University of Potsdam

Bachelor Thesis

Investigation of Functionals of the Eigenvalues of Unitary Matrices

Author:
Carina Seidel

Supervisor:
Dr. Thomas Mach

July 9, 2025

Abstract

The eigenvalues of large matrices play a crucial role in a wide range of applications. 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). While established methods exist for real symmetric matrices, this work extends these approaches to unitary matrices by employing the Cayley transform. The aim is to develop and analyze efficient techniques for approximating spectral functionals in this broader context.

Contents

1 Introduction **2**
1.1 Unitary Matrices 2
1.2 Cayley Transform 3
1.3 Spectral Density 4

2 Motivation **6**
2.1 Motivation 6

Introduction

To commence this investigation we begin 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 the literature, I_n denotes the identity matrix of size n.

Definition 1 (Orthogonal matrix)**.** Let A be a real, square matrix of size n. Then A is called *orthogonal* if $A^T \cdot A = I_n$.

For completeness' sake, we also define symmetric matrices.

Definition 2 (Symmetric matrix)**.** Let A be a real, square matrix of size n. Then A is called *symmetric* if $A^T = A$.

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.

We examine the eigenvalues of Hermitian matrices.

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

$$\lambda v^* v = v^* (\lambda v) = v^* (Av) = (A^* v)^* v = (Av)^* v = (\lambda v)^* v = \overline{\lambda} v^* v \tag{1.1}$$

Since we have that $v \neq \mathbf{0}$ 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.

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

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

We calculate

$$v^* v = v^* U^* U v = v^* \lambda^* \lambda v = \lambda^* \lambda v^* v = |\lambda|^2 v^* v$$

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

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

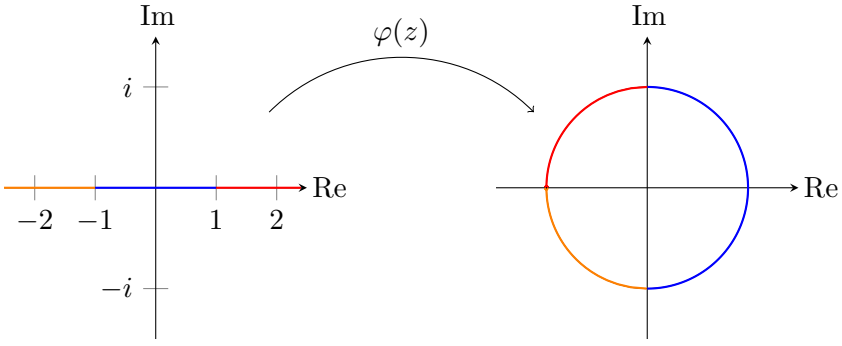
meaning that all eigenvalues of unitary matrices have a length of 1 and are thus situated on the unit circle. For orthogonal matrices, this means their eigenvalues are either 1 or -1. This property is crucial, as it enables the application of the Cayley transform, introduced in the following section.

1.2 Cayley Transform

The Cayley transform is given by the simple function

$$\varphi(z) = (i - z)(i + z)^{-1}$$

This function maps the real line to the unit circle, and more specifically, the interval $[-1, 1]$ to the semi circle $\{z = e^{i\theta} : \theta \in [-\pi/2, \pi/2]\}$ as shown below



This will be relevant later, as we can then use φ to transform unitary matrices into symmetric once and vice versa. Now, let's look how the spectral density is defined in the next section.

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 5 (linear functional)**.** Let V be a vectorspace over a field $\mathbb{K}$. A *linear functional* T is a linear function $T : V \to \mathbb{K}$. The space over all linear functionals $V \mapsto \mathbb{K}$ is called the *dual space* V'.

A simple example for such a functional would be

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

A more special case is the integral.

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

This definition leads to the concept of a distribution which was introduced to get a method to differentiate where differentiation in the classical sense is not possible.

Definition 6 (distribution)**.** Let $\emptyset \neq \Omega \subset \mathbb{R}^n$ be open. Let $\mathcal{E}$ be the space of *test functions* over Ω. A *distribution* T is a function $T : \mathcal{E} \to \mathbb{C}$ where for all $g, g_1, g_2, \{g_n\}_{n\in\mathbb{N}} \in \mathcal{E}$ with $\lim\limits_{n\to\infty} g_n \to g$ it holds:

$$T(g_1 + \lambda g_2) = T(g_1) + \lambda T(g_2) \quad \text{und} \quad \lim_{n\to\infty} T(g_n) \to T(g)$$

So, put shortly: A distribution T is a continuous and linear functional on $\mathcal{E}$. Now the path is clear to define the so called Dirac delta distribution

Definition 7 (Dirac delta function)**.** Let $\mathcal{E} = \mathcal{C}^\infty(\Omega)$ with $0 \in \Omega \subset \mathbb{R}^n$. Then

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

An important feature of this definition is:

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

This distribution is often misleadingly labled as a function inspite of being decidedly not that.
We can now finally define the spectral density.

Definition 8 (spectral density)**.** Let $A \in \mathbb{R}^{n\times n}$, $A^T = A$ and A sparce. Then, the spectral density is defined as

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

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

The number of eigenvalues in an intervall $[a, b]$ can then be expressed as follows:

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

Definition 9 (Schwartz-space over $\mathbb{R}$)**.** The Schwartz-space over $\mathbb{R}$ contains all smooth functions f, which fall fast enough against 0, when $|x|$ gets closer to ∞. [4] In Formeln

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

Im Weiteren werde ich das Symbol $\mathcal{S}$ als Abkürzung für $\mathcal{S}(\mathbb{R})$ benutzen, da sich diese Arbeit allein mit dem reellen Kontext befasst.

Motivation

2.1 Motivation

Calculating the spectral density of a matrix is trivial, when its eigenvalues are already known. However, this is mostly not the case and calculating eigenvalues of very large matrices is time- and energy intensive. At the same time, the DOS as a kind of probability density over the distribution of eigenvalues is of great interest in many fields. Thus, there is a need for methods which approximate the spectral density at low cost. The problem with this is that $\phi(t)$ the Delta distribution is not a *function* as we know it, that can be evaluated at each point.
A more intuitive idea would be to choose an intervall $I \in \mathbb{R}$ such that the spectrum of A, $\sigma(A)$, is a subset of I. Now choose k points t_i in I, such that the intervall is divided in sub intervals:

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

Now count the eigenvalues in every sub interval. Then calculate the average value of $\phi(t)$ in every intervall with $\nu_{[a,b]}$ from equation 1.6. The results is histograms, which with increasingly smaller subintervalls, that is to say bigger k and $(t_{i+1} - t_i) \longrightarrow 0$, approach the spectral density.
To count the eigenvalues in the intervals, there is means like for example the Sylvestreschen Trägheitssatz. The details of this method are not part of this work, it would be necessary to calculate a decomposition of $A - t_i I = LDL^T$ for all t_i [3]. We prefer a method in which A is multiplied with vectors, which is in bigger dimensions.
For simplicity we are going to assume, that A is szmmetric and real. The extension to hermetian matrices is simple in comparison.

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