LECTURE 15: *Chebyshev Polynomials for Optimal Interpolation*

2.3 *Optimal Interpolation Points via Chebyshev Polynomials*

As an application of the minimax approximation procedure, we consider how best to choose interpolation points $\{x_j\}_{j=0}^n$ to minimize

$$\|f - p_n\|_\infty,$$

where $p_n \in \mathcal{P}_n$ is the interpolant to f at the specified points.

Recall the interpolation error bound developed in Section 1.6: If $f \in C^{n+1}[a,b]$, then for any $x \in [a,b]$ there exists some $\xi \in [a,b]$ such that

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{j=0}^{n}(x - x_j).$$

Taking absolute values and maximizing over $[a,b]$ yields the bound

$$\|f - p_n\|_\infty = \max_{\xi \in [a,b]} \frac{|f^{(n+1)}(\xi)|}{(n+1)!} \max_{x \in [a,b]} \Big| \prod_{j=0}^{n}(x - x_j) \Big|.$$

For Runge's example, $f(x) = 1/(1+x^2)$ for $x \in [-5,5]$, we observed that $\|f - p_n\|_\infty \to \infty$ as $n \to \infty$ if the interpolation points $\{x_j\}$ are uniformly spaced over $[-5,5]$. However, Marcinkiewicz's theorem (Section 1.6) guarantees there is always some scheme for assigning the interpolation points such that $\|f - p_n\|_\infty \to 0$ as $n \to \infty$. While there is no fail-safe *a priori* system for picking interpolations points that will yield uniform convergence for *all* $f \in C[a,b]$, there is a distinguished choice that works exceptionally well for just about every function you will encounter in practice. We determine this set of interpolation points by choosing those $\{x_j\}_{j=0}^n$ that *minimize the error bound* (which is distinct from – but hopefully akin to – minimizing the error itself, $\|f - p_n\|_\infty$). That is, we want to solve

$$\min_{x_0,\ldots,x_n} \max_{x \in [a,b]} \Big| \prod_{j=0}^{n}(x - x_j) \Big|. \tag{2.9}$$

Notice that

$$\begin{aligned}\prod_{j=0}^{n}(x - x_j) &= x^{n+1} - x^n \sum_{j=0}^{n} x_j + x^{n-1} \sum_{j=0}^{n}\sum_{k=0}^{n} x_j x_k - \cdots + (-1)^{n+1} \prod_{j=0}^{n} x_j \\ &= x^{n+1} - r(x),\end{aligned}$$

where $r \in \mathcal{P}_n$ is a degree-n polynomial depending on the interpolation nodes $\{x_j\}_{j=0}^n$.

For example, when $n = 1$,

$$(x - x_0)(x - x_1) = x^2 - \Big((x_0 + x_1)x - x_0 x_1\Big) = x^2 - r_1(x),$$

where $r_1(x) = (x_0 + x_1)x - x_0 x_1$. By varying x_0 and x_1, we can obtain make r_1 any function in $\mathcal{P}_1$.

To find the optimal interpolation points according to (2.9), we should solve

$$\min_{r \in \mathcal{P}_n} \max_{x \in [a,b]} |x^{n+1} - r(x)| = \min_{r \in \mathcal{P}_n} \|x^{n+1} - r(x)\|_\infty.$$

Here the goal is to approximate an $(n+1)$-degree polynomial, x^{n+1}, with an n-degree polynomial. The method of solution is somewhat indirect: we will produce a class of polynomials of the form $x^{n+1} - r(x)$ that satisfy the requirements of the Oscillation Theorem, and thus $r(x)$ must be the minimax polynomial approximation to x^{n+1}. As we shall see, the roots of the resulting polynomial $x^{n+1} - r(x)$ will fall in the interval $[a, b]$, and can thus be regarded as 'optimal' interpolation points. For simplicity, we shall focus on the interval $[a, b] = [-1, 1]$.

Definition 2.1. The degree-n *Chebyshev polynomial* is defined for $x \in [-1, 1]$ by the formula

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

At first glance, this formula may not appear to define a polynomial at all, since it involves trigonometric functions. But computing the first few examples, we find

$$n = 0: \quad T_0(x) = \cos(0 \cos^{-1} x) = \cos(0) = 1$$
$$n = 1: \quad T_1(x) = \cos(\cos^{-1} x) = x$$
$$n = 2: \quad T_2(x) = \cos(2 \cos^{-1} x) = 2\cos^2(\cos^{-1} x) - 1 = 2x^2 - 1.$$

Furthermore, it doesn't apply if $|x| > 1$. For such x one can define the Chebyshev polynomials using hyperbolic trigonometric functions, $T_n(x) = \cosh(n \cosh^{-1} x)$. Indeed, using hyperbolic trigonometric identities, one can show that this expression generates for $x \notin [-1, 1]$ the same polynomials we get for $x \in [-1, 1]$ from the standard trigonometric identities. We discuss this point in more detail at the end of the section.

For $n = 2$, we employed the identity $\cos 2\theta = 2\cos^2\theta - 1$, substituting $\theta = \cos^{-1} x$. More generally, use the cosine addition formula

$$\cos\alpha + \cos\beta = 2\cos\left(\frac{\alpha+\beta}{2}\right)\cos\left(\frac{\alpha-\beta}{2}\right)$$

to get the identity

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

This formula implies, for $n \geq 2$,

$$T_{n+1}(x) = 2x T_n(x) - T_{n-1}(x),$$

a formula related to the three term recurrence used to construct orthogonal polynomials.

In fact, Chebyshev polynomials are orthogonal polynomials on $[-1, 1]$ with respect to the inner product

$$\langle f, g \rangle = \int_a^b \frac{f(x)g(x)}{\sqrt{1-x^2}}\,\mathrm{d}x,$$

a fact we will use when studying Gaussian quadrature later in the semester.

Chebyshev polynomials exhibit a wealth of interesting properties, of which we mention just three.

Theorem 2.4. Let T_n be the degree-n Chebyshev polynomial

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

for $x \in [-1, 1]$.

- $|T_n(x)| \leq 1$ for $x \in [-1, 1]$.
- The roots of T_n are the n points $\xi_j = \cos \frac{(2j-1)\pi}{2n}$, $j = 1, \ldots, n$.
- For $n \geq 1$, $|T_n(x)|$ is maximized on $[-1, 1]$ at the $n+1$ points $\eta_j = \cos(j\pi/n)$, $j = 0, \ldots, n$:

$$T_n(\eta_j) = (-1)^j.$$

Proof. These results follow from direct calculations. For $x \in [-1, 1]$, $T_n(x) = \cos(n \cos^{-1}(x))$ cannot exceed one in magnitude because cosine cannot exceed one in magnitude. To verify the formula for the roots, compute

$$T_n(\xi_j) = \cos\left(n \cos^{-1} \cos\left(\frac{(2j-1)\pi}{2n}\right)\right) = \cos\left(\frac{(2j-1)\pi}{2}\right) = 0,$$

since cosine is zero at half-integer multiples of π. Similarly,

$$T_n(\eta_j) = \cos\left(n \cos^{-1} \cos\left(\frac{j\pi}{n}\right)\right) = \cos(j\pi) = (-1)^j.$$

Since $T_n(\eta_j)$ is a nonzero degree-n polynomial, it cannot attain more than $n+1$ extrema on $[-1, 1]$, including the endpoint: we have thus characterized all the maxima of $|T_n|$ on $[-1, 1]$. ■

Figure 2.5 shows Chebyshev polynomials T_n for nine different values of n.

2.3.1 *Interpolation at Chebyshev Points*

Finally, we are ready to solve the key minimax problem that will reveal optimal interpolation points. Looking at the above plots of Chebyshev polynomials, with their striking equioscillation properties, perhaps you have already guessed the solution yourself.

We defined the Chebyshev polynomials so that

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

with $T_0(x) = 1$ and $T_1(x) = x$. Thus T_{n+1} has the leading coefficient 2^n for $n \geq 0$. Define

$$\widehat{T}_{n+1} = 2^{-n} T_{n+1}$$

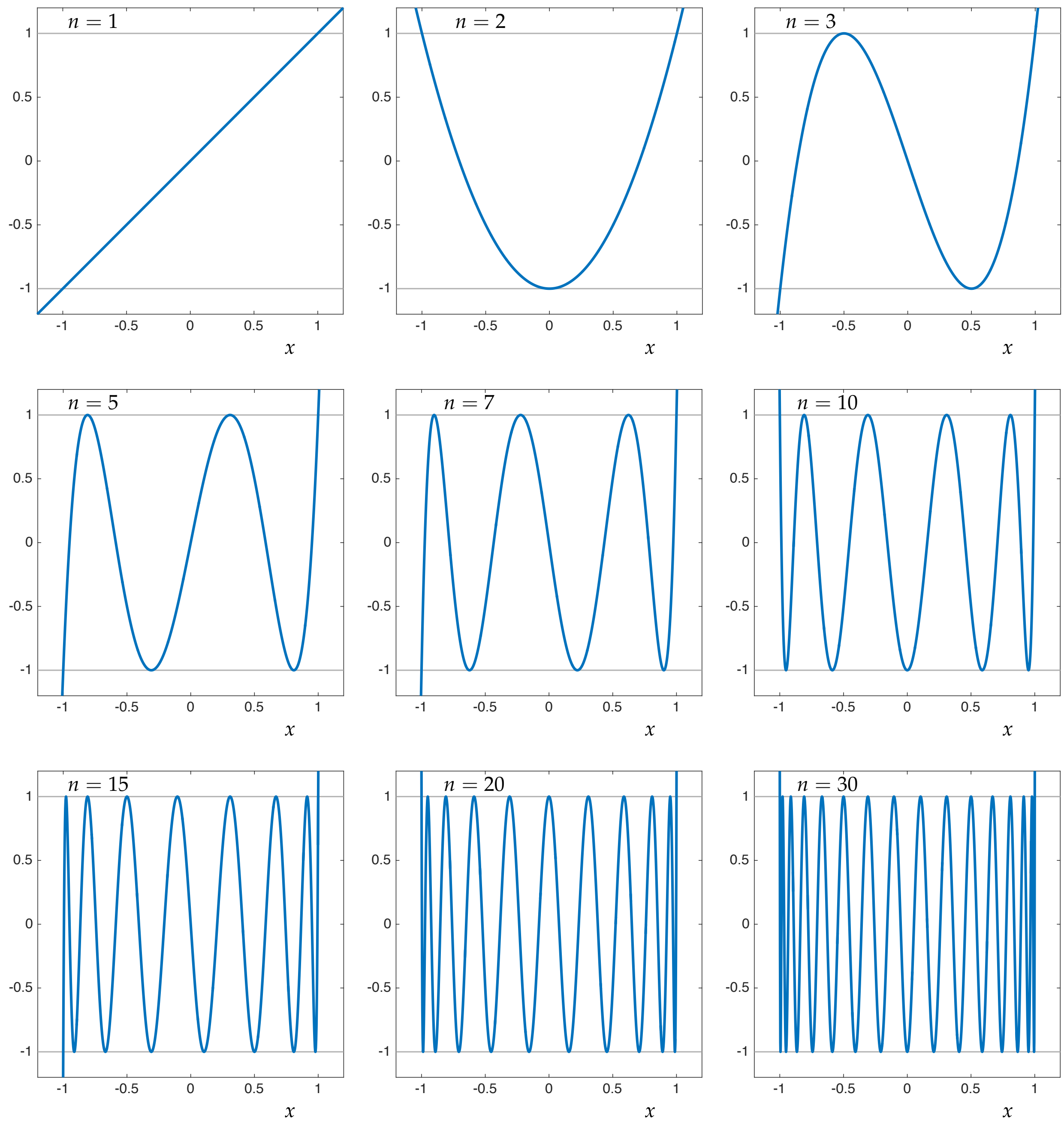


Figure 2.5: Chebyshev polynomials T_n of degree $n = 1, 2, 3$ (top), $n = 5, 7, 10$ (middle), and $n = 15, 20, 30$ (bottom). Note how rapidly these polynomials grow outside the interval $[-1, 1]$.

for $n \geq 0$, with $\widehat{T}_0(x) = 1$. These *normalized* Chebyshev polynomials are *monic*, i.e., the leading term in $\widehat{T}_{n+1}(x)$ is x^{n+1}, rather than $2^n x^{n+1}$ as for $T_{n+1}(x)$. Thus, we can write

$$\widehat{T}_{n+1}(x) = x^{n+1} - r_n(x)$$

for some polynomial $r_n(x) = x^{n+1} - \widehat{T}_{n+1}(x) \in \mathcal{P}_n$. We do not especially care about the particular coefficients of this r_n; our quarry will be the *roots* of $\widehat{T}_{n+1}$, the optimal interpolation points.

For $n \geq 0$, the polynomials $\widehat{T}_{n+1}(x)$ oscillate between $\pm 2^{-n}$ for $x \in [-1, 1]$, with the maximal values attained at

$$\eta_j = \cos\left(\frac{j\pi}{n+1}\right)$$

for $j = 0, \ldots, n+1$. In particular,

$$\widehat{T}_{n+1}(\eta_j) = (\eta_j)^{n+1} - r_n(\eta_j) = (-1)^j 2^{-n}.$$

Thus, we have found a polynomial $r_n \in \mathcal{P}_n$, together with $n+2$ distinct points, $\eta_j \in [-1, 1]$ where the maximum error

$$\max_{x \in [-1,1]} |x^{n+1} - r_n(x)| = 2^{-n}$$

is attained with alternating sign. Thus, by the oscillation theorem, we have found the minimax approximation to x^{n+1}.

Theorem 2.5 (Optimal approximation of x^{n+1})**.**
The optimal approximation to x^{n+1} from $\mathcal{P}_n$ for $x \in [-1, 1]$ is

$$r_n(x) = x^{n+1} - \widehat{T}_{n+1}(x) = x^{n+1} - 2^{-n} T_{n+1}(x) \in \mathcal{P}_n.$$

Thus, the optimal interpolation points are those $n+1$ roots of $x^{n+1} - r_n(x)$, that is, the roots of the degree-$(n+1)$ Chebyshev polynomial:

$$\xi_j = \cos\left(\frac{(2j+1)\pi}{2n+2}\right), \quad j = 0, \ldots, n.$$

For generic intervals $[a, b]$, a change of variable demonstrates that the same points, appropriately shifted and scaled, will be optimal.

Similar properties hold if interpolation is performed at the $n+1$ points

$$\eta_j = \cos\left(\frac{j\pi}{n}\right), \quad j = 0, \ldots, n,$$

which are also called Chebyshev points and are perhaps more popular due to their slightly simpler formula. (We used these points to successfully interpolate Runge's function, scaled to the interval $[-5, 5]$.) While these points differ from the roots of the Chebyshev polynomial, they *have the same distribution* as $n \to \infty$. That is the key.

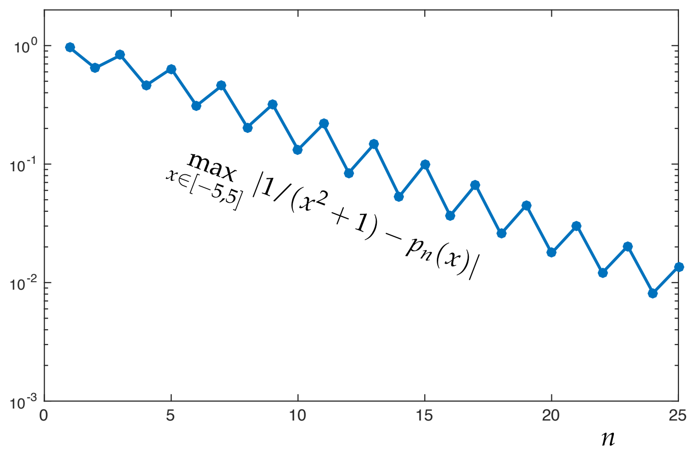
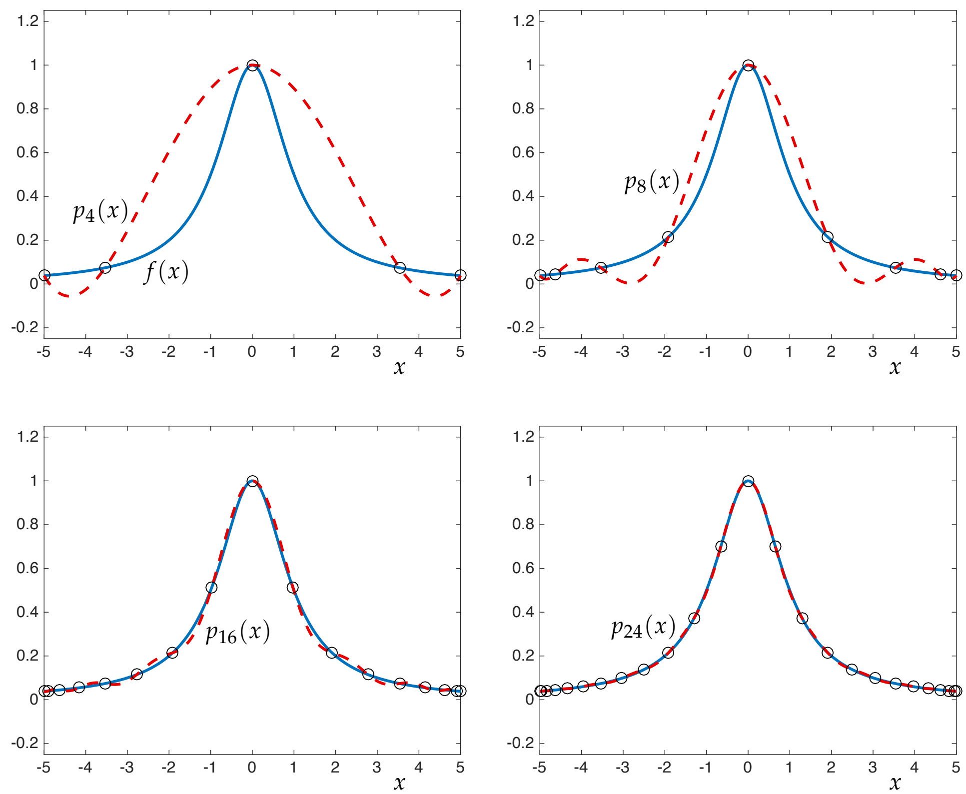


Figure 2.6: Repetition of Figure 1.8, interpolating Runge's function $1/(x^2+1)$ on $x \in [-5,5]$, but now using Chebyshev points $x_j = 5\cos(j\pi/n)$. The top plot shows this convergence for $n = 0,\ldots,25$; the bottom plots show the interpolating polynomials p_4, p_8, p_{16}, and p_{24}, along with the interpolation points that determine these polynomials (black circles). Unlike interpolation at uniformly spaced points, these interpolants *do converge* to f as $n \to \infty$. Notice how the interpolation points cluster toward the ends of the domain $[-5,5]$.



We emphasize the utility of interpolation at Chebyshev points by quoting the following result from Trefethen's excellent *Approximation Theory and Approximation Practice* (SIAM, 2013). Trefethen emphasizes that worst-case results like Faber's theorem (Theorem 1.4) give misleadingly pessimistic concerns about interpolation. If the function $f \in C[a,b]$ has just a bit of smoothness (i.e., bounded derivatives), interpolation in Chebyshev points is 'bulletproof'. The following theorem consolidates aspects of Theorem 7.2 and 8.2 in Trefethen's book.

The results are stated for $[a, b] = [-1, 1]$ but can be adapted to any real interval.

Theorem 2.6 (Convergence of Interpolants at Chebyshev Points).

For any $n > 0$, let p_n denote the interpolant to $f \in C[-1, 1]$ at the Chebyshev points

$$x_j = \cos\left(\frac{j\pi}{n}\right), \quad j = 0, \ldots, n.$$

- Suppose $f \in C^\nu[-1, 1]$ for some $\nu \geq 1$, with $f^{(\nu)}$ having variation $V(\nu)$, i.e.,

  $$V(\nu) := \max_{x\in[-1,1]} f^{(\nu)}(x) - \min_{x\in[-1,1]} f^{(\nu)}(x).$$

  Then for any $n > \nu$,

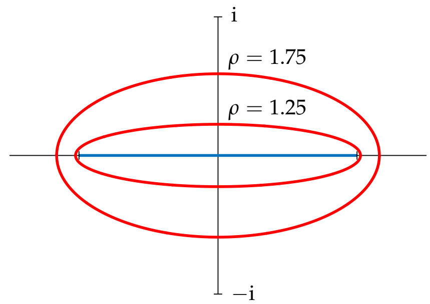
  $$\|f - p_n\|_\infty \leq \frac{4V(\nu)}{\pi(\nu(n-\nu)^\nu)}.$$

- Suppose f is *analytic* on $[-1, 1]$ and can be analytically continued (into the complex plane) onto the region bounded by the ellipse

  $$E_\rho := \left\{\frac{\rho e^{i\theta} + e^{-i\theta}/\rho}{2} : \theta \in [0, 2\pi)\right\}.$$

  Suppose further that $|f(z)| \leq M$ on and inside E_ρ. Then

  $$\|f - p_n\|_\infty \leq \frac{2M\rho^{-n}}{\rho - 1}.$$



Interval $[-1, 1]$ (blue), with two ellipses E_ρ for $\rho = 1.25$ and $\rho = 1.75$.

For example, the first part of this theorem implies that if f' exists and is bounded, then $\|f - p_n\|_\infty$ must converge at least as fast as $1/n$ as $n \to \infty$. While that is not such a fast rate, it does indeed show convergence of the interpolant. The second part of the theorem ensures that if f is well behaved in the region of the complex plane around $[-1, 1]$, the convergence will be extremely fast: the larger the area of $\mathbb{C}$ in which f is well behaved, the faster the convergence.

2.3.2 *Chebyshev polynomials beyond* $[-1, 1]$

Another way of interpreting the equioscillating property of Chebyshev polynomials is that T_n solves the approximation problem

$$\|T_n\|_\infty = \min_{\substack{p\in\mathcal{P}_n \\ p \text{ monic}}} \|p\|_\infty,$$

over the interval $[-1, 1]$, where a polynomial is *monic* if it has the form $x^n + q(x)$ for $q \in \mathcal{P}_{n-1}$.

In some applications, such as the analysis of iterative methods for solving large-scale systems of linear equations, one needs to bound the size of the Chebyshev polynomial *outside the interval* $[-1,1]$. Figure 2.5 shows that T_n grows very quickly outside $[-1,1]$, even for modest values of n. How fast?

To describe Chebyshev polynomials outside $[-1,1]$, we must replace the trigonometric functions in the definition $T_n(x) = \cos(n\cos^{-1} x)$ with hyperbolic trigonometric functions:

$$T_n(x) = \cosh(n\cosh^{-1} x), \qquad x \notin (-1,1). \tag{2.10}$$

Is this definition is consistent with

$$T_n(x) = \cos(n\cos^{-1} x), \qquad x \in [-1,1]$$

used previously? Trivially one can see that the new definition also gives $T_0(x) = 1$ and $T_1(x) = x$. Like standard trigonometric functions, the hyperbolic functions also satisfy the addition formula

$$\cosh\alpha + \cosh\beta = 2\cosh\Big(\frac{\alpha+\beta}{2}\Big)\cosh\Big(\frac{\alpha-\beta}{2}\Big),$$

and so

$$\cosh\big((n+1)\theta\big) = 2\cosh\theta\cosh n\theta - \cosh\big((n-1)\theta\big),$$

leading to the same three-term recurrence as before:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

Thus, the definitions are consistent.

We would like a more concrete formula for $T_n(x)$ for $x \notin [-1,1]$ than we could obtain from the formula (**??**). Thankfully Chebyshev polynomials have infinitely many interesting properties to lean on. Consider the change of variables

$$x = \frac{w+w^{-1}}{2},$$

which allows us to write

$$x = \frac{\mathrm{e}^{\log w} + \mathrm{e}^{-\log w}}{2} = \cosh(\log w).$$

Thus work from the definition to obtain

$$\begin{aligned} T_n(x) &= \cosh(n\cosh^{-1}(x)) \\ &= \cosh(n\log w) \\ &= \cosh(\log w^n) = \frac{\mathrm{e}^{\log(w^n)} + \mathrm{e}^{-\log(w^n)}}{2} = \frac{w^n + w^{-n}}{2}. \end{aligned}$$

We emphasize this last important formula:

$$T_n(x) = \frac{w^n + w^{-n}}{2}, \qquad x = \frac{w + w^{-1}}{2} \notin (-1, 1). \tag{2.11}$$

We have thus shown that $|T_n(x)|$ *will grow exponentially in n for any $x \notin (-1, 1)$ for which $|w| \neq 1$.* When does $|w| = 1$? Only when $x = \pm 1$. Hence,

$$|T_n(x)| \text{ grows exponentially in } n \text{ for all } x \notin [-1, 1].$$

Example 2.3. We want to evaluate $T_n(2)$ as a function of n. First, find w such that $2 = (w + w^{-1})/2$, i.e.,

$$w^2 - 4w + 1 = 0.$$

Solve this quadratic for

$$w_\pm = 2 \pm \sqrt{3}.$$

We take $w = 2 + \sqrt{3} = 3.7320\ldots$. Thus by (2.11)

$$T_n(2) = \frac{(2 + \sqrt{3})^n + (2 - \sqrt{3})^n}{2} \approx \frac{(2 + \sqrt{3})^n}{2}$$

as $n \to \infty$, since $(2 - \sqrt{3})^n = (0.2679\ldots)^n \to 0$.

Which $\pm$ choice should you make? It does not matter. Notice that $(2 - \sqrt{3})^{-1} = 2 + \sqrt{3}$, and this happens in general: $w_\pm = 1/w_\mp$.

Take a moment to reflect on this: We have a beautifully concrete way to write down $|T_n(x)|$ that does not involve any hyperbolic trigonometric formulas, or require use of the Chebyshev recurrence relation. Formulas of this type can be very helpful for analysis in various settings. You will see one such example on Problem Set 3.