"There is only one bullet in the gun. It's called a Chebyshev polynomial."
— Rocco Servedio via Moritz Hardt

"Chebyshev polynomials are everywhere dense in numerical analysis."
— Philip Davis & George Forsythe

The **Chebyshev polynomials** appear frequently in numerical analysis and are incredibly useful for analyzing and accelerating the convergence of iterative methods. One might even say that Chebyshev polynomials are the *best* polynomials, a fact which can be made precise in a variety of different ways. In these notes, we define Chebyshev polynomials and their basic properties, before discussing their utility in **minimax approximation theory**, which was the subject of a previous set of notes.
Note: The material below draws heavily from the first three chapters of (Mason and Handscomb 2003) and in some places has been copied directly.

Trigonometric Formulation

Definition 1. The Chebyshev polynomial $T_n(x)$ of the **first kind** is a polynomial in x of degree n, defined for $x = \cos\theta \in [-1, 1]$ by the relation

$$T_n(x) = \cos n\theta = \cos n \cos^{-1} x \tag{1}$$

Applying **de Moivre's formula**, $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$, makes it clear that $T_n(x)$ is indeed a polynomial of degree n in $x = \cos\theta$. This form allows us to evaluate $T_n(x)$ outside of $[-1, 1]$ but obscures the connection to trigonometric functions. The first several Chebyshev polynomials are shown below.

$$\begin{aligned}
T_0(x) &= \cos 0 && = 1 \\
T_1(x) &= \cos\theta && = x \\
T_2(x) &= \cos 2\theta = 2\cos^2\theta - 1 && = x^2 - 1 \\
T_3(x) &= \cos 3\theta = 4\cos^3\theta - 3\cos\theta && = 4x^3 - 3x \\
T_4(x) &= \cos 4\theta = 8\cos^4\theta - 8\cos^2\theta + 1 && = 8x^4 - 8x^2 + 1
\end{aligned} \tag{2}$$

Compare the graphs of $T_5(x)$ and $\cos 5\theta$ in Figure 4. The shape of both functions is similar, and in particular both oscillate between six extrema of equal magnitude with alternating signs. However, there are three key differences, as observed by [1]:

- The polynomial $T_5(x)$ corresponds to $\cos 5x$ *reversed*. If the range of the variable x is the interval $[-1, 1]$, then the range of the corresponding variable θ can be taken as $[0, \pi]$. These ranges are traversed in opposite directions, since $x = -1$ corresponds to $\theta = \pi$ and $x = 1$ corresponds to $\theta = 0$.
- The extrema of $T_5(x)$ at the endpoints $x = \pm 1$ do not correspond to zero gradients as they do for $\cos 5\theta$, but rather to rapid changes in the polynomial as a function of x. In fact, the Chebychev polynomials increase more quickly outside the interval $[-1, 1]$ than any other polynomial of similar magnitude within the interval, as noted in [2].
- The zeros and extrema of $T_5(x)$ are clustered towards the endpoints ± 1, whereas the zeros and extrema of $\cos 5\theta$ are equally spaced, due to the fact that $x = \cos\theta$ is a nonlinear transformation of the domain that *contracts* with increasing magnitude towards the endpoints.

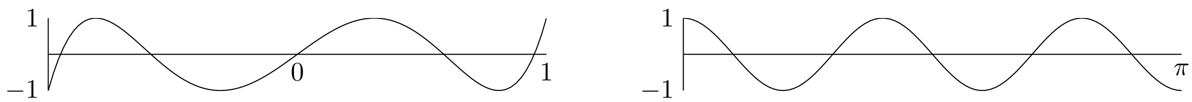

Figure 1: Comparison of the Chebychev polynomial $T_5(x)$ on the left with $\cos 5\theta$ on the right.

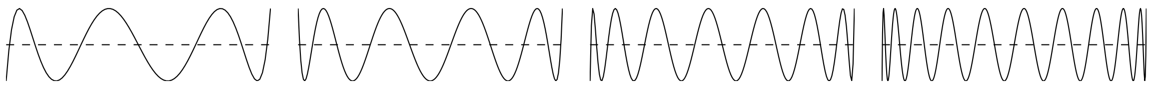

Figure 2: Comparison of $T_n(x)$ for $n = 7, 10, 15, 20$.

Recurrence Relation

Proposition 1. *For any integer* $n \in \mathbb{Z}$, $\cos n\theta + \cos(n-2)\theta = 2\cos\theta\cos(n-1)\theta$.

Proof. Recall the identity $\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$, easily deduced by Euler's formula. Then,

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

Now for the usual trick. We recognize the lefthand side of our identity as the real part of a complex exponential and apply de Moivre's formula. Having done this, we make the above substitution.

$$\begin{aligned}
\cos n\theta + \cos(n-2)\theta &= \mathfrak{Re}\left[e^{ni\theta} + e^{(n-2)i\theta}\right] = \mathfrak{Re}\left[e^{(n-2)i\theta}(1 + e^{2i\theta})\right] \\
&= \mathfrak{Re}\left[(\cos\theta + i\sin\theta)^{n-2}(1 + (\cos\theta + i\sin\theta)^2)\right] \\
&= \mathfrak{Re}\left[(\cos(n-2)\theta + i\sin(n-2)\theta)\cdot(2\cos^2\theta + 2i\cos\theta\sin\theta)\right] \\
&= 2\cos\theta\left[\cos\theta\cos(n-2)\theta - \sin\theta\sin(n-2)\theta\right] \\
&= 2\cos\theta\cos(n-1)\theta
\end{aligned}$$

□

Corollary 1. *The Chebyshev polynomials satisfy the following recurrence relation:*

$$\begin{cases} T_0(x) = 1 \\ T_1(x) = x \\ T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x) \quad \text{for } n \geq 2 \end{cases} \tag{3}$$

Zeros of $T_n(x)$

The Chebychev polynomials have precisely n zeros and $n+1$ local extrema in the interval $[-1,1]$, and nowhere else. From our trigonometric definition, the zeros for $T_n(x)$ in $[-1,1]$ correspond to the zeros in $[0,\pi]$ of $\cos n\theta$, so that

$$n\theta = \left(k - \frac{1}{2}\right)\pi \qquad \text{for } k = 1, 2, \ldots, n \tag{4}$$

Hence, the zeros of $T_n(x)$ are

$$x = x_k \equiv \cos\frac{(k-\frac{1}{2})\pi}{n} \qquad \text{for } k = 1, 2, \ldots, n \tag{5}$$

As defined, $x_1 > \cdots > x_n$ in decreasing order. Note $x = 0$ is a zero of $T_n(x)$ for all odd n, but not for even n, and that zeros are symmetrically placed in pairs on either side of $x = 0$, with increasing density towards the endpoints. These points, called **Chebyshev nodes**, are often used as nodes in polynomial interpolation because they can be shown to minimize oscillatory approximation errors due to **Runge's phenomenon**.

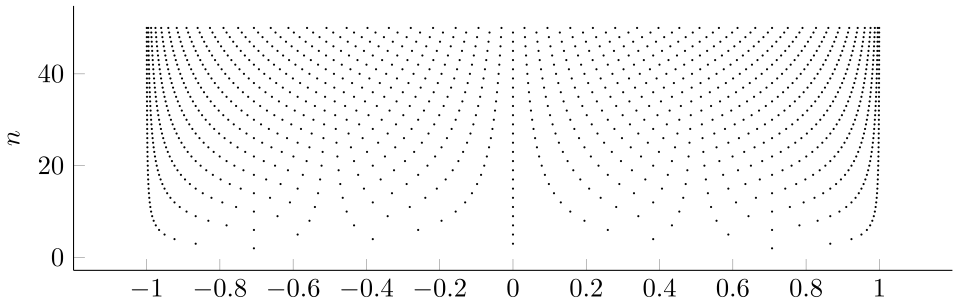

Figure 3: Zeros $\cos \frac{(k-\frac{1}{2})\pi}{n}$ of the Chebyshev polynomial $T_n(x)$ plotted for several values of n.

Extrema of $T_n(x)$

The internal extrema of $T_n(x)$ correspond to the extrema of $\cos n\theta$, namely the zeros of $\sin n\theta$, since

$$\frac{\mathrm{d}}{\mathrm{d}x} T_n(x) = \frac{\mathrm{d}}{\mathrm{d}x} \cos n\theta = \frac{\mathrm{d}}{\mathrm{d}\theta} \cos n\theta \Big/ \frac{\mathrm{d}x}{\mathrm{d}\theta} = \frac{-n \sin n\theta}{-\sin\theta} \tag{6}$$

Hence, including those at $x = \pm 1$, the extrema of $T_n(x)$ on $[-1, 1]$ occur when $\theta = \frac{k\pi}{n}$, that is,

$$x = \cos \frac{k\pi}{n} \qquad \text{for } k = 0, 1, \ldots, n \tag{7}$$

Minimax Property

Note that the $n + 1$ extrema of $T_n(x)$ are all of equal magnitude and alternate in sign between -1 and $+1$. We will use this fact in conjunction with the **Chebyshev equioscillation theorem**, reproduced below, to show that $\frac{1}{2^{n-1}} T_n(x)$ is the monic polynomial of degree n with smallest L^∞-norm on the interval $[-1, 1]$.

Theorem 1 (Chebyshev Equioscillation)**.** *A polynomial $p^* \in \mathcal{P}_n$ is a best (minimax) approximation to $f \in \mathcal{C}[a, b]$ (in the L^∞ sense) if and only if there is an alternating set for $f - p$ consisting of at least $n + 2$ points. The best approximation is unique.*

Proof Sketch. High-degree polynomials have more opportunity to reverse direction than do lower-degree polynomials. Alternatively, because (nonzero) polynomials are necessarily unbounded, we have more *control* over the values of higher-degree polynomials before they blow up. Due to these observations, it turns out that approximations which violate the alternating property leave extra *wiggle room* to correct for error, allowing a better approximation to be constructed. For full details, see my previous notes or [3, 4, 5]. □

The equioscillation theorem gives us a means to *identify* best approximations but not a way to *find* them. The **Remez algorithm** is a simple numerical method for computing minimax approximations that exploits the equioscillation theorem to assess convergence. The theorem can also be useful for direct proofs, as below.

Example 1. Consider the minimax polynomial approximation $p^*_{n-1} \in \mathcal{P}_{n-1}$ to $f(x) = x^n$ over $[-1, 1]$.

(1) Notice that the Chebyshev polynomial $T_n(x)$ has degree n with leading coefficient $\frac{1}{2^{n-1}}$. If we choose $p^*_{n-1}(x) = x^n - \frac{1}{2^{n-1}} T_n(x)$, then the error function is the monic polynomial

$$\varphi(x) = x^n - p^*_{n-1}(x) = \frac{1}{2^{n-1}} T_n(x)$$

(2) We know $\varphi(x)$ has an alternating set of $n + 1$ points, corresponding to the extrema of the polynomial $T_n(x)$. By the equioscillation theorem, p^*_{n-1} is the unique best approximation to x^n on $[-1, 1]$. Because $|T_n(x)| \le 1$ on the interval, the approximation error is no larger than $\frac{1}{2^{n-1}}$. More explicitly,

$$\max_{x \in [-1,1]} \left| \frac{1}{2^{n-1}} T_n(x) \right| \le \max_{x \in [-1,1]} |x^n - q(x)| \qquad \forall\, q \in \mathcal{P}_{n-1} \tag{8}$$

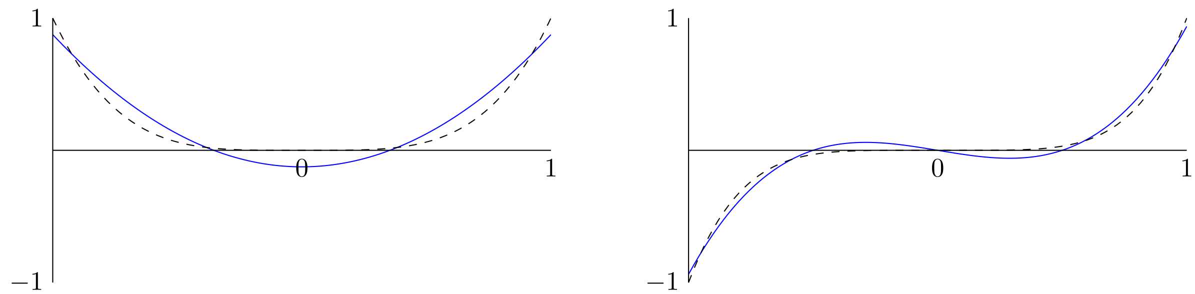

Figure 4: The function $x^n - \frac{1}{2^{n-1}} T_n(x)$ is a fantastic degree-n approximation to the x^n.

Since every monic polynomial of degree n takes the form $x^n - q(x)$ for some $q \in \mathcal{P}_{n-1}$, we have just proven the following useful theorem. In a future set of notes, we will see that this result fully accounts for the appearance of Chebyshev polynomials in the convergence analysis of iterative methods.

Theorem 2 (Minimax Property of $T_n(x)$)**.** *The normalized Chebyshev polynomial $\frac{1}{2^{n-1}}$ is the minimax approximation on $[-1,1]$ to the zero function by a monic polynomial of degree n; that is, it is the monic polynomial of degree n with the smallest L^∞-norm on $[-1,1]$.*

Growth Rate Beyond $[-1, 1]$

The following *extremal* property of Chebyshev polynomials states that Chebyshev polynomials increase in magnitude more quickly outside the range $[-1,1]$ than any other polynomial with magnitude bounded by one in the same interval. Loosely speaking, Chebyshev polynomials increase as quickly as possible outside the interval. For a direct proof that avoids the equioscillation theorem, see (Carothers 1998); for several applications, see (Musco 2015).

Theorem 3. *Let $q \in \mathcal{P}_n$ with $|q(x)| \leq 1$ for all $x \in [-1,1]$. Then, for any $y \notin [-1,1]$, $|q(x)| \leq |T_n(x)|$.*

Proof. Assume towards contradiction that there exist $q \in \mathcal{P}_n$ and $y \notin [-1,1]$ with $|q(y)| > |T_n(y)|$. Consider

$$h(x) \equiv \frac{T_n(y)}{q(y)} q(x) \quad \text{and} \quad \varphi(x) \equiv T_n(x) - h(x) \quad \text{in } \mathcal{P}_n \tag{9}$$

Clearly $\varphi(y) = 0$. Further, because $|h(x)| \leq 1$ for $x \in [-1,1]$ and $T_n(x)$ alternates between ± 1 at least n times, the continuous functions T_n and h must intersect no fewer than n times. Hence $\varphi \in \mathcal{P}_n$ has at least $(n+1)$ zeros, a contradiction! □

Chebyshev Polynomials on $[a, b]$

It is usually necessary work with Chebyschev polynomials on a general interval $[a,b]$, which we simply map to $[-1,1]$ under the linear change of variable

$$s = \frac{2x - (a+b)}{b-a} \tag{10}$$

The Chebyshev polynomials of the first kind appropriate to $x \in [a,b]$ are thus $T_n(s)$.

Trigonometric-Free Form

Chebyshev polynomials can be expressed in a closed form without the use of trigonometric functions. We give two proofs below; the first relies on the recurrence relation established earlier, while the second begins from the trigonometric definition and reveals a connection to the hyperbolic cosine function. Notice first that the expression below is indeed a polynomial, since the odd powers of $\sqrt{x^2-1}$ cancel in the binomial expansions of each term.

Theorem 4. *For all $x \in \mathbb{R}$ and $n \in \mathbb{N}$,*

$$T_n(x) = \frac{1}{2}\left[(x + \sqrt{x^2-1})^n + (x - \sqrt{x^2-1})^n\right] \tag{11}$$

First Proof. (from [7]) We proceed by induction using the recurrence from in Corollary 1.

(1) Let $z = x+\sqrt{x^2-1}$. Note $z^{-1} = x-\sqrt{x^2-1}$. We wish to prove that for all $n \in \mathbb{N}$, $T_n(x) = \frac{1}{2}(z^n+z^{-n})$.

(2) Base Case. The result is clearly true for $n = 0, 1$, since

$$T_0(x) = 1 = \frac{1}{2}(z^0+z^0) \qquad T_1(x) = x = \frac{1}{2}(z^1+z^{-1})$$

(3) Induction. Assume $T_k(x) = \frac{1}{2}(z^k+z^{-k})$ for all $k = 0, \dots, n$. Then,

$$\begin{aligned} T_{n+1}(x) &= 2xT_n(x) - T_{n-1}(x) \\ &= 2x\left[\frac{1}{2}(z^n+z^{-n})\right] - \left[\frac{1}{2}(z^{n-1}+z^{-n-1})\right] \\ &= \frac{1}{2}\left[z^{n-1}(2xz-1) + z^{-(n-1)}(2xz^{-1}-1)\right] = \boxed{\frac{1}{2}(z^{n+1}+z^{-(n+1)})} \end{aligned}$$

since $2xz-1 = z^2+1$ and $2xz^{-1} = z^{-2}+1$ as a result of the base case. By induction, we are done! □

Second Proof. (from [4]) Let $x = \cos\theta$ and expand $T_n(x)$ as a complex exponential,

$$\begin{aligned} T_n(x) = T_n(\cos\theta) = \cos n\theta &= \frac{e^{in\theta}+e^{-in\theta}}{2} \\ &= \frac{1}{2}\left[(\cos\theta + i\sin\theta)^n + (\cos\theta - i\sin\theta)^n\right] \\ &= \frac{1}{2}\left[(x + i\sqrt{1-x^2})^n + (x - i\sqrt{1-x^2})^n\right] \\ &= \frac{1}{2}\left[(x + \sqrt{x^2-1})^n + (x - \sqrt{x^2-1})^n\right] \end{aligned}$$

We've shown that these two polynomials agree for $|x| \le 1$, so they must agree for all x (real or complex). □

Corollary 2. *For all $n \in \mathbb{N}$ and $|x| \ge 1$, $T_n(x) = \cosh n \cosh^{-1}(x)$.*

References

[1] John C. Mason and David C. Handscomb. *Chebyshev Polynomials*. CRC Press, 2003.

[2] Jonathan Richard Shewchuk. An Introduction to the Conjugate Gradient Method Without the Agonizing Pain, 1994.

[3] Douglas N. Arnold. A Concise Introduction to Numerical Analysis. *Institute for Mathematics and its Applications, Minneapolis*, 2001.

[4] Neal L. Carothers. A Short Course on Approximation Theory. *Bowling Green State University*, 1998.

[5] Alexei Shadrin. Approximation Theory, Lecture 5. University of Cambridge, Mathematical Tripos Part III, 2005. URL http://www.damtp.cam.ac.uk/user/na/PartIIIat/b05.pdf.

[6] Cameron Musco. Chebyshev Polynomials and Approximation Theory in Theoretical Computer Science and Algorithm Design. Talk for MIT's Danny Lewin Theory Student Retreat, October 2015. URL http://www.cameronmusco.com/personal_site/pdfs/retreatTalk.pdf.

[7] Sushant Sachdeva and Nisheeth K Vishnoi. Faster Algorithms via Approximation Theory. *Theoretical Computer Science*, 9(2):125–210, 2013.