THE LIMITING DISTRIBUTION OF A RANDOM VARIABLE TRANSFORMED BY CHEBYSHEV POLYNOMIALS

JAVIER CHICO VÁZQUEZ* AND ANDREW J. HORNING †

Abstract. In this paper we present the result of successively applying a Chebyshev polynomial to a continuous random variable. In particular we show that under mild assumptions the limiting distribution will be the same as the weight with respect to which Chebyshev polynomials are orthogonal, namely $\sim (1 - t^2)^{-\frac{1}{2}}$.

Key words. Chebyshev polynomials, Convergence in distribution

1. Introduction. This paper investigates the behaviour of a random variable X when it is transformed by Chebyshev polynomials. Mathematically, we are interested in the distribution of $T_k(X)$, in particular on how this behaves as $k \to \infty$. In particular, we will show that this converges in distribution to a continuous random variable whose density function is the same as the weight function under which Chebyshev polynomials are orthogonal. Further, we show that the limiting distribution is invariant when transformed by T_k.

The structure of this paper is the following: First we review necessary concepts in section 2. Next we state the main result in section 3, with an interesting asymptotic analysis of the error and the statement for a lemma which serves as a cornerstone for the proof. This is followed by a presentation of the numerical results. In the next section a rigorous proof for the theorem is presented, and the final remarks include results that seem to hold from the numerics but have not been shown formally.. In the final appendix the proof for the lemma is presented.

2. Preliminaries. We introduce some important definitions and results before presenting the main results. First, for convenience, we define the kth Chebyshev polynomial of the first kind as [6, 2, 4]

$$(2.1) \qquad T_k(x) = \cos(k \cos^{-1}(x))$$

where $\cos^{-1} = \arccos$. Chebyshev polynomials are orthogonal in the following sense: [2]

$$(2.2) \qquad \int_{-1}^{1} T_k(x) T_l(x) w(x) \mathrm{d}x = 0 \text{ when } k \neq l \text{ and } w(x) = \frac{1}{\sqrt{1 - x^2}}$$

The integral of a Chebyshev polynomial over $[-1, 1]$ is given by [1]

$$(2.3) \qquad \int_{-1}^{1} T_k(x) \mathrm{d}x = \begin{cases} \frac{(-1)^k + 1}{1 - k^2} & k \neq 1 \\ 0 & k = 1 \end{cases}$$

The following identities for the sums of cosines and sines will be useful later.

---
*Imperial College London
†Massachusetts Institute of Technology

These are standard results from trigonometry [1].

$$(2.4a) \qquad \sum_{j=1}^{n} \cos(jx) = \frac{\sin(nx/2)\cos((n+1)x/2)}{\sin(x/2)}$$

$$(2.4b) \qquad \sum_{j=1}^{n} \sin(jx) = \frac{\sin(nx/2)\sin((n+1)x/2)}{\sin(x/2)}$$

Once the above has been collected we can proceed by stating the main results and the proof.

3. Main results. As mentioned previously, we are interested in the distribution of $T_k(X)$ as k gets large. In particular, under mild assumptions, the following fact holds: $T_k(X)$ will converge in distribution to a random variable whose density function is the same as the weight function for Chebyshev polynomials.

THEOREM 3.1. *Suppose X is a continuous random variable with density function f_X on $[-1,1]$. As it is continuous, it can be expanded as a Chebyshev series that converges absolutely [6, 5]*

$$(3.1) \qquad f_X(x) = \sum_{l=0}^{\infty} \mu_l T_l(x)$$

then $T_k(X)$ converges in distribution to a random variable Γ with probability density function as $k \to \infty$

$$f_\Gamma(z) = \frac{1}{\pi} \frac{1}{\sqrt{1-z^2}}$$

This limiting random variable Γ is a scaled and shifted Beta distribution with parameters $\alpha = \beta = \frac{1}{2}$. In particular, if $Y \sim \text{Beta}(\alpha = \frac{1}{2}, \beta = \frac{1}{2})$, then $\Gamma = 2Y - 1$.

For a more general X defined on $[a,b]$ we can re-scale the Chebyshev polynomials on that interval, with a suitably re-defined weight function as well. The theorem rests on the following result, which provides us with an analytic formula for the probability density function of $T_k(X)$

LEMMA 3.2. *Let f_X and F_X be the pdf and cdf of X, a continuous random variable define on $[-1,1]$. Then the probability density function of $T_k(X)$ when $k = 2m$ is even is*

$$(3.2) \qquad \begin{aligned} f_k(z) &= \frac{1}{\sqrt{1-z^2}} \sum_{j=1}^{m=\frac{k}{2}} \left\{ f_{\Psi_k}(2\pi j - \cos^{-1}(z)) + f_{\Psi_k}(2\pi(j-1) + \cos^{-1}(z)) \right\} \\ &= \frac{1}{\sqrt{1-z^2}} S_k(z) \end{aligned}$$

where $S_k(z)$ is defined as the sum in the top line. When $k = 2m + 1$ is odd there is an extra term in the sum

$$\begin{aligned} f_k(z) = &\frac{1}{\sqrt{1-z^2}} \sum_{j=1}^{m=\frac{k-1}{2}} \left\{ f_{\Psi_k}(2\pi j - \cos^{-1}(z)) + f_{\Psi_k}(2\pi(j-1) + \cos^{-1}(z)) \right\} \\ &+ \frac{1}{\sqrt{1-z^2}} f_{\Psi_k}(2\pi m + \cos^{-1}(z)) \end{aligned}$$

where f_{Ψ_k} is the pdf of $\Psi_k = k\cos^{-1}(X)$, so that

$$f_{\Psi_k}(\psi) = \frac{1}{k} f_X(\cos(\psi/k))\sin(\psi/k)$$

3.1. Asymptotic analysis of the convergence. If we expand the expression for the complete sum in the general case, (5.3a), around $k \to \infty$ we obtain the following expansion,

(3.4)

$$S_k(z) \approx \frac{1}{\pi} + \frac{1}{k^2}\left(\frac{\pi}{3} - \frac{(\pi - \cos^{-1}(z))^2}{\pi}\right)\left[\mu_0 + \sum_{l=2}^{\infty}\left\{\mu_l \cos\left(\frac{l\pi}{2}\right)^2\right\}\right] + \mathcal{O}\left(\frac{1}{k^4}\right)$$

$$= \frac{1}{\pi} + \frac{1}{k^2}\left(\frac{\pi}{3} - \frac{(\pi - \cos^{-1}(z))^2}{\pi}\right)\sum_{l=0}^{\infty}\mu_{2l} + \mathcal{O}\left(\frac{1}{k^4}\right)$$

In particular, we see that the error is quadratic in $\frac{1}{k}$, thus explaining the rapid convergence found numerically in the upcoming sections. Furthermore, only the even Chebyshev moments contribute to the error.

4. Numerical Results.

4.1. Dance around the origin. For distributions where the bulk of the probability mass is centered at the origin (i.e. Gaussian, Cauchy, ...), the second Chebyshev polynomial $T_2(t) = 2t^2 - 1$ will map most of that probability mass to $T_2(0) = -1$, so the distribution will become skewed to the left. From there on, as $T_k(1) = 1$ and $T_k(-1) = (-1)^k$, the probability mass will remain mostly on $t = -1$ for $k = 3$, but will migrate to $t = 1$ at $k = 4$, and then oscillate between -1 and 1 changing place every two iterations, but loosing skewedness as k increases, as we know that the distribution eventually converges to $\frac{1}{\pi\sqrt{1-t^2}}$ which is symmetric. This oscillation is visualized in Figure 1, where the initial distribution is a Gaussian.
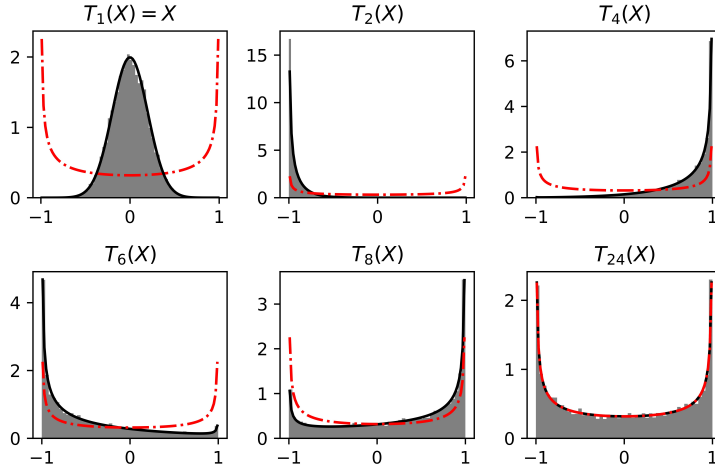

FIG. 1. *Dance around the origin for an initial Gaussian distribution. The solid black lines represent the analytic solution as given by Lemma 3.2 and the dashed red line represents the limiting distribution*

In Figure 1 we further observe that after only 24 iterations the distribution has essentially converged (very quickly!), indeed, as we will see in the next section this is because the convergence is quadratic in $\frac{1}{k}$, so that the error decreases with $\mathcal{O}(\frac{1}{k^2})$.

5. Proofs for the main results. Before proving the general case we will show the proof for the case $X \sim \Gamma$ (in this particular case, instead of convergence we will have that the distribution is invariant, i.e. $\Gamma = T_k(\Gamma)$ in distribution. For this it is necessary to obtain a general formula for the probability density function of $T_k(X)$, denoted as $f_k(z)$. More formally:

THEOREM 5.1 (The invariant measure). *If* $X \sim \text{Beta}(1/2, 1/2)$ *then* $T_k(X) \sim$ Beta$(1/2, 1/2)$. *This means that the distribution* Beta$(1/2, 1/2)$ *is invariant under the transformation* $T_k : [-1, 1] \longrightarrow [-1, 1]$

Proof. In this case

$$f_X(x) = \frac{1}{\pi\sqrt{1 - x^2}}$$

then

$$f_{\Psi_k}(\psi) = f_\Theta(\psi/k)/k = \frac{1}{k}f_X(\cos(\psi/k))\sin(\psi/k) = \frac{\sin(\psi/k)}{\pi k\sqrt{1 - \cos^2(\psi/k)}} = \frac{1}{k\pi}$$

So this is a Uniform distribution!! (which depends on k) Finally, as $f_{\Psi_k}(\psi)$ is constant, the sum is really simple in the case k even.

$$f_k(z) = \frac{1}{\sqrt{1 - z^2}}\sum_{j=1}^{m} f_{\Psi_k}(2\pi j - \cos^{-1}(z)) + f_{\Psi_k}(2\pi(j-1) + \cos^{-1}(z)) =$$

$$= \frac{1}{\sqrt{1 - z^2}}\sum_{j=1}^{k/2}\left(\frac{1}{k\pi} + \frac{1}{k\pi}\right) = \frac{2}{k\pi\sqrt{1 - z^2}}\sum_{j=1}^{k/2} 1 = \frac{1}{\pi\sqrt{1 - z^2}}$$

For k odd we will just have one extra term (but k is one unit greater) so the result is the same and we are done. □

We can now turn our attention to the proof for the general statement. We will do this by first exploring the constraint on the Chebyshev moments that arises from normalisation, and we will use that to our advantage to show that the limit in (7.2) is indeed $\frac{1}{\pi}$.

General case. We have that

$$f_X(x) = \sum_{l=0}^{\infty} \mu_l T_l(x)$$

But f_X is a pdf so in particular it satisfies

$$1 = \int_{-1}^{1} f_X(x)dx = \sum_{l=0}^{\infty} \mu_l \int_{-1}^{1} T_l(x)dx$$

We can use (2.3) so that the constraint becomes

(5.1) $$1 = 2\mu_0 + \sum_{l=2}^{\infty} \mu_l \frac{(-1)^l + 1}{1 - l^2}$$

This will be useful later. We will now proceed as in the simpler cases, by computing the limit in (7.2).

$$f_{\Psi_k}(\psi) = \frac{\sin(\psi/k)}{k} f_X(\cos(\psi/k)) = \frac{\sin(\psi/k)}{k} \sum_{l=0}^{\infty} \mu_l T_l(\cos(\psi/k))$$

But $T_l(\cos(\psi/k)) = \cos\left(\frac{l}{k}\psi\right)$ so that

$$f_{\Psi_k}(\psi) = \frac{\sin(\psi/k)}{k} \sum_{l=0}^{\infty} \mu_l \cos(l\psi/k)$$

If we focus on $k = 2m$ even, then the limit is

$$\sum_{j=1}^{k/2} f_{\Psi_k}(2\pi j - \cos^{-1}(z)) + f_{\Psi_k}(2\pi(j-1) + \cos^{-1}(z)) =$$

(5.2)

$$= \frac{1}{k} \sum_{j=1}^{k/2} \sum_{l=0}^{\infty} \mu_l \sin(x_k j - \beta_k) \cos(l(x_k j - \beta_k)) + \sin((x_k(j-1) + \beta_k) \cos(l(x_k(j-1) + \beta_k)$$

defining $\beta_k = \cos^{-1}(z)$ as before and $x_k = \frac{2\pi}{k}$. We can exchange the order of summation as the infinite series converges absolutely.

$$\sum_{l=0}^{\infty} \mu_l \frac{1}{k} \sum_{j=1}^{k/2} \{\sin(x_k j - \beta_k) \cos(l(x_k j - \beta_k)) + \sin(x_k(j-1) + \beta_k) \cos(l(x_k(j-1) + \beta_k))\}$$

thus, we can focus by fixing l and computing the limit of the inner sum as k gets large. We will do this for three separate cases: $l = 0, l = 1$ and $l \geq 2$. For $l = 0$ the sum is simplified as the cosines have zero argument so become factors of 1. Thus we have to consider

$$\frac{1}{k} \sum_{j=1}^{k/2} \{\sin(x_k j - \beta_k) + \sin(x_k(j-1) + \beta_k)\} = \frac{2\cos\left(\frac{\pi - \cos^{-1}(z)}{k}\right)}{k \sin(\pi/k)}$$

In the limit as $k \to \infty$ this becomes $\frac{2}{\pi}$. We can see this because $\sin(\pi/k) \sim \frac{\pi}{k}$ and the cosine in the denominator converges to 1 as the argument converges to zero.

Moving on to the case $l = 1$ we can regroup the terms and obtain that the sum (and thus the limit) is identically zero. The summand in this case can be expressed as

$$\sin(x_k j - \beta_k) \cos(x_k j - \beta_k) + \sin(x_k(j-1) + \beta_k) \cos(x_k(j-1) + \beta_k) =$$
$$= \cos(\frac{2(\pi - \cos^{-1}(z))}{k}) \sin(\frac{2\pi(2j-1)}{k})$$

So when we sum over j the first term can be moved outside as a constant and the second term, using the formula in (2.4b) gives us 0.

Finally, for the case $l \geq 2$, the sum will give us

$$\frac{\cos(l\pi/2)^2}{2k \sin(\pi/k(l+1)) \sin(\pi/k(1-l))} [\sin((2\pi + (l-1)\cos^{-1}(z))/k) +$$
$$+ \sin((2\pi l + (1-l)\cos^{-1}(z))/k) + \sin((2\pi - (l+1)\cos^{-1}(z))/k) +$$
$$+ \sin((-2\pi l + (l+1)\cos^{-1}(z))/k)]$$

As all the sines have an argument that is divided by k, in the limit they can be replaced by their argument. This means there is massive cancellation. In particular, in the limit,

$$\frac{k^2 \cos(l\pi/2)^2}{2k\pi^2(1-l^2)}[4\pi/k] = \frac{1}{\pi}\frac{1+\cos(l\pi)}{1-l^2} = \frac{1}{\pi}\frac{1+(-1)^l}{1-l^2}$$

but this is precisely the result in (2.3) multiplied by $\frac{1}{\pi}$. Upon assembling the different case we thus get that in the limit

$$\frac{2}{\pi}\mu_0 + 0 \cdot \mu_1 + \sum_{l=2}^{\infty} \mu_l \frac{1}{\pi}\frac{1+(-1)^l}{1-l^2} = \frac{1}{\pi}\left[2\mu_0 + \sum_{l=2}^{\infty}\mu_l\frac{1+(-1)^l}{1-l^2}\right] = \frac{1}{\pi} \qquad \square$$

from the constraint in (5.1), so we are done.

Before, moving on, we can reassemble the sum before passing onto the limit, for it will be useful later.

$$(5.3\text{a}) \qquad\qquad S_k(z) = \frac{2\cos\left(\frac{\pi-\cos^{-1}(z)}{k}\right)}{k\sin(\pi/k)}+$$

$$(5.3\text{b}) \qquad +\frac{\cos(l\pi/2)^2}{2k\sin(\pi/k(l+1))\sin(\pi/k(1-l))}[\sin((2\pi+(l-1)\cos^{-1}(z))/k)+$$

$$(5.3\text{c}) \qquad +\sin((2\pi l + (1-l)\cos^{-1}(z))/k) + \sin((2\pi-(l+1)\cos^{-1}(z))/k)+$$

$$(5.3\text{d}) \qquad +\sin((-2\pi l + (l+1)\cos^{-1}(z))/k)]$$

6. Final Remarks. All in all we have shown the main fact for mild assumptions on X, in particular that it has a Chebyshev Expansion. In particular, we showed that if the initial distribution is the same as the limiting then it is conserved. Furthermore, we have explored the oscillatory behaviour at the initial stages of the convergence.

In the future it would be interesting to analytically study the convergence of discontinuous distributions, such as

$$f_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{else} \end{cases}$$

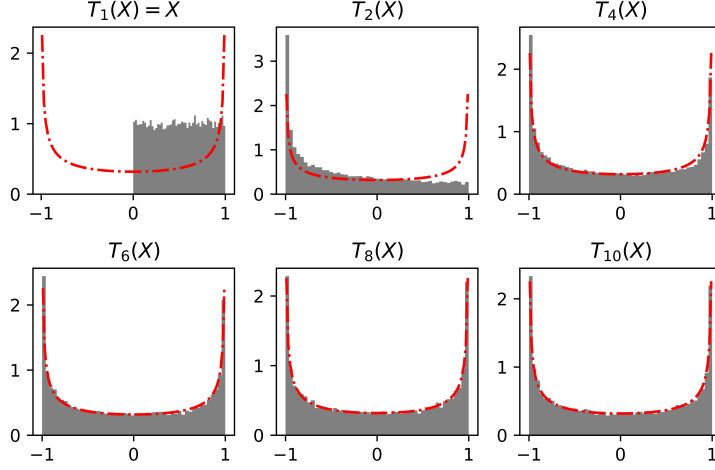as our numerical testing strongly suggests the same result applies, as seen in Figure 2

FIG. 2. *Numerical evidence for the convergence for a discontinuous random variable.*

Furthermore, we would like to explore how this result can be used to accelerate bayesian methods spectral methods like [3]

7. Appendix. A constructive proof of Lemma 3.2 is first given, as it is essential for the proof of the main theorem.

Proof. We want to compute

$$T_k(X) = \cos(k \cos^{-1}(X))$$

From **??** we can motivate computing this quantity in two steps. First, figuring out the distribution of the "inner" part, $k \cos^{-1}(X)$, and then transforming this with the outer cosine. Thus, let $\Theta = \cos^{-1}(X)$. Then

$$F_\Theta(\theta) = \Pr(\Theta \le \theta) = \Pr(\cos^{-1}(X) \le \theta) = \Pr(X \ge \cos(\theta))$$

as the cosine is monotonically decreasing in arccosine [-1,1], so we change the sign of the inequality. Then

$$F_\Theta(\theta) = 1 - \Pr(X \le \cos(\theta)) = 1 - F_X(\cos(\theta))$$

and thus, taking a derivative

$$f_\Theta(\theta) = f_X(\cos(\theta)) \sin \theta$$

Now define $\Psi_k = k\Theta$. We can easily get the density and cumulative density of this new random variable, as it is just Θ scaled

$$F_{\Psi_k}(\psi) = F_\Theta(\psi/k) = 1 - F_X(\cos(\theta/k))$$

and taking a derivative

$$f_{\Psi_k}(\psi) = f_\Theta(\psi/k)/k = \frac{1}{k} f_X(\cos(\psi/k)) \sin(\psi/k)$$

so we have completed the first step of the process. Now we can compute, for $T_k(X) = \cos(\Psi_k)$. However, we will have to be more careful with the transformation. Let F_k be the cdf of $T_k(X)$, and f_k its pdf as mentioned in the statement. Then

$$F_k(z) = \Pr(\cos(\Psi_k) \leq z)$$

Let's look at that inequality in more detail in Figure 3. For now, assume k is even:
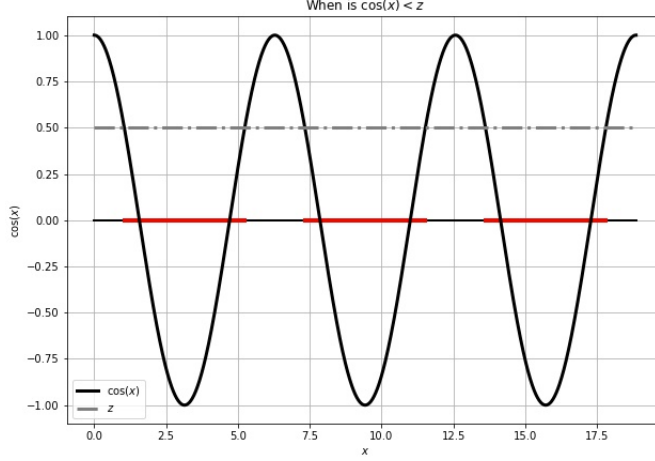

FIG. 3. *Cosine inequality visualized. It is satisfied on the regions of the x-axis highlighted in red.*

So we can see it is satisfied in the intervals:

$$\cos^{-1}(z) \leq \Psi_k \leq 2\pi - \cos^{-1}(z)$$

$$2\pi + \cos^{-1}(z) \leq \Psi_k \leq 4\pi - \cos^{-1}(z)$$

$$\vdots$$

$$2\pi(j-1) + \cos^{-1}(z) \leq \Psi_k \leq 2\pi j - \cos^{-1}(z)$$

$$\vdots$$

$$2(k-1)\pi + \cos^{-1}(z) \leq \Psi_k \leq k\pi - \cos^{-1}(z)$$

Let $m = k/2$. Then, as the intervals are disjoint, we have that

$$F_k(z) = \Pr(\cos(\Psi_k) \leq z) = \sum_{j=1}^{m} \Pr(2\pi(j-1) + \cos^{-1}(z) \leq \Psi_k \leq 2\pi j - \cos^{-1}(z))$$

We can express the probability for each summand in terms of the cdf of Ψ_k:

$$\Pr(2\pi(j-1) + \cos^{-1}(z) \leq \Psi_k \leq 2\pi j - \cos^{-1}(z)) = \Pr(\Psi_k \leq 2\pi j - \cos^{-1}(z)) - \Pr(\Psi_k \leq 2\pi(j-1) + \cos^{-1}(z))$$

$$= F_{\Psi_k}(2\pi j - \cos^{-1}(z)) - F_{\Psi_k}(2\pi(j-1) + \cos^{-1}(z))$$

So we get

$$F_k(z) = \sum_{j=1}^{m} F_{\Psi_k}(2\pi j - \cos^{-1}(z)) - F_{\Psi_k}(2\pi(j-1) + \cos^{-1}(z))$$

taking a derivative

(7.1a)
$$f_k(z) = \sum_{j=1}^{m} f_{\Psi_k}(2\pi j - \cos^{-1}(z))\frac{1}{\sqrt{1-z^2}} + f_{\Psi_k}(2\pi(j-1) + \cos^{-1}(z))\frac{1}{\sqrt{1-z^2}} =$$

(7.1b)
$$= \frac{1}{\sqrt{1-z^2}} \sum_{j=1}^{m} f_{\Psi_k}(2\pi j - \cos^{-1}(z)) + f_{\Psi_k}(2\pi(j-1) + \cos^{-1}(z))$$

Which is what we wanted to show.   □

Hence we can already see the main part of the pdf we want to reach in the limit $k \to \infty$. Therefore, to proof the statement we just need to show that the sum converges (pointwise) to $1/\pi$. Thus we want to show that for a fixed z,

(7.2)
$$\lim_{k\to\infty} \left\{ \sum_{j=1}^{k/2} f_{\Psi_k}(2\pi j - \cos^{-1}(z)) + f_{\Psi_k}(2\pi(j-1) + \cos^{-1}(z)) \right\} = \frac{1}{\pi}$$

For k odd we just have an extra term in the sum. Before proving the general case we will look at two particular examples of great interest: X uniformly distributed and $X \sim \text{Beta}(1/2, 1/2)$. For convenience, we define this sum we are taking the limit as $S_k(z)$

$$S_k(z) = \sum_{j=1}^{k/2} f_{\Psi_k}(2\pi j - \cos^{-1}(z)) + f_{\Psi_k}(2\pi(j-1) + \cos^{-1}(z))$$

REFERENCES

[1] *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Wiley, New York ;, 1964 edition ed., 1972 - 1964.
[2] J. P. BOYD, *Chebyshev and Fourier Spectral Methods Second Revised Edition*, Dover Books on Mathematics, Dover Publications, Newburyport, 2nd ed. ed., 2013.
[3] A. CADONNA, A. KOTTAS, AND R. PRADO, *Bayesian mixture modeling for spectral density estimation*, Statistics probability letters, 125 (2017), pp. 189–195.
[4] L. FOX, *Chebyshev polynomials in numerical analysis*, Oxford mathematical handbooks, Oxford U.P, 1968.
[5] M. J. D. M. J. D. POWELL, *Approximation theory and methods*, Cambridge University Press, Cambridge, 1981.
[6] L. N. L. N. TREFETHEN, *Approximation theory and approximation practice*, SIAM, Society for Industrial and Applied Mathematics, Philadelphia, extended edition. ed., 2020.