

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 $\sim (1 - t^2)^{-1}$.

Key words. Chebyshev Polynomials, Convergence in distribution

1. Introduction. Chebyshev Polynomials are an essential tool for many areas of Numerical Analysis. In particular, when estimating the spectral density of large matrices we use iterative methods to approximate the Chebyshev Moments of the density of states, which is defined as

$$\nu_A(t) = \sum_{j=1}^N \delta(t - \lambda_j)$$

where λ_j are the (real) eigenvalues of the (symmetric) matrix $A \in \mathbb{R}^{N \times N}$. In particular, the Kernel Polynomial Method approximates the scaled density of states as a Chebyshev Series

$$\hat{\phi}(t) = \sqrt{1 - t^2} \nu_A(t) = \sum_{k=1}^{\infty} \mu_k T_k(t)$$

Dividing by $\sqrt{1 - t^2}$, multiplying by $T_l(t)$ and integrating in $[-1, 1]$, and exploiting the orthogonality of Chebyshev polynomials the Chebyshev coefficients can be estimated by

$$\mu_k \sim \frac{1}{N} \sum_{j=1}^N T_k(\lambda_j) = \frac{1}{N} \text{Trace}(A)$$

In the above, we see why we would be interested in understanding in the average value of $T_k(\lambda_j)$ as k gets large, and even in how the eigenvalue distribution changes as it gets transformed by T_k . In this paper we will present what happens to a continuous random variable X when it is transformed via Chebyshev Polynomials. Mathematically, we are interested in the distribution of $T_k(X)$, and we are especially focused on how this behaves as $k \rightarrow \infty$.

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](#), followed by the proof for two simpler cases. At the end of the section a proof for the general case is given.

2. Preliminaries.

DEFINITION 2.1 (Chebyshev polynomials). *We define the k -th polynomial by*

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

where $\cos^{-1} = \arccos$.

*Imperial College London

†Massachusetts Institute of Technology

LEMMA 2.2. *The following sums of cosines and sines will be useful later. These are standard results.*

$$(2.2a) \quad \sum_{j=1}^n \cos(jx) = \frac{\sin(nx/2) \cos((n+1)x/2)}{\sin(x/2)}$$

$$(2.2b) \quad \sum_{j=1}^n \sin(jx) = \frac{\sin(nx/2) \sin((n+1)x/2)}{\sin(x/2)}$$

LEMMA 2.3 (Definite integral of a Chebyshev Polynomial). *The integral of a Chebyshev polynomial over $[-1, 1]$ is given by*

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

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.

THEOREM 3.1. *Suppose X is a continuous random variable with density function f_X that can be expanded as a Chebyshev series*

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

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

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

This 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 $\Gamma = 2Y - 1$.

Before proving the general case we will show the proof for the case $X \sim \text{Uniform}(-1, 1)$, and the case $X \sim \Gamma$ (in this particular case, instead of convergence we will have that the distribution is conserved, i.e. $\Gamma = T_k(\Gamma)$). For this it is necessary to obtain a general formula for the probability density function of $T_k(X)$, denoted as $f_k(z)$.

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*

$$(3.2) \quad f_k(z) = \frac{1}{\sqrt{1 - z^2}} \sum_{j=1}^{m=\frac{k}{2}} \{f_{\Psi_k}(2\pi j - \cos^{-1}(z)) + f_{\Psi_k}(2\pi(j-1) + \cos^{-1}(z))\}$$

and when $k = 2m + 1$ is odd we have one more term in the sum

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

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

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

Proof. We want to compute

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

From [Definition 2.1](#) 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

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

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

$$F_{\Theta}(\theta) = 1 - \Pr(X \leq \cos(\theta)) = 1 - F_X(\cos(\theta))$$

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 1](#). For now, assume k is even:

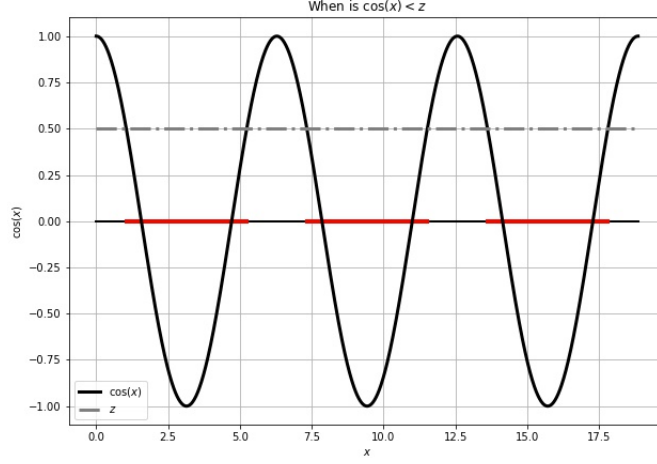


FIG. 1. *Cosine inequality visualized. It is satisfied on the regions of the x-axis highlighted in red.*

So we can see it is satisfied in the intervals:

$$\begin{aligned}
 \cos^{-1}(z) &\leq \Psi_k \leq 2\pi - \cos^{-1}(z) \\
 2\pi + \cos^{-1}(z) &\leq \Psi_k \leq 4\pi - \cos^{-1}(z) \\
 &\vdots \\
 2\pi(j-1) + \cos^{-1}(z) &\leq \Psi_k \leq 2\pi j - \cos^{-1}(z) \\
 &\vdots \\
 (k-2)\pi + \cos^{-1}(z) &\leq \Psi_k \leq k\pi - \cos^{-1}(z)
 \end{aligned}$$

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 have

$$\begin{aligned}
 \Pr(2\pi(j-1) + \cos^{-1}(z) \leq \Psi_k \leq 2\pi j - \cos^{-1}(z)) &= \Pr(\Psi_k \leq 2\pi j - \cos^{-1}(z)) - \Pr(\Psi_k \leq 2\pi(j-1) + \cos^{-1}(z)) \\
 &= F_{\Psi_k}(2\pi j - \cos^{-1}(z)) - F_{\Psi_k}(2\pi(j-1) + \cos^{-1}(z))
 \end{aligned}$$

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

(3.4a)

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

$$(3.4b) \quad = \frac{1}{\sqrt{1-z^2}} \sum_{j=1}^m f_{\Psi_k}(2\pi j - \cos^{-1}(z)) + f_{\Psi_k}(2\pi(j-1) + \cos^{-1}(z))$$

Which is what we wanted to show. \square

Hence we can already see the main part of the pdf we want to reach in the limit $k \rightarrow \infty$. 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 ,

$$(3.5) \quad \lim_{k \rightarrow \infty} \left\{ \sum_{j=1}^{k/2} f_{\Psi_k}(2\pi j - \cos^{-1}(z)) + f_{\Psi_k}(2\pi(j-1) + \cos^{-1}(z)) \right\} = \frac{1}{\pi}$$

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 $X \sim \text{Beta}(1/2, 1/2)$.

COROLLARY 3.3 ($X \sim \text{Beta}(1/2, 1/2)$). *If $X \sim \text{Beta}(1/2, 1/2)$ then $T_k(X) \sim \text{Beta}(1/2, 1/2)$. This means that the distribution $\text{Beta}(1/2, 1/2)$ is invariant under the transformation $T_k : [-1, 1] \rightarrow [-1, 1]$*

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.

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

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. \square

COROLLARY 3.4 ($X \sim \text{Uniform}(-1, 1)$).

Proof. Assume k is even, so that $k = 2m$. Then $f_X(x) = 1/2$. Thus

$$f_{\Psi_k}(\psi) = \frac{\sin(\frac{\psi}{k})}{2k}$$

Then the limit to be computed is (for a fixed z , let $\beta_k = \cos^{-1}(z)/k$)

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

Recall equations 2.2b, and take $x = \frac{2\pi}{k}$ so that

$$\frac{1}{2k} \sum_{j=1}^{k/2} \sin(2\pi j/k - \beta_k) + \sin(2\pi(j-1)/k + \beta_k) = \frac{1}{2k} \sum_{j=1}^{k/2} \sin(xj - \beta_k) + \sin(x(j-1) + \beta_k)$$

This can be expanded with half angle formulas.

$$\begin{aligned} \sin(xj - \beta_k) + \sin(x(j-1) + \beta_k) &= \sin(xj) \cos \beta_k - \sin \beta_k \cos(xj) + \sin(x(j-1)) \cos \beta_k + \sin \beta_k \cos(x(j-1)) \\ &= \sin(xj) \cos \beta_k - \sin \beta_k \cos(xj) + c(\beta_k)(s(xj)c(x) - s(x)c(xj)) + s(\beta_k)(c(xj)c(x) + s(xj)s(x)) \end{aligned}$$

So when we sum over j , from the formulas above it is clear we will have a $\sin(x/2) = \sin(\pi/k)$ in the denominator.

Overall (I will write out the details for this later) [Mathematica]

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

When we take the limit of this as k gets large for a fixed z , $\cos^{-1}(z)$ is just a constant number between 0 and π , so $\beta_k \rightarrow 0$. Hence, the argument of the cosine goes to zero, and the cosine thus goes to 1, and using $\sin(\pi/k) \approx \pi/k$ for large k

$$\frac{2 \cos(\pi/k - \beta_k)}{k \sin(\pi/k)} \approx \frac{\cos(\pi/k - \beta_k)}{k \sin(\pi/k)} \approx \frac{1}{k\pi/k} = \frac{1}{\pi}$$

and we are done as we have shown the limit in (3.5) is $\frac{1}{\pi}$ □

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 (3.5) 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

$$(3.6) \quad 1 = 2\mu_0 + \sum_{l=2}^{\infty} \mu_l \frac{(-1)^l + 1}{1 - l^2}$$

This will be useful later. We will now proceed as in the simpler cases, by computing the limit in (3.5).

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

But $T_l(\cos(\psi/k)) = \cos(\frac{l}{k}\psi)$ 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

$$\begin{aligned} & \sum_{j=1}^{k/2} f_{\Psi_k}(2\pi j - \cos^{-1}(z)) + f_{\Psi_k}(2\pi(j-1) + \cos^{-1}(z)) = \\ (3.7) \quad & = \frac{1}{k} \sum_{j=1}^{k/2} \sum_{l=0}^{\infty} \mu_l \sin(x_k j - \beta_k) \cos(l(x_k j - \beta_k)) + \sin((x_k(j-1) + \beta_k) \cos(l(x_k(j-1) + \beta_k)) \end{aligned}$$

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 $k \rightarrow \infty$ 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

$$\begin{aligned} & \sin(x_k j - \beta_k) \cos(x_k j - \beta_k) + \sin(x_k(j-1) + \beta_k) \cos(x_k(j-1) + \beta_k) = \\ & = \cos\left(\frac{2(\pi - \cos^{-1}(z))}{k}\right) \sin\left(\frac{2\pi(2j-1)}{k}\right) \end{aligned}$$

So when we sum over j the first term can be moved outside as a constant and the second term, using the formula in (2.2b) gives us 0.

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

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

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

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

from the constraint in (3.6), so we are done.

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

$$(3.8a) \quad S_k(z) = \frac{2 \cos\left(\frac{\pi - \cos^{-1}(z)}{k}\right)}{k \sin(\pi/k)} +$$

$$(3.8b) \quad + \frac{\cos(l\pi/2)^2}{2k \sin(\pi/k(l+1)) \sin(\pi/k(1-l))} [\sin((2\pi + (l-1) \cos^{-1}(z))/k) +$$

$$(3.8c) \quad + \sin((2\pi l + (1-l) \cos^{-1}(z))/k) + \sin((2\pi - (l+1) \cos^{-1}(z))/k) +$$

$$(3.8d) \quad + \sin((-2\pi l + (l+1) \cos^{-1}(z))/k)]$$

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 2, where the initial distribution is a Gaussian.

In Figure 2 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. Asymptotic analysis of the convergence. If we expand the expression for the complete sum in the general case, (3.8a), around $k \rightarrow \infty$ we obtain the following expansion,

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

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

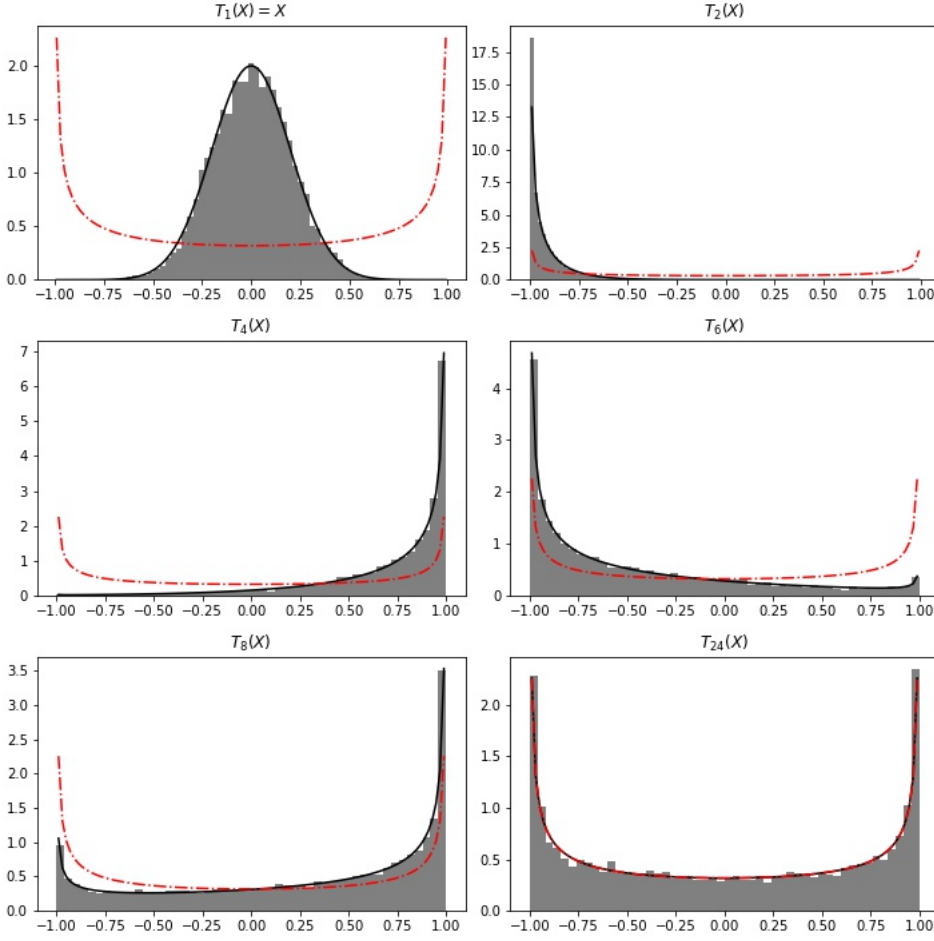


FIG. 2. Dance around the origin for an initial Gaussian distribution. The solid black lines represent the analytic solution as found in the previous section and the dashed red line represents the limiting distribution

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

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

as our numerical testing strongly suggests the same result applies.

Acknowledgments. We would like to acknowledge the MIT Mathematics Department for their support during our UROP.

REFERENCES