases, we reach

$$\sigma (e^{i\nu})^{p+q-(s+t)} (e^{-i\nu})^{s+t} (e^{\mu})^{p+q-(s+t)} (e^{-\mu})^{s+t},$$

which by rearrangement and properties of exponents is equivalent to

$$\sigma (e^{i\nu} e^{\mu})^{p+q-(s+t)} (e^{-i\nu} e^{-\mu})^{(s+t)} = \sigma e^{(i\nu+\mu)(p+q)} e^{(i\nu+\mu)(-2(s+t))}.$$

Substituting for σ yields

$$x^p y^q = 4\zeta \sum_{s=0}^{p} \sum_{t=0}^{q} \binom{p}{s} \binom{q}{t} e^{(i\nu+\mu)(p+q)} e^{(i\nu+\mu)(-2(s+t))}. \tag{5.3}$$

However, if s and t are nonzero, then

$$\binom{p}{s} \binom{q}{t} = \left(\frac{p!}{(p-s)!s!} \right) \left(\frac{q!}{(q-t)!t!} \right),$$

which is equivalent to

$$\frac{(p(p-1)\cdots(p-(s-1)))\,(q(q-1)\cdots(q-(t-1)))}{s!t!}. \tag{5.4}$$

Otherwise,

$$\binom{p}{s} \binom{q}{t} = \begin{cases} p & \text{if } s = 1, t = 0 \\ q & \text{if } s = 0, t = 1 \\ 1 & \text{if } s = t = 0. \end{cases} \tag{5.5}$$

According to [7], letting $l = s + t$ enables us to combine exponential terms with fixed l into a single term with coefficient C_l, but we remark that (5.4) and (5.5) imply that introducing l makes it impossible to determine the exact value of C_l. Substituting for ζ in (5.3) implies

$$x^p y^q = 4\left(\frac{a^{p+q}}{2^{(2p+2q)} i^q}\right) e^{(i\nu+\mu)(p+q)} \sum_{l=0}^{p+q} C_l e^{(i\nu+\mu)(-2l)},$$

and letting $m = p + q$ implies

$$x^p y^q = \frac{a^m e^{(i\nu+\mu)m}}{2^{(2m-2)} i^q} \sum_{l=0}^{m} C_l e^{(i\nu+\mu)(-2l)}.$$

Finally we use (5.2) to arrive at

$$f(a,\mu,\nu) = \sum_{m=0}^{\infty} a^m C_m \frac{e^{(i\nu+\mu)m}}{2^{(2m-2)} i^q} \sum_{l=0}^{m} C_l e^{(i\nu+\mu)(-2l)} = \sum_{m=0}^{\infty} a^m \sum_{l=0}^{m} C_{lm} e^{(i\nu+\mu)(-2l)}.$$

Hence,

$$f(a,\mu,\nu) = \sum_{m=0}^{\infty} a^m \sum_{l=0}^{m} C_{lm} e^{(i\nu+\mu)(-2l)}. \tag{5.6}$$

Acknowledgments

I would like to thank my advisors — Professor Jannick Rolland, Renuka Manjula, and Yiwen Fan — for introducing me to the world of freeform optics, inviting me to events and laboratories, and recommending papers and code to guide me through the subject. I would also like to thank Connie Jones for organizing this REU. Finally, I would like to thank my family for giving me positive support to make this all possible.

References

[1] P. Chebyshev, "Théorie des mécanismes connus sous le nom de parallélogrammes", 1854.

[2] V. Kaplunovsky, https://web2.ph.utexas.edu/ vadim/Classes/2016s/mme.pdf, 2016.

[3] B. Ibrahimoglu, "Lebesgue functions and Lebesgue constants in polynomial interpolation", 2016.

[4] H. Wang, "How much faster does the best polynomial approximation converge than Legendre projection?", 2021.

[5] M. Kline, *Mathematical Thought From Ancient to Modern Times*, Volume 2, 1990.

[6] A. Yen, "Straightforward path to Zernike polynomials", 2021.

[7] E. Muslimov et al., "Combining freeform optics and curved detectors for wide field imaging: a polynomial approach over squared aperture", 2017.

[8] J. Boyd et al., "The Relationships Between Chebyshev, Legendre and Jacobi Polynomials: The Generic Superiority of Chebyshev Polynomials and Three Important Exceptions", 2014.

[9] F. Zhang, "Comparison of Chebyshev and Legendre Polynomial Expansion of Phase Function of Cloud and Aerosol Particles", 2017.

[10] N. Patanarapeelert et al., "Comparison Study of Series Approximation and Convergence between Chebyshev and Legendre Series", 2013.