UNIVERSITY OF POTSDAM

BACHELOR THESIS

Investigation of Functionals of the Eigenvalues of Unitary Matrices

Author:
Carina SEIDEL

Supervisor:
Dr. Thomas MACH

July 14, 2025

Abstract

The eigenvalues of large matrices are of interest in a large variety of use cases. However, their computation becomes increasingly challenging as matrix size grows. This thesis focuses on the investigation of functionals over the eigenvalues of unitary matrices, with particular emphasis on the spectral density, or density of states (DoS) among those. While established methods exist for real symmetric matrices, this work extends these approaches to unitary matrices with the help of the Cayley transform. The aim is to develop and analyze efficient techniques for approximating spectral functionals in this broader context.

Contents

Introduction

We begin this thesis by examining unitary matrices and their fundamental properties. Then we revisit the Cayley transform to finally close in on the spectral density. Upon this we should be well-equipped to proceed with the investigation afterwards.

1.1 Unitary Matrices

We first recall two definitions for important real matrices that we then extend to complex matrices. The index ^{T} marks the transpose of a matrix. As common in literature, I_n denotes the identity matrix of size n.

Definition 1 (Symmetric matrix). Let A be a real, square matrix of size n. Then A is called *symmetric* if $A^T = A$.

The following definition is for matrices with their transpose as their inverse.

Definition 2 (Orthogonal matrix). Let A be a real, square matrix of size n. Then A is called *orthogonal* if $A^T \cdot A = I_n$.

The complex equivalent of a real symmetric matrix is a *Hermitian matrix*. Note that A^* is *conjugate transpose* of the matrix A with all of its entries complex conjugated and transposed.

Definition 3 (Hermitian matrix). Let A be a complex square matrix of size n. Then A is called *Hermitian* if $A^* = A$.

Throughout this thesis, A will denote a complex, square matrix of size n, unless stated otherwise. To reference a Hermitian matrix, we will use the letter H.

We now examine the eigenvalues of Hermitian matrices.

Let $H = H^*$ and $Hv = \lambda v$ for a complex vector $v \neq 0$ of size n and a scalar $\lambda \in \mathbb{C}$. Consider now the inner product v^*v.

$$\lambda v^* v = v^* \left(\lambda v \right) = v^* \left(H v \right) = \left(v^* H \right) v = \left(H^* v \right)^* v = \left(H v \right)^* v = \left(\lambda v \right)^* v = \overline{\lambda} v^* v \quad (1.1)$$

Since we have that $v \neq 0$ it follows that $v^*v \neq 0$ and therefore $\lambda = \overline{\lambda}$, that is to say λ is real. This means that all eigenvalues of Hermitian matrices are real numbers. It follows that all eigenvalues of symmetric matrices are also real numbers, since they are a special case of Hermitian matrices. Now we can define the complex equivalent of orthogonal matrices:

Definition 4 (Unitary matrix)**.** A matrix A is called *unitary* if $A^* \cdot A = I_n$.

We will oftentimes denote unitary matrices by using U as a reference. It is easy to see that orthogonal matrices are a special case of unitary matrices, since $A^T = A^*$ for all real matrices.

Consider a unitary matrix U and an eigenpair (λ, v) of U. The complex conjugate of the eigenvalue equation $Uv = \lambda v$ is

$$v^* U^* = v^* \overline{\lambda} = \overline{\lambda} v^* \tag{1.2}$$

We calculate

$$v^* v = v^* U^* U v = v^* \overline{\lambda} \lambda v = \overline{\lambda} \lambda v^* v = |\lambda|^2 v^* v$$

Similarly to above, we can divide by $v^* v$ to obtain

$$1 = |\lambda|^2 = |\lambda| \tag{1.3}$$

meaning that all eigenvalues of unitary matrices have a length of 1 and are thus situated on the unit circle. As they are a special case of unitary matrices, the same goes for orthogonal matrices. This property is crucial, as it enables the application of the Cayley transform, introduced in the following section.

There is a special group of matrices that all the matrices we have defined so far belong to.

Definition 5 (Normal matrix)**.** A matrix A is called *normal* if it commutes with its conjugate transpose, that is to say $A^* A = A A^*$.

It is easy to see that both Hermitian and unitary matrices are normal matrices and that the notion includes real symmetric and orthogonal matrices as special cases. The spectral theorem states that normal matrices can be diagonalized by a unitary matrix. That means, for any normal matrix A, there exists a unitary matrix U such that $A = U \Lambda U^*$, where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with $\lambda_1, \ldots, \lambda_n$ being the eigenvalues of A. This is essential for applying a function to a matrix, which we will get to in the next section.

1.2   Cayley Transform

Before giving the central definition of this section, we will first enxtend the concept of functions to normal matrices, allowing us to map one matrix to another.

Definition 6 (Matrix function on normal matrices)**.** Let A be a normal matrix and let $f : \mathbb{C} \to \mathbb{C}$ be a function that is defined on the spectrum of A, $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}$. Then the *matrix function* $f(A)$ is defined as

$$f(A) := U f(\Lambda) U^* = U \, \mathrm{diag}(f(\lambda_1), \ldots, f(\lambda_n)) U^*$$

where U is the matrix of eigenvectors of A and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix of eigenvalues.

Now we can define the *Cayley transform*, which is a specific matrix function, that establishes a correspondence between Hermitian and unitary matrices, allowing spectral properties to be translated between these two important classes.

For a complex number $z \in \mathbb{C}$ with $z \neq -i$, the Cayley transform is defined as

$$\varphi(z) = \frac{i - z}{i + z}.$$

This function maps the real line to the unit circle in the complex plane.
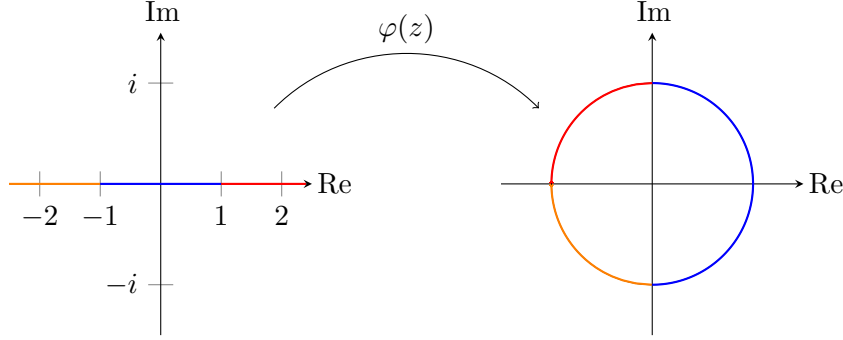
For matrices, the Cayley transform maps a Hermitian matrix H (with $i + H$ invertible) to a unitary matrix U via

$$U = (i - H)(i + H)^{-1}.$$

The condition of $i + H$ being invertible is the same as requiring that H does not have $-i$ as an eigenvalue. If (and only if) that were the case, then we had $Hv = -iv$ for some eigenvector $v \neq \mathbf{0}$, and therefore

$$(i + H)v = iv + Hv = iv + (-iv) = 0.$$

Since Hermitian matrices have only real eigenvalues as discussed above, $-i$ can never be an eigenvalue. Conversely, given a unitary matrix U (with $U \neq -I_n$), the inverse Cayley transform yields a Hermitian matrix:

$$H = i(I_n - U)(I_n + U)^{-1}.$$

This will be relevant later, as we can then use φ to transform unitary matrices into symmetric ones and vice versa. In particular, it enables the transfer of spectral density results, which will be explored in the following sections.

1.3 Spectral Density

To get to the notion of the spectral density, we will first need some more basic definitions to build on.

Definition 7 (Linear functional)**.** Let V be a vectorspace over $\mathbb{C}$. A *linear functional* T is a linear map $T : V \to \mathbb{C}$.

A simple example for such a functional is evaluation at a point:

$$T : \mathcal{C}^\infty(\mathbb{C}) \to \mathbb{C}, \qquad f \mapsto f(0) \tag{1.4}$$

where $\mathcal{C}^\infty(\mathbb{C})$ denotes the space of infinitely differentiable (smooth) functions over $\mathbb{C}$.

Another example is integration against a fixed complex-valued function g:

$$T_g : \mathcal{C}^\infty(\mathbb{C}) \to \mathbb{C}, \qquad f \mapsto \int_{\mathbb{C}} g \cdot f \, \mathrm{d}x \tag{1.5}$$

Now, let $\Omega \subset \mathbb{C}^n$ be open and non-empty. A *test function* is a smooth function with compact support defined on Ω. The space of all test functions on Ω is usually denoted by $\mathcal{E}$. *Distributions* generalize the concept of functions and allow for differentiation in cases where classical differentiation is not possible.

Definition 8 (Distribution). Let $\Omega \subset \mathbb{C}^n$ be open, and let $\mathcal{E}$ be the space of complex-valued test functions over Ω. A *distribution* T is a continuous linear functional $T : \mathcal{E} \to \mathbb{C}$. That is, for all $g_1, g_2 \in \mathcal{E}$ and $\lambda \in \mathbb{C}$,

$$T(g_1 + \lambda g_2) = T(g_1) + \lambda T(g_2)$$

and for any sequence $\{g_n\}_{n \in \mathbb{N}}$ in $\mathcal{E}$ with $g_n \to g$,

$$T(g_n) \to T(g).$$

An important example of a distribution is the *Dirac delta distribution*, which is defined as follows:

Definition 9 (Dirac delta distribution). Let $\mathcal{E} = \mathcal{C}^\infty(\Omega)$ with $0 \in \Omega \subset \mathbb{R}^n$. Then

$$\delta : \mathcal{E} \to \mathbb{R}, \quad f \mapsto f(0) \quad \text{with} \quad \delta(f) = \langle \delta, f \rangle = f(0)$$

this distribution is often mistakenly referred to as a function, although it is not a function in the classical sense.
The Dirac delta is characterized by the following property:

$$\int_{-\infty}^{\infty} f(x)\delta(x-a) \, \mathrm{d}x = \int_{-\infty}^{\infty} f(x)\delta(a-x) \, \mathrm{d}x = f(a) \implies \int_{-\infty}^{\infty} \delta(x-a) \, \mathrm{d}x = 1.$$

This means that the Dirac delta distribution is zero everywhere except at the point a, where it is infinitely high, such that the integral over it equals 1. We now have all the tools we need to define the central concept of this thesis:

Definition 10 (Spectral density). Let H be hermitian and sparse. The *spectral density* is then defined as

$$\phi(t) = \frac{1}{n} \sum_{j=1}^{n} \delta(t - \lambda_j)$$

where δ is the Dirac delta distribution and λ_j are the eigenvalues of H in non-descending order.

The number of eigenvalues in an interval $[a, b]$ can then be counted in the following manner:

$$\nu_{[a,b]} = \int\limits_a^b \sum_j \delta(t - \lambda_j) \, \mathrm{d}t \equiv \int\limits_a^b n\phi(t) \, \mathrm{d}t \tag{1.6}$$

Random matrix theory tells us that the eigenvalues of random unitary matrices are distributed uniformly on the unit circle. Thus, we would expect the spectral density to be uniformly distributed on the unit circle as well. This is important as we think of the choice of random unitary matrices.

If we generate a random complex square matrix A, there are multiple ways to obtain a unitary matrix U. We will now compare the svd with the qr decomposition.

Before we can proceed to the motice of this thesis, we will need to define the following space:

Definition 11 (Schwartz space over $\mathbb{R}$). The *Schwartz space* over $\mathbb{R}$ consists of all smooth functions f that decay rapidly to zero as $|x|$ approaches infinity [4]. Formally,

$$\mathcal{S}(\mathbb{R}) := \left\{ f \in \mathcal{C}^\infty(\mathbb{R}) \mid \forall p, k \in \mathbb{N}_0 : \sup_{x \in \mathbb{R}} \left| x^p f^{(k)}(x) \right| < \infty \right\}$$

Motivation

Calculating the spectral density of a matrix is straightforward when its eigenvalues are already known. However, this is rarely the case in practice, and computing the eigenvalues of very large matrices is both time-consuming and computationally expensive. At the same time, the density of states (DOS), which serves as a probability density for the distribution of eigenvalues, is of great interest in many fields. Therefore, there is a need for efficient methods to approximate the spectral density at low computational cost.

The challenge arises because $\phi(t)$, the delta distribution, is not a conventional *function* that can be evaluated pointwise.

One intuitive approach is to select an interval $I \in \mathbb{R}$ such that the spectrum of A, $\sigma(A)$, is contained within I. Next, choose k points t_i in I so that the interval is divided into subintervals:

$$\{t_i\}_{i=1}^k \subset I \quad \text{with} \quad \bigcup_{i=1}^{k-1}[t_i, t_{i+1}] = I$$

Count the number of eigenvalues in each subinterval. Then, calculate the average value of $\phi(t)$ in each interval using $\nu_{[a,b]}$ from equation 1.6. The result is a histogram, which, as the subintervals become smaller (i.e., as k increases and $(t_{i+1} - t_i) \longrightarrow 0$), approaches the true spectral density.

To count the eigenvalues in the intervals, one can use methods such as Sylvester's law of inertia. The details of this method are beyond the scope of this work; it would require computing a decomposition of $A - t_i I = LDL^T$ for all t_i [3]. Instead, we prefer a method in which A is multiplied by vectors, which is more efficient in higher dimensions.

Kernel Polynomial Method

3.1   Overview

There exists a whole class of methods, all of which called the kernel polynomial method, or KPM for short. They are powerful tools for approximating the spectral density of matrices. We will focus on the main approach in the following.

As the name suggests, the KPM is a polynomial extension of the spectral density. The coefficients of the polynomials are derived from the method of moments, in order to obtain an estimator function as in statistics. The method is based on a corollary of the following theorem:

Theorem 1. Let $A = A^T \in \mathbb{R}^{n \times n}$ with spectral decomposition

$$A = U \Lambda U^T \quad \text{where} \quad U U^T = I_n \text{ and } \Lambda = \text{diag}(\lambda_1, ..., \lambda_n)$$

Also, let $\beta, v \in \mathbb{R}^n$ with $v = U\beta$.

If $v_i \sim_{\text{i.i.d.}} \mathcal{N}(0,1)$ for the components $\{v_i\}_{i=1}^n$ of v, that is to say

$$\mathbb{E}[v] = 0 \text{ and } \mathbb{E}[vv^T] = I_n, \tag{3.1}$$

then

$$\mathbb{E}[\beta] = 0 \text{ and } E[\beta\beta^T] = I_n$$

Proof of Theorem 1. It holds that

$$\mathbb{E}[v] = \mathbb{E}[U\beta] = U\mathbb{E}[\beta] = 0 \implies \mathbb{E}[\beta] = 0$$

Furthermore it holds that

$$\mathbb{E}[vv^T] = \mathbb{E}[U\beta\beta^T U^T] = U\mathbb{E}[\beta\beta^T]U^T = U\mathbb{E}[\beta\beta^T]U^T = I_n$$

Since U is orthogonal, we can multiply both sides with U^T and U:

$$U^T \mathbb{E}[vv^T]U = \mathbb{E}[\beta\beta^T] = U^T I_n U = I_n$$

Thus, we have shown that $\mathbb{E}[\beta\beta^T] = I_n$. $\qquad\qquad\square$

This theorem has a nice corollary when investigating a matrix function $f(A)$. In that case, we have

$$\mathbb{E}\left[v^T f(A)v\right] = \mathbb{E}\left[(U\beta)^T f(U\Lambda U^T)(U\beta)\right] = \mathbb{E}\left[\beta^T U^T U f(\Lambda) U^T U\beta\right]$$
$$= \mathbb{E}\left[\beta^T f(\Lambda)\beta\right]$$
$$= \mathbb{E}\left[\sum_{j=1}^{n} \beta_j^2 f(\lambda_j)\right]$$
$$= \sum_{j=1}^{n} f(\lambda_j)\mathbb{E}\left[\beta_j^2\right]$$
$$= \sum_{j=1}^{n} f(\lambda_j)$$

or, put simpler

$$\mathbb{E}\left[v^T f(A)v\right] = \mathrm{Tr}(f(A)). \tag{3.2}$$

3.2 Polynomial Extension with Tschebyschev polynomials

We assume Tschebyshev polynomials to be a good fit for the polynomial extension of the Dirac delta distribution because of many of their properties. Using the trigonometrical functions, they can be expressed as follows:

$$T_k(t) = \begin{cases} \cos(k\arccos(t)) & \text{for } k \in [-1,1] \\ \cosh(k\,\mathrm{arcosh}(t)) & \text{for } k > 1 \\ (-1)^k \cosh(k\,\mathrm{arcosh}(-t)) & \text{for } k < -1 \end{cases}$$

We will only use the formula $T_k(t) = \cos(k\arccos(t))$. This means to only consider matrices, which have eigenvalues within the intervall $[-1,1]$. In the case that this condition should not be fulfilled, the eigenvalues can be transformed accordingly. For this, let λ_{lb} and λ_{ub} be the lower and upper bound for the eigenvalues of A, respectively. Define

$$c := \frac{\lambda_{lb} + \lambda_{ub}}{2} \quad \text{and} \quad d := \frac{\lambda_{ub} - \lambda_{lb}}{2}$$

Then, the matrix $B = \frac{A - c*I_n}{d}$ has eigenvalues in the interval $[-1,1]$. A visualization of this is linked in the appendix.

Tschebyschev polynomials can furthermore be calculated using the recursion formula

$$T_{k+1}(t) = 2t T_k(t) - T_{k-1}(t),$$

where the starting conditions are given by $T_0(t) = 1$ and $T_1(t) = x$.

Additionally, note that the resultat in equation 3.2 says, that

$$\mathbb{E}\left[v^T T_k(A)v\right] = \sum_{j=1}^{n} T_k(\lambda_j) = \mathrm{Tr}(T_k(A)) \tag{3.3}$$

holds. This is central as we proceed.

Now, let

$$h(x) = \frac{1}{\sqrt{1-t^2}} \tag{3.4}$$

be a weight function. Another property of Tschebyschev polynomials is that they are *orthogonal* in terms of the scalar product weighted with h:

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

This means that

$$\int_{-1}^{1} \frac{1}{\sqrt{1-t^2}} \cdot T_k(t) \cdot T_l(t) \, \mathrm{d}t = \begin{cases} 0 & \text{für } k \neq l \\ \pi & \text{für } k = l = 0 \\ \frac{\pi}{2} & \text{für } k = l \neq 0 \end{cases}$$

3.3 Approximating the spectral density

Now multiply the spectral density with the inverse of the weight function 3.4:

$$\hat{\phi}(t) = \sqrt{1-t^2}\phi(t) = \sqrt{1-t^2} \times \frac{1}{n} \sum_{j=1}^{n} \delta(t - \lambda_j)$$

Sei nun $g \in \mathcal{S}(\mathbb{R})$, the Schwartz space defined in definition 11, and $\mu_k \in \mathbb{R}$ coefficients to be determined such that the following equation holds:

$$\int_{-1}^{1} \hat{\phi}(t)g(t) \, \mathrm{d}t = \int_{-1}^{1} \sum_{k=0}^{\infty} \mu_k T_k(t)g(t) \, \mathrm{d}t \tag{3.5}$$

Is this true for arbitrary $g \in \mathcal{S}(\mathbb{R})$, we can simplify our equation 3.5 to

$$\hat{\phi}(t) = \sum_{k=0}^{\infty} \mu_k T_k(t) \tag{3.6}$$

Now utilize the orthogonality of the Tschebyschev polynomials, to calculate a specific coeffitient μ_k:

$$\sum_{l=0}^{\infty} \mu_l T_l(t) = \hat{\phi}(t)$$

$$\implies \left(\sum_{l=0}^{\infty} \mu_l T_l(t) \right) \cdot T_k(t) = \hat{\phi}(t) \cdot T_k(t)$$

$$\implies \int_{-1}^{1} \frac{1}{\sqrt{1-t^2}} \cdot \left(\sum_{l=0}^{\infty} \mu_l T_l(t) \right) \cdot T_k(t) \, \mathrm{d}t = \int_{-1}^{1} \frac{1}{\sqrt{1-t^2}} \cdot \hat{\phi}(t) \cdot T_k(t) \, \mathrm{d}t$$

$$\implies \mu_k \cdot \frac{\pi}{2 - \delta_{k0}} = \int_{-1}^{1} \frac{1}{\sqrt{1-t^2}} \cdot \sqrt{1-t^2} \cdot \phi(t) \cdot T_k(t) \, \mathrm{d}t$$

$$\implies \mu_k = \frac{2 - \delta_{k0}}{\pi} \cdot \int_{-1}^{1} \phi(t) \cdot T_k(t) \, \mathrm{d}t$$

By applying the Dirac delta distribution we obtain:

$$\mu_k = \frac{2 - \delta_{k0}}{\pi} \cdot \int_{-1}^{1} \phi(t) \cdot T_k(t) \, \mathrm{d}t = \frac{2 - \delta_{k0}}{\pi} \cdot \int_{-1}^{1} \frac{1}{n} \sum_{j=1}^{n} \delta(t - \lambda_j) \cdot T_k(t) \, \mathrm{d}t$$

$$= \frac{2 - \delta_{k0}}{n\pi} \sum_{j=1}^{n} T_k(\lambda_j)$$

$$= \frac{2 - \delta_{k0}}{n\pi} \operatorname{Tr}(T_k(A))$$

Now let $n_{vec} \in \mathbb{R}$ and consider vectors $v_0^{(1)}, v_0^{(2)}, \ldots, v_0^{(n_{vec})}$, which are chosen randomly from the standard normal distribution, that is to say $\mathbb{E}[v_0^{(k)}] = 0$ und $\mathbb{E}\left[v_0^{(k)} \left(v_0^{(k)}\right)^T\right] = I_n$. It follows from equation 3.3 that

$$\zeta_k = \frac{1}{n_{vec}} \sum_{l=1}^{n_{vec}} \left(v_0^{(l)}\right)^T T_k(A) v_0^{(l)}$$

is a good estimator for $\operatorname{Tr}(T_k(A))$ and therefore

$$\mu_k \approx \frac{2 - \delta_{k0}}{n\pi} \zeta_k$$

In order to determine the ζ_k, let $v_0 \equiv v_0^{(l)}$ Using the recursion formula for Tschebyschev polynomials, we can calculate

$$T_{k+1}(A)v_0 = 2A T_k(A)v_0 - T_{k-1}(A)v_0$$

For $v_k \equiv T_k(A)v_0$ it also holds that

$$v_{k+1} = 2A v_k - v_{k-1}$$

With this we are fully equpped for the final calculation and the goal of the KPM is reached: Instead of having to multiply matrices with other matrices, it now suffices to multiply matrices with vectors. Now we can approximate $\phi(t)$ as closely as we like. As aforementioned, it is not always desirable to have an infinitely exact approximation. Since it holds that

$$\lim_{k \to \infty} \mu_k \to 0$$

and we are only interested in $T_k(t)$ with $k \leq M$
Therefore we estimate ϕ with

$$\tilde{\phi}_M(t) = \frac{1}{\sqrt{1 - t^2}} \sum_{k=0}^{M} \mu_k T_k(t) \tag{3.7}$$

The following pseudo code is based on [2, p. 10] and summarizes the steps described above. The implementation is done in Python, and linked in the appendix.

Algorithm 1 The Kernel Polynomial Method

Require: $A = A^T \in \mathbb{R}^{n \times n}$ with eigenvalues in the intervall $[-1, 1]$
Ensure: Estimated spectral density $\{\tilde{\phi}_M(t_i)\}$

    for $k = 0 : M$ **do**
        $\zeta_k \leftarrow 0$
    end for
5: **for** $l = 1 : n_{\text{vec}}$ **do**
        Choose a random new vector $v_0^{(l)}$; $\triangleright\ v_{0_i}^{(l)} \sim$ i.i.d. $\mathcal{N}(0,1)$
        for $k = 0 : M$ **do**
            Calculate $\zeta_k \leftarrow \zeta_k + \left(v_0^{(l)}\right)^T v_k(l)$;
            if $k = 0$ **then**
10:             $v_1^{(l)} \leftarrow A v_0^{(l)}$
            else
               $v_{k+1}^{(l)} \leftarrow 2 A v_k^{(l)} - v_{k-1}^{(l)}$ $\triangleright$ Three term recursion
            end if
        end for
15: **end for**
    for $k = 0 : M$ **do**
        $\zeta_k \leftarrow \frac{\zeta_k}{n_{\text{vec}}}$
        $\mu_k \leftarrow \frac{2 - \delta_{k0}}{n\pi} \zeta_k$
    end for
20: Evaluate $\tilde{\phi}_M(t_i)$ with equation 3.7

Quality Analysis

Let $\tilde{\phi}(t)$ be a regular function, that estimates the spectral density. Since $\phi(t)$ is no function in the classical sense, but all approximations are continuous functions,

$$\left\| \phi(t) - \tilde{\phi}(t) \right\|_{L^p}$$

is not defined, to estimate the error. In the following, we will introduce two methods to circumvent this problem. The first one uses the fact that $\delta(t)$ is a distribution that can be applied to test functions from the Schwartz space $\mathcal{S}(\mathbb{R})$, while the second one regularizes δ-functions by replacing them with continuous and smooth functions. These two approaches are closely related, as we will see later.

4.1 Resolution

In practice, there is often no need to approximate all eigenvalues of A exactly. This would even be highly discontinuous, as we recall the intuitive approach with histograms. The accuracy of the approximation should be oriented towards the desired *resolution*. This means that often only the eigenvalues in a certain subinterval $[a, b] \subset \sigma(A)$ are of interest in a certain context. The size $b - a$ is then called the resolution of the estimate. This insight should be taken as a parameter in our further considerations.

4.2 Restriction on the Schwartz space

In this first method, we consider $\delta(t)$ as a distribution. Let $g \in \mathcal{C}^\infty(\mathbb{R})$ be a test function from the Schwartz space $\mathcal{S}(\mathbb{R})$ defined in Definition 11. Then, according to Definition 8, we have

$$\langle \delta(\cdot - \lambda), g \rangle = \int\limits_{-\infty}^{\infty} \delta(t - \lambda) g(t) \, \mathrm{d}t = g(\lambda)$$

and for all $p, k \in \mathbb{N}_0$

$$\sup_{t \in \mathbb{R}} |t^p g^{(k)}(t)| < \infty$$

Then the error can be measured as follows:

$$\epsilon_1 = \sup_{g \in \mathcal{S}(\mathbb{R})} \left| \langle \phi, g \rangle - \langle \tilde{\phi}, g \rangle \right|$$

To transfer the concept of resolution to this error, we first consider $\nu_{[a,b]}$ from equation 1.6 and define accordingly

$$\tilde{\nu}_{[a,b]} = \int\limits_a^b n\tilde{\phi}(t) \, \mathrm{d}t$$

with $\tilde{\phi}(t) \in \mathcal{C}^\infty(\mathbb{R})$. Assuming $n = 1$ and thus $\phi(t) = \delta(t)$, an infinite resolution would mean that for arbitrarily small intervals $[a,b]$ the term $\left|\nu_{[a,b]} - \tilde{\nu}_{[a,b]}\right|$ also becomes arbitrarily small. Let $a = -\varepsilon, b = \varepsilon$. From the definition of the δ-function it follows that

$$\lim_{\varepsilon \to 0+} \nu_{[-\varepsilon,\varepsilon]} = 1$$

while for smooth functions $\tilde{\phi}$ it is clear that

$$\lim_{\varepsilon \to 0+} \tilde{\nu}_{[-\varepsilon,\varepsilon]} = 0$$

This shows that no smooth function converges to the spectral density under continuous increase of the resolution. A carefully balanced approximation would therefore be as precise as a constant one. However, as noted earlier, a finite resolution is often sufficient. We can therefore restrict the Schwartz space $\mathcal{S}(\mathbb{R})$. For example, one could only consider Gaussian distributions of the form

$$g_\sigma(t) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} e^{-\frac{t^2}{2\sigma^2}}$$

and restrict $\mathcal{S}(\mathbb{R})$ to the subspace

$$\mathcal{S}(\mathbb{R})(\sigma; [\lambda_{us}, \lambda_{os}]) = \{g \mid g(t) \equiv g_\sigma(t - \lambda), \lambda \in [\lambda_{us}, \lambda_{os}]\}$$

Here, λ_{us} and λ_{os} are the upper and lower bounds of the eigenvalues of A as in the previous section, while the parameter σ represents the *target resolution*. We can now use the following metric for quality assessment:

$$E\left[\tilde{\phi}; \mathcal{S}(\mathbb{R})\left(\sigma; [\lambda_{lb}, \lambda_{ub}]\right)\right] = \sup_{g \in \mathcal{S}(\mathbb{R})(\sigma; [\lambda_{lb}, \lambda_{ub}])} \left|\langle\phi, g\rangle - \langle\tilde{\phi}, g\rangle\right| \tag{4.1}$$

4.3   Regularizing the Dirac delta distribution

In the second method, we regularize the δ-function. This means that we replace the δ-function with a smooth function $\phi(t)$ that approximates the spectral density. This is done by smoothing the δ-function. For example, we can use the Gaussian distribution with standard deviation σ:

$$\phi_\sigma(t) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} e^{-\frac{t^2}{2\sigma^2}}$$

This function is smooth and well-defined for all $t \in \mathbb{R}$. It is also normalized, meaning that

$$\int\limits_{-\infty}^{\infty} \phi_\sigma(t) \, \mathrm{d}t = 1$$

This means that $\phi_\sigma(t)$ can be interpreted as a probability density function. The parameter σ controls the width of the Gaussian distribution and thus the resolution of the approximation. This means that the smaller σ is, the more precise the approximation is, but also the more sensitive it is to noise in the data. This is a common approach in numerical analysis to approximate distributions with smooth functions. The function $\phi_\sigma(t)$ is then well-defined and can be used to calculate the error for $p = 1, 2$ and ∞:

$$\epsilon_2 = \left|\left|\phi_\sigma(t) - \tilde{\phi}(t)\right|\right|_p$$

Let

$$\phi_\sigma(t) = \langle \phi(\cdot), g_\sigma(\cdot - t) \rangle = \sum_{j=1}^{n} g_\sigma(t - \lambda_j)$$

This is the convolution of the spectral density with the Gaussian function g_σ. In similar manner, let

$$\tilde{\phi}_\sigma(t) = \langle \tilde{\phi}(\cdot), g_\sigma(\cdot - t) \rangle$$

Then

$$E\left[\tilde{\phi}; \mathcal{S}(\mathbb{R}) \left(\sigma; [\lambda_{lb}, \lambda_{ub}]\right)\right] = \sup_{g \in \mathcal{S}(\mathbb{R})(\sigma;[\lambda_{lb}, \lambda_{ub}])} \left|\phi_\sigma(t) - \tilde{\phi}_\sigma(t)\right|$$

is the L^∞ error between two well-defined functions. Just like in the other method, σ controls the resolution: the larger the σ, the smoother but also less accurate the approximation. For smaller values, one obtains coarser functions, where small peaks can be detected exactly where the eigenvalues are located. The following graphic is taken from [2, p. 6] and illustrates this effect for four different values of σ: Sei also

$$\phi_\sigma(t) = \langle \phi(\cdot), g_\sigma(\cdot - t) \rangle = \sum_{j=1}^{n} g_\sigma(t - \lambda_j)$$

Dies ist dann nicht anderes als die "Weichzeichnung" der Spektraldichte durch Gauß-Funktionen der Breite σ.

15
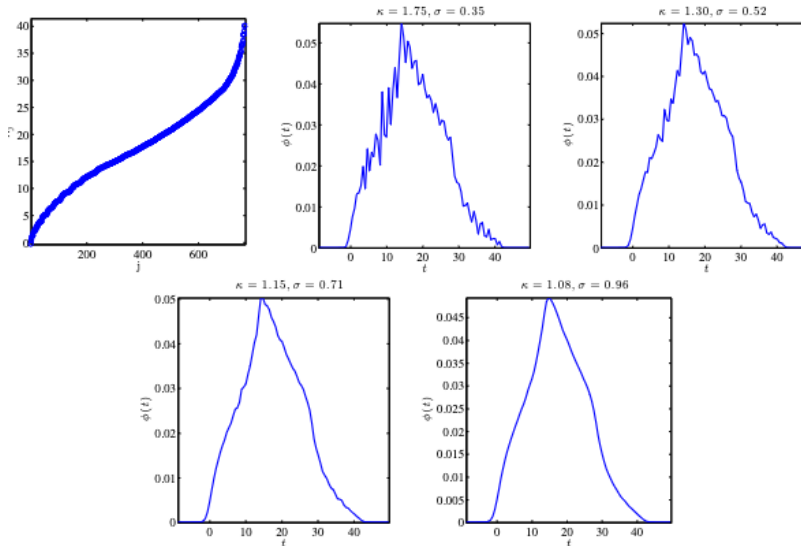
4.4 The condition of non-negativity

As a probability distribution, the spectral density is non-negative, meaning

$$\forall g \in \mathcal{S}(\mathbb{R}), g \geq 0 : \langle \phi, g \rangle \geq 0$$

Some numerical approximations break this property, including the Kernel Polynomial Method. This leads to large errors and must be considered carefully. Methods to mitigate these negative effects would exceed the scope of this work.

Bibliography

[1] CHRISTIAN BÄR, *Lineare Algebra und analytische Geometrie*, Springer Spektrum 2018

[2] LIN LIN, YOUSSEF SAAD AND CHAO YANG, *Approximating Spectral Densities of Large Matrices*, `https://arxiv.org/pdf/1308.5467v2`
5. Oktober 2014

[3] G. H. GOLUB UND C. F. VAN LOAN, *Matrix Computations, 4th Edition*, Johns Hopkins University Press, Baltimore, MD, 3rd ed., 2013

[4] R. D. RICHTMYER, *Principles of advanced mathematical physics*, vol 1, Springer Verlag, New York, 1981