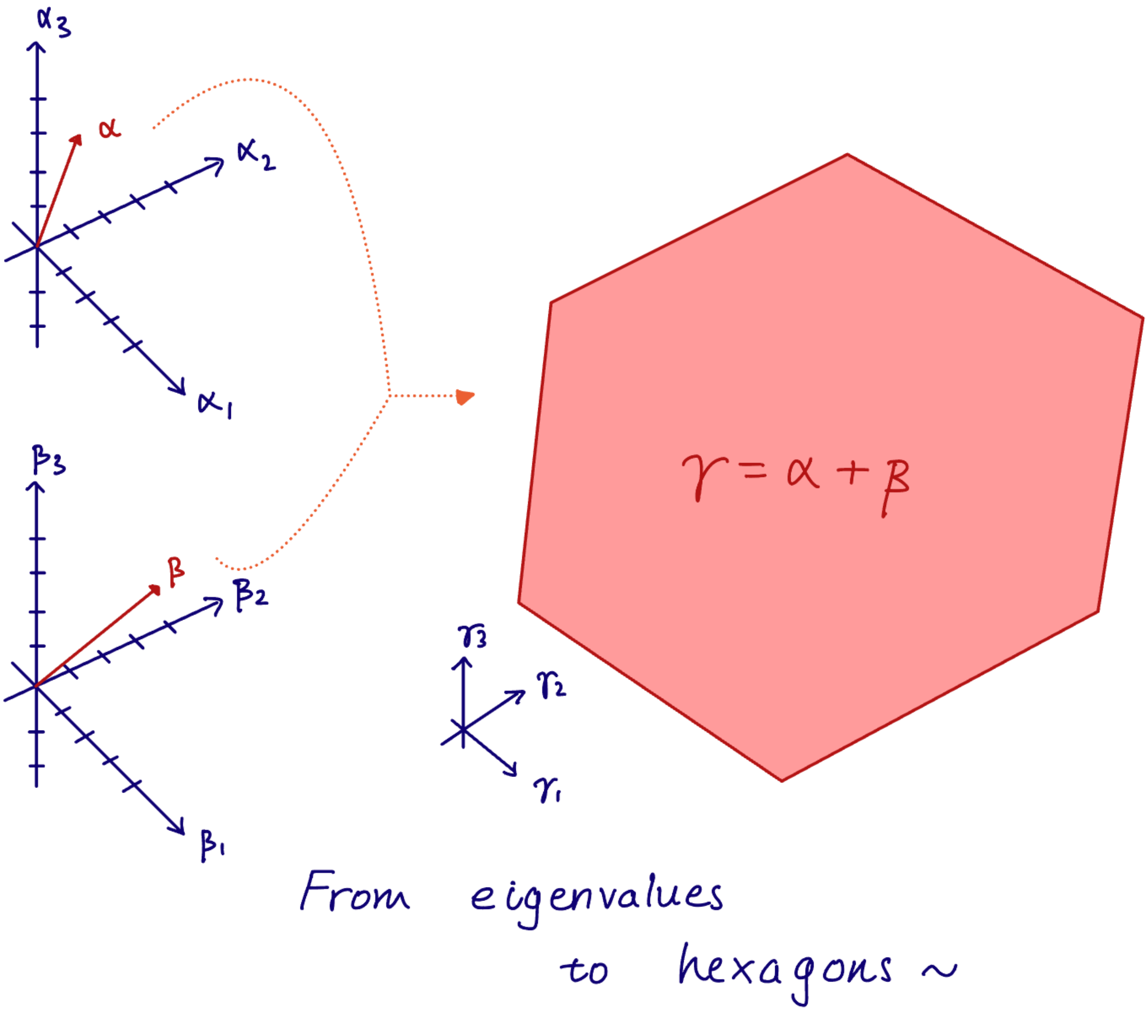
A Probabilistic Interpretation of the Horn Problem

Sean Yang



A University of Auckland Summer Project

Supervised by Professor Pedram Hekmati

08 March 2025

UNIVERSITY OF AUCKLAND
Waipapa Taumata Rau
NEW ZEALAND

Abstract

This summer project will be the breakdown of the Horn problem and a probabilistic re-interpretation of the Horn problem. We will be developing theory from an undergraduate perspective to tackle the results developed during the solving of the Horn problem and its probabilistic interpretation. The first part involves proving matrix theory results, and a basic Horn problem inequality; Weyl's inequalities by Hermann Weyl, to specific examples for $n = 2, 3$ case of the Horn problem. The second part involves the development of Radon and Haar measures where we prove the Riesz-Markov-Kakutani representation theorem, and Haar's theorem, to giving a construction of the unitary Haar measure, to a brief explanation of J. Faraut's work on the probabilistic Horn problem.

I thank my supervisor Pedram Hekmati for his continued support throughout my summer project.

I thank my mentor Dean Miller for his guidance on the mathematics of my summer project.

Contents

1 | Introduction

It is common stance that eigenvalues are one of the most important attribute regarding matrices. Thus a natural question to ask is how eigenvalues are preserved through matrix operations?

The Horn problem, conjectured by Alfred Horn, answers a part of this question, where it asks how eigenvalues are preserved through addition of Hermitian matrices. To put it briefly, the Horn problems answers the following: given Hermitian matrices A and B, one can be said about the eigenvalues of $A + B$?

Furthermore, we will be discussing the probability analog of the Horn problem, deemed as the probabilistic Horn problem. Where we further investigate the probability distribution of the eigenvalues of $A + B$ if A and B came from a family of Hermitian matrices with fixed eigenvalues.

In Chapter 3, we will be developing matrix theory results for the development of further chapters. While Chapter 3.3, the development of matrix roots, will be specifically be used to give an explicit construction of the unitary Haar measure in Chapter 7.3. Which is one of the foundational tools to answer the probabilistic Horn problem.

In Chapter 4 , we will give a thorough introduction to the Horn problem and simple results regarding the restriction for the eigenvalues of $A + B$. With those results, we can answer the Horn problem for the 2-dimensional case as presented in Chapter 4.2. Finally in Chapter 4.4, we will give the full statement of the Horn problem, but will not prove it, with the main goal of understanding the solution by providing an example for 3-dimensional matrices.

The theory involved to discuss the probability distribution of the eigenvalues will require a heavy machinery from measure theory. Haar measures, which are Radon measures on locally compact groups that are left (or right) translational invariant, are of interest, especially unitary Haar measures. Thus in Chapter 5, we will be developing the theory revolving Radon measures, starting with a brief introduction of Radon measures in Chapter 5.1, and a proof of the Riesz-Markov-Kakutani Representation Theorem −inspired by Gerald B. Folland−encapsulating the essence of Radon measures. Further relevant properties will be provided in Chapter 5.2. Additional relevant measure theoretic concepts: pushforwards and supports, will also be introduced and discussed in Chapter 5.3 and 5.4 respectively.

In Chapter 6, we will give a short introduction of topological groups and provide simple results which will be used in later chapters.

In Chapter 7, we will give an introduction of Haar measures and their relevant properties. Finally, providing a proof of the Haar's theorem in Chapter 7.2, which allows us to obtain unique (up to a multiplicative constant) left Haar measures on locally compact groups, hence finalizing the setup to attack the probabilistic Horn problem. As the existence of such Haar measures in the proof of Haar's theorem uses axiom of choice, we will also give an explicit construction of the unitary Haar measure that does not utilise axiom of choice, which provides a gateway to give explicit computations of the unitary Haar measure.

As a matter of fact that the results central to the probabilistic Horn problem is too technical for the Summer project, we will be only be giving brief explanations of such results, and the process of the solution towards the probabilistic Horn problem. The solutions we will be examining are from J. Faraut, which is what essentially Chapters 8 and 9 are; a breakdown of Faraut's approach to be digestible for undergraduates.

2 | Notations

This section will list down the common notations that will be used throughout the paper. The notations should be conventional, but will still be defined here to avoid confusion. New notations that are relevant to the topic of the paper will be introduced in further sections.

1 – Algebra

The symbols m and n will mean some element in $\mathbb{N} := \{1, 2, ...\}$, this definition would persist in the Matrix Theory Chapter 3. Otherwise, their meaning should remain the same.

Let F be a field. Given a vector $v \in F^n$ and $i \leq n$ (i.e. $i \in \{1, ..., n\}$), we denote v_i to be the ith component of v, i.e. $v = (v_1, ..., v_n)$.

Symbol	**Meaning**
$F^{m \times n}$	The set of $m \times n$ matrices over field F.
$\mathbb{K}$	Either $\mathbb{R}$ or $\mathbb{C}$.
$\mathbb{R}^n_{\geq 0}$	$\{x \in \mathbb{R}^n : x_i \geq 0, \text{for all } i \leq n\}$
$\mathbb{R}^n_{>0}$	$\{x \in \mathbb{R}^n : x_i > 0, \text{for all } i \leq n\}$
$\mathcal{H}(n)$	The set of $n \times n$ Hermitian matrices (over $\mathbb{C}$).
$\mathcal{U}(n)$	The set of $n \times n$ unitary matrices (over $\mathbb{C}$).

For finite indexing, we will refer $\sum_{i=1}^{n}$ as $\sum_{i \leq n}$, and similarly for other operations that uses indexing. Given $A \in \mathbb{K}^{m \times n}$, we denote $\|A\|$ to be the usual operator norm, i.e.

$$\|A\| = \sup_{\substack{x \in \mathbb{K}^m \\ \|x\| \leq 1}} \|Ax\|.$$

We shall denote $\mathbb{R}^n_{\geq 0}$ (respectively $\mathbb{R}^n_{>0}$) to be the set of real vectors of nonnegative (respectively positive) entries. We also apply the $\downarrow$ subscript to denote that vectors are in (weakly) decreasing order, e.g.

$$(\mathbb{R}^n)_{\downarrow} := \{(x_1, ..., x_n) \in \mathbb{R}^n : x_1 \geq ... \geq x_n\}.$$

We shall denote $\langle \cdot, \cdot \rangle$ to be the usual inner product in $\mathbb{C}$ that is linear in the first entry.

Given a matrix A over F.

Symbol	**Meaning**
A^*	If $F = \mathbb{C}$: the conjugate-transpose of some matrix A.
$\sigma(A)$	The set of eigenvalues of A.
$\sigma_{\downarrow}(A)$	The vector of (need not to be unique) eigenvalues of A in decreasing order.
$\chi_A(\lambda)$	The characteristic polynomial of A with variable $\lambda \in F$.

I_n	The $n \times n$ identity matrix.

Given a $v \in F^n$, we shall denote $\operatorname{diag}(v)$ to be the $n \times n$ diagonal matrix over F with the diagonal entries corresponding to the coordinate entries in v.

Given a $\lambda \in \sigma(A)$, denote $\operatorname{alg}_A(\lambda)$ and $\operatorname{geo}_A(\lambda)$ to the algebraic and geometric multiplicities of λ respectively. The subscript may be omitted if the context is clear.

2 – Analysis

The notations used here will be relevant in Chapter 5 and beyond.

Given a topological spaces X, Y, $x \in X$, $A \subseteq X$ and $I \subseteq \mathbb{C}$ (be nonempty!).

Symbol	**Meaning**
$\overline{A}$	Closure of A
A°	Interior of A
$\mathcal{N}_X[x]$	The set of all neighbourhoods containing x. The subscript X may be omitted if the context is clear.
$\mathcal{B}(X)$	Borel σ-algebra of X.
$C(X, Y)$	The set of continuous functions from X to another topological space Y
$C(X)$	$C(X, \mathbb{C})$
$C_c(X, I)$	$\{f \in C(X, I) : \operatorname{supp}(f) \text{ is compact}\}$.
$C_c(X)$	$C_c(X, \mathbb{C})$

It follows that $C_c(X)$ is a normed space with the uniform norm:

$$\|f\|_\infty := \sup_{x \in X} |f(x)| = \max_{x \in X} |f(x)| \quad \text{for all } f \in C_c(X).$$

We adapt the convention(s):

- $\sup(\emptyset) = -\infty$ and $\inf(\emptyset) = \infty$.

Given a (nonempty) set X, a σ-algebra Σ on X, a measure μ on (X, Σ), and $p \in [1, \infty]$, we will abbreviate the usual L^p space equipped with the L^p-norm $\|\cdot\|_p$, $L^p(X, \Sigma, \mu)$, to just $L^p(\mu)$ or even L^p if the context is clear.

Given a subset A of a set X, we shall denote $1_A : X \to \{0, 1\}$ to be the indicator function. So $1_\emptyset = 0$ and $1_X = 1$.

Given a group G, whenever e is noted in the same context, it will always refer to the identity of G.

3 | Matrix Theory

In this section, we will briefly develop the linear algebra theory required for the Horn problem, and its probabilistic interpretation.

1 – Unitarily Similarity

We aim to prove that a commuting family of normal matrices are simultaneously unitarily diagonalizable, i.e. unitarily similar to diagonal matrices. This is essentially the generalization of the spectral theorem and it provides all of the diagonalization results we need in this paper.

Note that a matrix $A \in \mathbb{K}^{n\times n}$ is **normal** if $A^*A = AA^*$.

Given a matrices $A, B \in \mathbb{C}^{n\times n}$, we say A is **unitarily similar** to B if there is a $U \in \mathcal{U}(n)$ such that $U^*AU = B$. One has a nice result that every matrix is unitarily similar to an upper triangular matrix with matching eigenvalues.

Schur's Decomposition 1. *Any matrix $A \in \mathbb{C}^{n\times n}$ is unitarily similar to an upper triangular matrix. In particular, given $\lambda_1, ..., \lambda_n \in \sigma(A)$, then there is a $U \in \mathcal{U}(n)$ such that U^*AU is upper triangular with the diagonal being $\lambda_1, ..., \lambda_n$.*

PROOF. Clearly the result holds for $n = 1$, so consider induction on $n > 1$. By the fundamental theorem of algebra, A has an eigenvalue $\lambda \in \mathbb{C}$. Let $x \in \mathbb{C}^n$ be the unit eigenvector corresponding to λ, and by the Gram-Schmidt algorithm, one has a unitary matrix $U_1 = (x, u_2, \cdots, u_n)$ (where $u_i \in \mathbb{C}^n$). Now one has

$$U_1^*AU_1 = \begin{pmatrix} x^* \\ u_2^* \\ \vdots \\ u_n^* \end{pmatrix} (Ax \;\; Au_2 \; \dots \; Au_n) = \begin{pmatrix} \lambda & x^*Au_2 & \dots & x^*Au_n \\ \lambda u_2^* x & & & \\ \vdots & & A_1 & \\ \lambda u_n^* x & & & \end{pmatrix} = \begin{pmatrix} \lambda & a \\ 0 & A_1 \end{pmatrix}.$$

Now $A_1 \in \mathbb{C}^{(n-1)\times(n-1)}$, $a \in \mathbb{C}^{1\times(n-1)}$, then by inductive hypothesis, there is a $U_2 \in \mathcal{U}(n-1)$ such that $U_2^*AU_2$ is upper triangular with the main diagonal being eigenvalues of A_1. Define a unitary matrix:

$$U = U_1 \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix}.$$

So one has

$$U^*AU = \begin{pmatrix} 1 & 0 \\ 0 & U_2^* \end{pmatrix} \begin{pmatrix} \lambda & a \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix} = \begin{pmatrix} \lambda & a \\ 0 & U_2^*A_1U_2 \end{pmatrix}$$

and the rest follows by noting that $\sigma(A) = \sigma(A_1) \cup \{\lambda\}$. ■

Schur's decomposition theorem gives a nice correlation between the trace and determinant of matrices in relation to its eigenvalues.

Corollary 2. *Let $A \in \mathbb{C}^{n\times n}$. Then*

$$\operatorname{tr}(A) = \sum_{\lambda\in\sigma(A)} \lambda,$$

and

$$\det(A) = \prod_{\lambda\in\sigma(A)} \lambda.$$

PROOF. By Schur's Decomposition 3.1.1, there is a $U \in \mathcal{U}(n)$ such that UAU^* is upper triangular with eigenvalues of A on its diagonal. As tr and det are invariant under matrix similarity, then the result follows. ∎

Given $\mathcal{F} \subseteq \mathbb{K}^{n\times n}$, we say $\mathcal{F}$ is a **commuting family** if $AB = BA$ for all $A, B \in \mathcal{F}$.

We shall extend the result of Schur's decomposition theorem to a commuting family of matrices.

Lemma 3. *Let F be an algebraically closed field, and $\mathcal{F} \subseteq F^{n\times n}$ be a commuting family. Then there is an nonzero vector in F^n that is an eigenvector for all $A \in \mathcal{F}$.*

PROOF. Let $M = \{\dim(V) : V \leq F^n, V \text{ is } \mathcal{F}\text{-invariant}\}$ (by $\mathcal{F}$-invariant, we mean $Ax \in V$ for all $x \in V$ and $A \in \mathcal{F}$). Note that $M \neq \emptyset$ as F^n is $\mathcal{F}$-invariant. Let $k = \min(M)$ and $V \leq F^n$ ($\leq$ means linear subspace here) be $\mathcal{F}$-invariant such that $\dim(V) = k$. Fix a $A \in \mathcal{F}$, and so there is an eigenvalue $\lambda \in F$ with an associated eigenvector in V. Consider $W_{A,\lambda} = \{x \in V : Ax = \lambda x\} \leq V$ (which is nontrivial), and let $x \in W$ and observe that

$$A(Bx) = (AB)x = B(Ax) = \lambda(Bx),$$

hence $Bx \in W$, which holds for all $B \in \mathcal{F}$. Thus W is $\mathcal{F}$-invariant, and so $k \leq \dim(W) \leq \dim(V) = k$ shows that $\dim(W) = k$, hence $W_{A,\lambda} = V$. Thus it follows that V gives a nontrivial space of eigenvectors for all $A \in \mathcal{F}$ ∎

Theorem 4. *Let $\mathcal{F} \subseteq \mathbb{C}^{n\times n}$ be a commuting family. Then there is a $U \in \mathcal{U}(n)$ such that U^*AU is upper triangular for all $A \in \mathcal{F}$.*

PROOF. Clearly it is true for $n = 1$, so consider induction on $n > 1$.
By preceding lemma, let $v \in \mathbb{C}^n \setminus \{0\}$ be an eigenvector for all $A \in \mathcal{F}$, then following from the proof in Schur's Decomposition 3.1.1, one has a $U \in \mathcal{U}(n)$ with the first column being v such that UAU^* is of the form

$$\begin{pmatrix} \lambda_A & a_A \\ 0 & C_A \end{pmatrix} \quad \text{where } \lambda_A \in \mathbb{C}, a_A \in \mathbb{C}^{1\times(n-1)} \text{ and } C_A \in \mathbb{C}^{(n-1)\times(n-1)}$$

for all $A \in \mathcal{F}$. Now given $A, B \in \mathcal{F}$, one has

$$\begin{aligned} 0 = UABU^* - UBAU^* &= (UAU^*)(UBU^*) - (UBU^*)(UAU^*) \\ &= \begin{pmatrix} \lambda_A & a_A \\ 0 & C_A \end{pmatrix}\begin{pmatrix} \lambda_B & a_B \\ 0 & C_B \end{pmatrix} - \begin{pmatrix} \lambda_B & a_B \\ 0 & C_B \end{pmatrix}\begin{pmatrix} \lambda_A & a_A \\ 0 & C_A \end{pmatrix} \\ &= \begin{pmatrix} \lambda_A\lambda_B & \lambda_A a_B + a_A C_B \\ 0 & C_A C_B \end{pmatrix} - \begin{pmatrix} \lambda_B\lambda_A & \lambda_B a_A + a_B C_A \\ 0 & C_B C_A \end{pmatrix}, \end{aligned}$$

so $C_A C_B = C_B C_A$, hence the family $\{C_A : A \in \mathcal{F}\}$ is a commuting family of $(n-1) \times (n-1)$ matrices over $\mathbb{C}$. Hence the rest follows from induction in the same manner as the proof in Schur's Decomposition 3.1.1. ■

Finally, we show that triangular normal matrices are diagonal, which immediately follows that a commuting family of normal matrices is simultaneously unitarily diagonalizable.

Lemma 5. *Let $A \in \mathbb{C}^{n\times n}$ be a normal matrix. If A is upper triangular, then A is diagonal.*

PROOF. Clearly it is true for $n = 1$. So consider induction on $n > 1$. Now

$$A = \begin{pmatrix} a & B \\ 0 & C \end{pmatrix} \text{ where } a \in \mathbb{C}, B \in \mathbb{C}^{1\times(n-1)}, \text{and upper triangular } C \in \mathbb{C}^{(n-1)\times(n-1)}.$$

So one has

$$\begin{aligned} 0 = A^*A - AA^* &= \begin{pmatrix} \overline{a} & 0 \\ B^* & C^* \end{pmatrix}\begin{pmatrix} a & B \\ 0 & C \end{pmatrix} - \begin{pmatrix} a & B \\ 0 & C \end{pmatrix}\begin{pmatrix} \overline{a} & 0 \\ B^* & C^* \end{pmatrix} \\ &= \begin{pmatrix} |a|^2 & \overline{a}B \\ aB^* & B^*B + C^*C \end{pmatrix} - \begin{pmatrix} |a|^2 + BB^* & BC^* \\ CB^* & CC^* \end{pmatrix} \end{aligned}$$

hence it follows that $0 = BB^*$, i.e. $B = 0$, so one has $0 = C^*C - CC^*$, so C is normal. Then by induction on C, it follows that A is diagonal. ■

Corollary 6. *Let $\mathcal{F} \subseteq \mathbb{C}^{n\times n}$ be a commuting family of normal matrices. Then $\mathcal{F}$ is* ***simultaneously unitarily diagonalizable****, that is there is a $U \in \mathcal{U}(n)$, such that U^*AU is diagonal for all $A \in \mathcal{F}$.*

Note that unitary and Hermitian matrices are normal matrices, and $\mathcal{F} = \{A\}$ for any $A \in \mathbb{C}^{n\times n}$ is a commuting family, it follows that we have the spectral theorem for both unitary and Hermitian matrices.

2 – Matrix Spaces

Here we aim to make clear on the types of matrix spaces relevant to our discussion. We will also be providing brief inequalities revolving around eigenvalues in relation to the analytical structure of those matrix spaces.

Recall that the space of matrices $\mathbb{K}^{n\times n}$ can be made into a normed space, hence a Banach space as it is finite-dimensional, endowed with the operator norm.

It is clear that $\mathbb{K}^{n\times n}$ can be identified with $\mathbb{K}^{n^2}$, where the identification is any coodinate projection map from $\mathbb{K}^{n\times n}$ to $\mathbb{K}^{n^2}$, which gives a linear homeomorphism.

One of the key properties of the operator norm we will be using frequently is that it bounds eigenvalues, and attains them for unitarily diagonalizable matrices.

Lemma 1. *Let $X \in \mathbb{K}^{n\times n}$, then*

$$\max_{\lambda\in\sigma(X)} |\lambda| \leq \|X\|.$$

In particular, if X is unitarily diagonalizable, then $\|X\| = \max_{\lambda\in\sigma(X)} |\lambda|$.

PROOF. Let $\lambda \in \sigma(X)$ and $x \in \mathbb{K}^n$ be a unit eigenvector of λ. Then one has

$$|\lambda| = \|\lambda x\| = \|Xx\| \leq \|X\|.$$

If X is unitarily diagonalizable, then there is an orthonormal basis of eigenvectors $x_1, ..., x_n \in \mathbb{K}^n$ corresponding to eigenvalues $\lambda_1, ..., \lambda_n \in \mathbb{K}$. Thus for each unit vector $x \in \mathbb{K}^n$, one has $x = \sum_{i \leq n} c_i x_i$ for some $c_1, ..., c_n \in \mathbb{K}$, and so

$$\|Xx\|^2 = \left\| \sum_{i \leq n} c_i \lambda_i x_i \right\|^2 = \sum_{i \leq n} |c_i \lambda_i|^2 \leq \max_{i \leq n} |\lambda_i|^2 \|x\|^2 = \max_{i \leq n} |\lambda_i|^2,$$

hence

$$\|Xx\| \leq \max_{i \leq n} |\lambda_i|.$$

Thus $\|X\| \leq \max_{\lambda \in \sigma(X)} |\lambda|$, as required. ■

One of the most common and important bilinear forms on $\mathbb{K}^n$ are the maps $x \mapsto \langle Ax, x \rangle$ for a fixed matrix $A \in \mathbb{K}^{n \times n}$. Here is another result to see how eigenvalues bounds such bilinear forms if the matrix is Hermitian.

Rayleigh Quotient Bound 2. *Let $X \in \mathcal{H}(n)$, and $\lambda = \sigma_{\downarrow}(X)$. One has*

$$\langle Xx, x \rangle \in \mathbb{R} \quad \text{for all } x \in \mathbb{C}^n,$$

and moreover,

$$\langle Xx, x \rangle \in [\lambda_n, \lambda_1] \quad \text{for all unit vector } x \in \mathbb{C}^n,$$

where the maxmimum/minimum is attained.

PROOF. As X is Hermitian, one has $\langle Xx, x \rangle = \langle x, Xx \rangle = \overline{\langle Xx, x \rangle}$, hence $\langle Xx, x \rangle \in \mathbb{R}$ for all $x \in \mathbb{C}^n$. Clearly maximum and minimum is attained via its unit eigenvectors. Let $x \in \mathbb{C}^n$ be a unit vector, then $x = \sum_{i \leq n} a_i v_i$ for $a_i \in \mathbb{C}$ and some set of orthonormal eigenvectors, $\{v_i\}_{i \leq n}$, of A. Now one has

$$\lambda_n = \lambda_n \|x\|^2 = \lambda_n \sum_{i \leq n} |a_i|^2 \leq \langle Xx, x \rangle = \sum_{i \leq n} \lambda_i |a_i|^2 \leq \lambda_1 \sum_{i \leq n} |a_i|^2 = \lambda_1 \|x\|^2 = \lambda_1$$

as required. ■

The space of Hermitian matrices $\mathcal{H}(n)$ and unitary matrices $\mathcal{U}(n)$ will be the most common spaces of matrices we will be working with. As stated before, these spaces can be identified as subspaces of $\mathbb{C}^{n^2}$. It is immediate that $\mathcal{H}(n)$ is a closed subspace as it is a preimage of the continuous map

$$\mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n} : X \mapsto X - X^*$$

under the set $\{0\}$. Similarly, $\mathcal{U}(n)$ is closed as it is a preimage of the continuous map

$$\mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n} : X \mapsto XX^*$$

under the set $\{I_n\}$.

It is clear that $\mathcal{H}(n)$ is unbounded as the matrix

$$\begin{pmatrix} k & 0 & \dots & 0 \\ 0 & 0 & & \\ \vdots & & \ddots & \vdots \\ 0 & & \dots & 0 \end{pmatrix}$$

has an arbitrarily large norm as $k \to \infty$.

We will show that $\mathcal{U}(n)$ is bounded. First, we need to consider a different norm.

Lemma 3. *The map*

$$\|\cdot\|_F : \mathbb{C}^{n\times n} \to \mathbb{R} : X \mapsto \sqrt{\operatorname{tr}(X^* X)}$$

is a norm on $\mathbb{C}^{n\times n}$.

PROOF. Let $X = (c_1, ..., c_n) = (x_{i,j})_{i,j\leq n} \in \mathbb{C}^{n\times n}$ (where $c_i \in \mathbb{C}^n$, $x_{ij} \in \mathbb{C}$), then by direct computation:

$$\|X\|_F = \sqrt{\operatorname{tr}\begin{pmatrix} \|c_1\|^2 & & \\ & \ddots & \\ & & \|c_n\|^2 \end{pmatrix}} = \sqrt{\sum_{i,j\leq n} |x_{ij}|^2}.$$

Then it is clear that $\|\cdot\|_F$ defines a norm on $\mathbb{C}^{n\times n}$ as this norm is the $\mathbb{C}^{n^2}$ norm. ■

The norm $\|\cdot\|_F$ defined in the lemma is called the **Frobenius norm**, with this we can show that $\mathcal{U}(n)$ is also bounded in $(\mathbb{C}^{n\times n}, \|\cdot\|)$. Note that all norms are equivalent in finite-dimensional spaces, so it suffices to show that $\mathcal{U}(n)$ is bounded under the Frobenius norm. Indeed,

$$\|U\|_F = \sqrt{\operatorname{tr}(I_n)} = \sqrt{n} \quad \text{for all } U \in \mathcal{U}(n).$$

From here, we have the immediate result.

Proposition 4. *The space* $\mathcal{U}(n)$ *is compact.*

PROOF. Since $\mathcal{U}(n)$ resides in a finite-dimensional normed space, and is closed and bounded. Then the result is immediate by Heine-Borel theorem. ■

The space $\mathcal{U}(n)$ has another nice topological property.

Proposition 5. *The space* $\mathcal{U}(n)$ *is path-connected.*

PROOF. Let $U \in \mathcal{U}(n)$, then U is unitarily diagonalizable. So there is a $V \in \mathcal{U}(n)$ and a $\alpha \in \mathbb{C}^n$ such that $U = V\operatorname{diag}(\alpha)V^*$. Note that $\det(U) = 1$, so it follows that

$$\alpha_1\alpha_2...\alpha_n = 1. \tag{3.1}$$

Now using polar representation of complex numbers, there are $\theta_1, ..., \theta_2 \in \mathbb{R}$ such that $\alpha_j = e^{i\theta_j}$ for all $j \leq n$ with $\sum_{j\leq n} \theta_j = 1$, and this sum is due to (3.1). Now consider the map

$$\gamma : [0,1] \to \mathcal{U}(n) : t \mapsto V\operatorname{diag}(e^{it\theta_1}, ..., e^{it\theta_n})V^*,$$

which is well-defined and continuous with $\gamma(0) = I_n$ and $\gamma(1) = U$. So this shows that every $U \in \mathcal{U}(n)$ has a path to I_n, so it follows that $\mathcal{U}(n)$ is path-connected. ■

3 – Roots of Positive Semi-definite Matrices

This section is exclusively for the concrete construction of the unitary Haar measure introduced in Chapter 7.3. To do that, we introduce the concept of matrix roots for positive semi-definite matrices. Given a matrix $A \in \mathbb{K}^{n\times n}$, we say $B \in \mathbb{K}^{n\times n}$ is a **square root** of A if $B^2 = A$.

Now it is clear that we do not have a natural square root map for general matrices, but it turns out we do for positive semi-definite Hermitian matrices as presented in the next theorem.

Theorem 1. *Let A be a $n \times n$ Hermitian positive semi-definite matrix over $\mathbb{C}$. Then for each $k \in \mathbb{N}$, there is a unique Hermitian positive semi-definite $B \in \mathbb{C}^{n\times n}$ such that $A = B^k$ such that:*
1. $\operatorname{rank}(A) = \operatorname{rank}(B)$;
2. There is a $p \in \mathbb{R}[x]$ such that $p(A) = B$.

Moreover, if A is invertible, then $(A^{-1})^{\frac{1}{k}} = \left(A^{\frac{1}{k}}\right)^{-1}$. If A is a matrix over $\mathbb{R}$, then everything can be replaced with $\mathbb{R}$ here.

Proof. One has $A = PDP^*$ where $P \in \mathcal{U}(n)$ and D is a diagonal matrix of nonnegative entries. Let $k \in \mathbb{N}$, and let $B = PD^{\frac{1}{k}}P^*$ where $D^{\frac{1}{k}}$ means component-wise kth root, so it is clear that $B^k = A$, $\operatorname{rank}(A) = \operatorname{rank}(B)$, and B is Hermitian positive semi-definite.
Let $\lambda_1, ..., \lambda_n \geq 0$ be the eigenvalues of A (the eigenvalues are nonnegative as A is positive semi-definite), and there is a (by Lagrange interpolation) polynomial $p \in \mathbb{R}[x]$ such that $p(\lambda_i) = \lambda_i^{\frac{1}{k}}$ for all $i \leq n$. Then

$$p(A) = p(PDP^*) = Pp(D)P^* = PD^{\frac{1}{k}}P^* = B.$$

Finally, if $C \in \mathbb{C}^{n\times n}$ is another Hermitian positive semi-definite matrix such that $A = C^k$, then $B = p(A) = p(C^k)$ shows that B commutes with C. So there is a unitary $U \in \mathcal{U}(n)$ such that $UBU^* = D_1$ and $UCU^* = D_2$ are diagonal of nonnegative entries by Corollary 3.1.6. Hence one has

$$D_1^k = UAU^* = D_2^k$$

which gives $D_1 = D_2$, hence $B = C$.
Note that $A^{-1} = PD^{-1}P^*$ (if it exists), so

$$A^{\frac{1}{k}}(A^{-1})^{\frac{1}{k}} = \left(PD^{\frac{1}{k}}P^*\right)\left(PD^{-\frac{1}{k}}P^*\right) = I_n$$

and it follows that $\left(A^{\frac{1}{k}}\right)^{-1} = (A^{-1})^{\frac{1}{k}}$. ■

Finally, we can define a root map on $SPH(n)$, which denotes the set of $n \times n$ positive semi-definite Hermitian matrices as a subspace of $\mathbb{C}^{n\times n}$.

Firstly, note that given a $X \in \mathbb{C}^{n\times n}$, then one has $X^*X \in SPH(n)$. As root preserves matrix rank, it follows that if $X \in \mathrm{GL}(n, \mathbb{C})$, then $(XX^*)^{\frac{1}{2}}$ is invertible. Thus we can consider the map:

$$F : \mathrm{GL}(n, \mathbb{C}) \to \mathcal{U}(n) : X \mapsto X(X^*X)^{-\frac{1}{2}},$$

which will be used to construct our unitary Haar measure. First of all, note that

$$F(X)F(X)^* = X(X^*X)^{-\frac{1}{2}}(X^*X)^{-\frac{1}{2}}X^* = X(X^*X)^{-1}X^* = I_n,$$

so $F(X) \in \mathcal{U}(n)$, and hence F is well-defined.

Now to show that F is Borel, it suffices to show that F is continuous, which boils down to showing that the matrix root map is continuous. Hence we just need the following proposition.

Proposition 2. *The map*

$$SPH(n) \to SPH(n) : X \mapsto X^{\frac{1}{2}}$$

is uniformly continuous.

PROOF. Let $X, Y \in SPH(n)$. and $\varepsilon > 0$. Let $x \in \mathbb{C}^n$ be a unit vector. Then following from the Cauchy-Schwarz inequality, one has

$$\langle (X - Y)x, x\rangle \leq \|X - Y\|.$$

Now suppose x is an eigenvector of $X^{\frac{1}{2}} - Y^{\frac{1}{2}}$ with eigenvalue $\mu \in \mathbb{R}$, one has that

$$\begin{aligned}\langle (X - Y)x, x\rangle = x^*(X - Y)x = x^*\left(\left(X^{\frac{1}{2}} - Y^{\frac{1}{2}}\right)X^{\frac{1}{2}} + Y^{\frac{1}{2}}\left(X^{\frac{1}{2}} - Y^{\frac{1}{2}}\right)\right)x \\ = \left(\left(X^{\frac{1}{2}} - Y^{\frac{1}{2}}\right)x\right)^* X^{\frac{1}{2}}x + \mu x^* Y^{\frac{1}{2}}x \\ = \mu x^*\left(X^{\frac{1}{2}} + Y^{\frac{1}{2}}\right)x \\ = \mu\langle\left(X^{\frac{1}{2}} + Y^{\frac{1}{2}}\right)x, x\rangle.\end{aligned}$$

By Lemma 3.2.1, choose μ such that $|\mu| = \left\|X^{\frac{1}{2}} - Y^{\frac{1}{2}}\right\|$ (as $X^{\frac{1}{2}} - Y^{\frac{1}{2}}$ is Hermitian), one has

$$\begin{aligned}\left\|X^{\frac{1}{2}} - Y^{\frac{1}{2}}\right\|^2 = |\mu|^2 = \left|\langle\left(X^{\frac{1}{2}} - Y^{\frac{1}{2}}\right)x, x\rangle\right|^2 \\ \leq \left|\langle\left(X^{\frac{1}{2}} - Y^{\frac{1}{2}}\right)x, x\rangle\right|\langle\left(X^{\frac{1}{2}} + Y^{\frac{1}{2}}\right)x, x\rangle \\ = |\mu|\langle\left(X^{\frac{1}{2}} + Y^{\frac{1}{2}}\right)x, x\rangle \\ = |\langle (X - Y)x, x\rangle| \\ \leq \|X - Y\|\end{aligned}$$

Thus choose $\delta = \varepsilon^2$, then whenever $\|X - Y\| < \delta$, one has $\left\|X^{\frac{1}{2}} - Y^{\frac{1}{2}}\right\| < \varepsilon$, as required. ∎

4 | The Horn Problem

Let $\alpha, \beta, \gamma \in (\mathbb{R}^n)_\downarrow$. The meaning of these symbols will remain the same or play the same role throughout this section. Given $A, B, C \in \mathcal{H}(n)$, we say they have **compatible eigenvalues** α, β, γ if $\sigma_\downarrow(A) = \alpha$, $\sigma_\downarrow(B) = \beta$, and $\sigma_\downarrow(C) = \gamma$.

The Horn problem tackles the question:

Given Hermitian matrices $A, B \in \mathcal{H}(n)$ with compatible eigenvalues α, β. What are the possible eigenvalues of $A + B$?

Now A. Horn conjectured a list of inequalities, dictated by something called the **admissible triples** alongside with a simple trace condition which answers the Horn question. The result is proven in a paper by Alexander Klyachko in 1998 [1], and the other by Allen Knutson and Terence Tao in 1999 [2].

In this section, we will briefly discuss what the "series of inequalities" and "admissible triples" are, therefore discussing the geometry of the Horn sets. Along the way, we will also be proving the easier necessary conditions of the Horn problem using basic linear algebra techniques.

1 – Basic Necessary Conditions of the Horn Problem

We say the triple (α, β, γ) is **Horn solvable** if there are Hermitian matrices A, B, and C with compatible eigenvalues α, β, γ, such that $A + B = C$. We also say that (A, B, C) **solves the Horn problem** for (α, β, γ). The solutions to the Horn problem can also be interpreted as finding the **Horn set**:

$$\operatorname{Horn}(\alpha, \beta) := \left\{ \lambda \in (\mathbb{R}^n)_\downarrow : (\alpha, \beta, \lambda) \text{ is Horn solvable} \right\}.$$

An immediate observation is that the Horn set is never empty as given α and β, we can choose $\gamma = \alpha + \beta \in (\mathbb{R}^n)_\downarrow$, then the Hermitian matrices $\operatorname{diag}(\alpha)$, $\operatorname{diag}(\beta)$, and $\operatorname{diag}(\gamma)$ solves the Horn problem for (α, β, γ).

Another topic of interest is to consider the set, called the **orbit** (with respect to α)

$$\mathcal{O}_\alpha := \left\{ X \in \mathcal{H}(n) : \sigma_\downarrow(X) = \alpha \right\},$$

then by the spectral theorem, one has

$$\mathcal{O}_\alpha = \{ U \operatorname{diag}(\alpha) U^* : U \in \mathcal{U}(n) \}.$$

This gives another interpretation of the Horn set:

$$\operatorname{Horn}(\alpha, \beta) = \sigma_\downarrow(\mathcal{O}_\alpha + \mathcal{O}_\beta) \quad \text{where} \quad \mathcal{O}_\alpha + \mathcal{O}_\beta = \left\{ A + B : A \in \mathcal{O}_\alpha, B \in \mathcal{O}_\beta \right\}.$$

These intepretations will be useful for further analyses on the Horn set, especially regarding its geometry and topological structure.

As the trace of a matrix over $\mathbb{C}$ is the sum of its eigenvalues (up to multiplicity) and is a linear functional on $\mathbb{C}^{n \times n}$, we immediately have the following lemma, called the trace condition, by applying the trace map onto the equation $A + B = C$.

Trace Condition 1. *If $A, B, C \in \mathcal{H}(n)$ with compatible eigevalues α, β, γ and $A + B = C$, then*

$$\sum_{i \leq n} \alpha_i + \sum_{i \leq n} \beta_i = \sum_{i \leq n} \gamma_i.$$

Hence one of the Horn problem's necessity condition is already proven. Then from here, it is already immediate that (α, β, γ) is not in general Horn solvable, for example: $\alpha = \beta = 0$, and $\gamma = (1, ..., 1)$.

Though it is clear that if $n = 1$, the triple (α, β, γ) is Horn solvable if, and only if, the trace condition is satisfied.

It is well-known that a $n \times n$ matrix (over some field) is diagonalizable if, and only if, the sum of the geometric multiplicities of its eigenvalues is n. And the algebraic multiplicity of an eigenvalue is at least its geometric multiplicity. Combining these two results, we have the following lemma:

Lemma 2. *Given a diagonalizable $n \times n$ matrix X over some field F. Then* $\operatorname{geo}(\lambda) = \operatorname{alg}(\lambda)$ *for all* $\lambda \in \sigma(X)$.

PROOF. Suppose there is a $\mu \in \sigma(X)$ such that $\operatorname{geo}(\mu) < \operatorname{alg}(\mu)$. As $\operatorname{geo}(\lambda) \leq \operatorname{alg}(\lambda)$ for all $\lambda \in \sigma(X)$, we obtain a contradiction:

$$n = \sum_{\lambda \in \sigma(X)} \operatorname{geo}(\lambda) < \sum_{\lambda \in \sigma(X)} \operatorname{alg}(\lambda) = n.$$

So it follows that $\operatorname{geo}(\mu) = \operatorname{alg}(\mu)$, as required. ■

Corollary 3. *Given a diagonalizable $n \times n$ matrix X over some field F, if $\sigma(X) = \{\lambda\}$ for some $\lambda \in F$, then $X = \lambda I_n$.*

PROOF. Note that $\det(X - tI_n) = (\lambda - t)^n$ for all $t \in F$ as λ is the only eigenvalue. Using the preceding lemma, one has

$$\dim(\ker(X - \lambda I_n)) =: \operatorname{geo}(\lambda) = \operatorname{alg}(\lambda) = n,$$

so $X - \lambda I_n = 0$, as required. ■

Using this, we can show that the Horn problem is not necessarily solvable for $n > 1$, despite (α, β, γ) satisfying the Trace Condition 4.1.1.

Choosing $\alpha = (1, ..., 1)$, $\beta = (0, ..., 0, -n)$, and $\gamma = 0$, so $\alpha, \beta, \gamma \in (\mathbb{R}^n)_\downarrow$. Hence Trace Condition 4.1.1 is satisfied. Given $A, B, C \in \mathcal{H}(n)$ with compatible eigenvalues α, β, γ.

Then by the preceding corollary, that means $A = I_n$ and $C = 0$. Hence one has

$$A + B = C \text{ imples } B = C - A = -I_n.$$

Thus $\sigma_\downarrow(B) = (-1, ..., -1) \neq \beta$.

So (α, β, γ) is not Horn solvable. From here, it is clear that

$$\operatorname{Horn}(\alpha, \beta) = \{\alpha + \beta\} \text{ if } n = 1 \text{ and } \{\alpha + \beta\} \subseteq \operatorname{Horn}(\alpha, \beta) \subsetneq \mathbb{R}^n \text{ for all } n > 1.$$

Thus trace condition is not sufficient. Now due to Hermann Weyl, there are more inequalities that the triple (α, β, γ) must satisfy for them to be Horn solvable [3, p. 291]. For now, we need a lemma.

Lemma 4. *Let X be a n-dimensional vector space and $U, V, W \subset X$ be subspaces. If*

$$\dim(U) + \dim(V) + \dim(W) \geq 2n + 1,$$

then $U \cap V \cap W \neq \{0\}$.

PROOF. By the dimension formula, note that

$$\dim(U \cap V) + \dim(W) - \dim(U \cap V \cap W) = n,$$

and

$$\dim(U) + \dim(V) - \dim(U \cap V) = n.$$

So one has

$$2n + 1 \leq \dim(U) + \dim(V) + \dim(W) = 2n + \dim(U \cap V \cap W),$$

i.e. $\dim(U \cap V \cap W) \geq 1$. So $U \cap V \cap W \neq 0$, as required. ∎

Then we have the Weyl's inequalities.

Weyl's Inequalities 5. *If $A + B = C$, then for $1 \leq i, j \leq n$, one has*

$$\gamma_{i+j-1} \leq \alpha_i + \beta_j \quad \text{if } i + j \leq n + 1,$$
$$\gamma_{i+j-n} \geq \alpha_i + \beta_j \quad \text{if } i + j \geq n + 1.$$

PROOF. Let $u_1, ..., u_n$, $v_1, ..., v_n$, and $w_1, ..., w_n$ be a basis of eigenvectors of A, B, and C respectively. For $i + j \leq n + 1$, define

$$U = \operatorname{span}(u_i, ..., u_n) \quad \text{so} \quad \dim(U) = n - i + 1$$
$$V = \operatorname{span}(v_j, ..., v_n) \quad \text{so} \quad \dim(V) = n - j + 1$$
$$W = \operatorname{span}(w_1, ..., w_{i+j-1}) \quad \text{so} \quad \dim(W) = i + j - 1.$$

Now $\dim(U) + \dim(V) + \dim(W) = 2n + 1$, so there is a unit vector $x \in U \cap V \cap W$ by preceding lemma. By Rayleigh Quotient Bound 3.2.2, one has

$$\langle Ax, x\rangle \leq \alpha_i, \quad \langle Bx, x\rangle \leq \beta_i, \quad \text{and} \quad \langle Cx, x\rangle \geq \gamma_{i+j-1},$$

hence,

$$\gamma_{i+j-1} \leq \langle Cx, x\rangle = \langle Ax, x\rangle + \langle Bx, x\rangle \leq \alpha_i + \beta_i.$$

For $i + j \geq n + 1$. Consider

$$U = \operatorname{span}(u_1, ..., u_i) \quad \text{so} \quad \dim(U) = i$$
$$V = \operatorname{span}(v_1, ..., v_j) \quad \text{so} \quad \dim(V) = j$$
$$W = \operatorname{span}(w_{i+j-n}, ..., w_n) \quad \text{so} \quad \dim(W) = 2n - i - j + 1.$$

Then it follows similarly as above. ∎

We shall see that in Chapter 4.3 that the Trace Condition 4.1.1 together with Weyl's Inequalities 4.1.5 is sufficient for the triple (α, β, γ) to be Horn solvable for $n = 2$.

2 – Unitary Group and the Horn Problem

In this chapter, we will look at how the Horn problem can be simplified and reformulated using the properties of unitary and Hermitian matrices.

Observe that $\mathcal{U}(n)$ is a subgroup of $\mathrm{GL}(n, \mathbb{C})$ under matrix multiplication, hence one can define an equivalence relation $\sim$ on $\mathcal{H}(n)$ such that $X \sim Y$ if, and only if, there is some $U \in \mathcal{U}(n)$ such that $X = UYU^* = UYU^{-1}$, and we shall say that X and Y are **unitarily similar**. As this relation is a conjugation relation under unitary matrices, it is indeed an equivalence relation.

Similarly we define $\sim$ on the set $\{(X, Y, Z) \in \mathcal{H}(n)^3 : X + Y = Z\}$ where $(X_1, X_2, X_3) \sim (Y_1, Y_2, Y_3)$ if, and only if, there is a $U \in \mathcal{U}(n)$ such that $X_i = UY_iU^*$ for all $i \leq 3$, and we shall say the triple (X_1, X_2, X_3) is **unitarily similar** to (Y_1, Y_2, Y_3). As matrix multiplication is distributive over matrix addition, and again this is a conjugation relation, it immediately follows that this relation is well-defined and is an equivalence relation. We shall denote $[(X_1, X_2, X_3)]$ to be the equivalence class under $\sim$, which is called a **unitary class** of (X_1, X_2, X_3).

Since the relation $\sim$ preserves matrix similarity, it follows that the eigenvalues are also preserved. Hence we have the following proposition.

Proposition 1. *The Hermitian matrices (A, B, C) solves the Horn problem for (α, β, γ) if, and only if, each $(X, Y, Z) \in [(A, B, C)]$ solves the Horn problem for (α, β, γ).*

PROOF. As discussed above, note that $A + B = C$ if, and only if, $UAU^* + UBU^* = UCU^*$, and $\sigma_\downarrow(A) = \sigma_\downarrow(UAU^*)$ etc. for B and C, for all $U \in \mathcal{U}(n)$. Then the proposition follows. ■

This proposition is great because if one triple of Hermitian matrices (A, B, C) solves a Horn problem, then we will have a family of matrices, which are unitarily similar to (A, B, C) which also solves the Horn problem. Define $\mathcal{O} := \{[(X, Y, Z)] : X, Y, Z \in \mathcal{H}(n)\}$ and the map

$$\mathrm{Eigen} : \mathcal{O} \to (\mathbb{R}^n)^3_\downarrow : [(X, Y, Z)] \to \big(\sigma_\downarrow(X), \sigma_\downarrow(Y), \sigma_\downarrow(Z)\big).$$

Thus the solution to Horn problem is the same as the solutions to the preimages of this map. It is already shown in Part 4.1 that this map is not surjective for $n > 1$.

Therefore the natural question that arises is if the Eigen map is injective, that is if (α, β, γ) is necessarily Horn solvable from only one unitary class of Hermitian matrices. Now as $U(1) = \{z \in \mathbb{C} : |z| = 1\}$, it is clear that the question is true for $n = 1$, but it is not that case for $n > 1$.

Let $\alpha = (1, ..., 1, 0)$, $\beta = (1, 0, ..., 0)$, and $\gamma = (1, ..., 1)$, so $\alpha, \beta, \gamma \in (\mathbb{R}^n)_\downarrow$. Clearly $\mathrm{diag}(\alpha) + \mathrm{diag}(0, ..., 0, 1) = \mathrm{diag}(\gamma) = I_n$ gives a solution to the Horn Problem for (α, β, γ).

Now given $P \in \mathcal{U}(n)$, observe that $P^*\mathrm{diag}(0, ..., 0, 1)P$, and $P^*\ \mathrm{diag}(\alpha)P$ must have nonnegative entries on the main diagonal. Now consider

$$A = \begin{pmatrix} A' & 0 \\ 0 & I_{n-2} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} B' & 0 \\ 0 & 0 \end{pmatrix}$$

where

$$A' = \begin{pmatrix} 2 & 2 \\ -1 & -1 \end{pmatrix} \quad \text{and} \quad B' = \begin{pmatrix} -1 & -2 \\ 1 & 2 \end{pmatrix}.$$

Note if $n = 2$, then take $A = A'$ and $B = B'$. So

$$A + B = \begin{pmatrix} A' + B' & 0 \\ 0 & I_{n-2} \end{pmatrix} = I_n$$

and $\chi_A(\lambda) = (1-\lambda)^{n-2}\chi_{A'}(\lambda) = -\lambda(1-\lambda)^{n-1}$, $\chi_B(\lambda) = \chi_{B'}(\lambda) = \lambda^{n-1}(\lambda - 1)$.

It follows that $\sigma_\downarrow(A) = \alpha$ and $\sigma_\downarrow(B) = \beta$. Now A and B has negative entries on the diagonal, so $(A, B, I_n) \notin [(\text{diag}(\alpha), \text{diag}(0, ..., 0, 1), \text{diag}(\gamma))]$, but $\text{Eigen}([(A, B, 1_n)]) = (\alpha, \beta, \gamma)$. So the noninjectivity of Eigen for $n > 1$ is shown.

Another benefit of identifying the solutions of the Horn problem through unitary classes is of the next proposition where it gives another depiction of the Horn set which can be useful for computing the Horn set.

Proposition 2. *The Horn set satisifes*

$$\text{Horn}(\alpha, \beta) = \{\sigma_\downarrow(\text{diag}(\alpha) + U\text{diag}(\beta)U^*) : U \in \mathcal{U}(n)\}.$$

PROOF. Given $\lambda \in \text{Horn}(\alpha, \beta)$, as (α, β, λ) is Horn solvable, there is a triple of Hermitian matrices (A, B, C) with compatible eigenvalues α, β, λ such that $A + B = C$. As $A \sim \text{diag}(\alpha)$ and $B \sim \text{diag}(\beta)$, there are $U_1, U_2 \in \mathcal{U}(n)$ such that $U_1 A U_1^* = \text{diag}(\alpha)$ and $B = U_2 \text{diag}(\beta) U_2^*$. Define $U = U_1^{-1} U_2$ and conjugate the equation $A + B = C$ under U_1^{-1} to get $\text{diag}(\alpha) + U\text{diag}(\beta)U^* = U_1^* C U_1$. Thus $\lambda = \sigma_\downarrow(\text{diag}(\alpha) + U\text{diag}(\beta)U^*)$.
Given $\lambda = \sigma_\downarrow(\text{diag}(\alpha) + U\text{diag}(\beta)U^*)$ for some $U \in \mathcal{U}(n)$, then it is clear that the Hermitian matrices $(\text{diag}(\alpha), U\text{diag}(\beta)U^*, \text{diag}(\alpha) + U\text{diag}(\beta)U^*)$ solves the Horn problem for (α, β, λ), i.e. $\lambda \in \text{Horn}(\alpha, \beta)$. Hence the proposition is proven. ■

With this, the Horn set can be computed stochastically by taking uniformly distributed samples of $\mathcal{U}(n)$.

By the preceding proposition, one has that $\text{Horn}(\alpha, \beta)$ is the image of the map

$$F : \mathcal{U}(n) \to \mathbb{R}^n : U \mapsto \sigma_\downarrow(\text{diag}(\alpha) + U\text{diag}(\beta)U^*).$$

We shall see that this map is continuous, so in particular, $\text{Horn}(\alpha, \beta)$ is a continuous image of $\mathcal{U}(n)$.

To show that F is continuous, it suffices to prove that the map $X \mapsto \sigma_\downarrow(X)$ on the set of Hermitian matrices is continuous as it is clear that the map $U \mapsto \text{diag}(\alpha) + U\text{diag}(\beta)U^*$ is continuous.

Lemma 3. *The 'eigenvalue' map $\sigma_\downarrow : \mathcal{H}(n) \to \mathbb{R}^n$ is Lipschitz.*

PROOF. Let $A, B \in \mathbb{C}^{n\times n}$ with compatible eigenvalues $\alpha, \beta \in \mathbb{R}^n$. For each $i \leq n$, define the component map $\sigma_i : \mathcal{H}(n) \to \mathbb{R}$ of $\sigma_\downarrow$, i.e. $\sigma_\downarrow = (\sigma_1, ..., \sigma_n)$. So it suffices to prove σ_i is continuous. Now note that $\beta_i = \sigma_i(B) = \sigma_i(--B) = -\sigma_{n-i+1}(-B)$, so $-\beta_i = -\sigma_i(B) = \sigma_{n-i+1}(-B)$. Note that by <u>Weyl's Inequalities 4.1.5</u>, one has

$$\sigma_n(A - B) \leq \sigma_i(A) + \sigma_{n-i+1}(-B) \quad \text{and} \quad \sigma_1(A - B) \geq \sigma_i(A) + \sigma_{n-i+1}(-B).$$

So

$$|\sigma_i(A) - \sigma_i(B)| \leq |\sigma_i(A) + \sigma_{n-i+1}(-B)| \leq \max\{|\sigma_n(A - B)|, |\sigma_1(A - B)|\} \leq \|A - B\|$$

shows that σ_i is Lipschitz. Thus it follows that $\sigma_\downarrow$ is also Lipschitz, as required. ■

Thus by the preceding lemma, F is continuous. As $\mathcal{U}(n)$ is compact and path-connected, we immediately see that $\mathrm{Horn}(\alpha, \beta)$ is also compact and path-connected. Note that we can also achieve the same result with the $\mathrm{Horn}(\alpha, \beta) = \sigma_{\downarrow}(\mathcal{O}_\alpha + \mathcal{O}_\beta)$ identification.

Since $\mathrm{Horn}(\alpha, \beta)$ is compact, then it is bounded. By Weyl's Inequalities 4.1.5 for $i = j = 1$, $i = n+1, j =$, and $i = 1, j = n$ in the statement, it is immediate that

$$\max(\alpha_1 + \beta_n, \alpha_n + \beta_1) \le \gamma_1 \le \alpha_1 + \beta_1, \tag{4.1}$$

i.e. $\|\gamma\| \le n|\gamma_1| \le nr$ for all $\gamma \in \mathrm{Horn}(\alpha, \beta)$ where r is the maximum of the absolute value of the bounding values in (4.1). Hence the (metric) diameter of $\mathrm{Horn}(\alpha, \beta)$ is contained in the closed ball of radius nr centered at the origin. Albeit this is not a precise approximation, but it is a simple bound.

3 – The $n = 2$ Horn Problem

For $n = 2$, the Trace Condition 4.1.1 states

$$\gamma_1 + \gamma_2 = \alpha_1 + \alpha_2 + \beta_1 + \beta_2 \tag{4.2}$$

and Weyl's Inequalities 4.1.5 gives

$$\begin{aligned} \alpha_2 + \beta_2 \le \gamma_2 &\le \min(\alpha_1 + \beta_2, \alpha_2 + \beta_1) \\ &\le \max(\alpha_1 + \beta_2, \alpha_2 + \beta_1) \le \gamma_1 \le \alpha_1 + \beta_1. \end{aligned} \tag{4.3}$$

Thus given $A, B, C \in \mathcal{H}(2)$ with compatible eigenvalues α, β, γ. For $A + B = C$ to hold, the conditions (4.2) and (4.3) must be satisfied. Observe that (4.2) tells us that γ lies on the line

$$x + y = \alpha_1 + \alpha_2 + \beta_1 + \beta_2 \quad (x, y \in \mathbb{R})$$

and (4.3) furthermore tells us that γ lies on the line segment between the points

$$(\max(\alpha_1 + \beta_2, \alpha_2 + \beta_1), \min(\alpha_1 + \beta_2, \alpha_2 + \beta_1)) \quad \text{and} \quad (\alpha_1 + \beta_1, \alpha_2 + \beta_2). \tag{4.4}$$

Now we claim that if α, β, γ satisfies (4.2) and (4.3), then (α, β, γ) is Horn solvable.

Let $A = \mathrm{diag}(\alpha)$, and $B = \mathrm{diag}(\beta)$. Define

$$U_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \quad \text{for all } \theta \in \mathbb{R}$$

so $U_\theta \in U(2)$. Let

$$C_\theta := A + U_\theta B U_\theta^* \in \mathcal{H}(2).$$

Look into Appendix 11.1 to see all of following the computations. We see that C_θ has eigenvalues

$$\lambda_1(\theta) = \overline{\alpha} + \overline{\beta} + \sqrt{\frac{\Delta(\theta)}{4}} \quad \text{and} \quad \lambda_2(\theta) = \overline{\alpha} + \overline{\beta} - \sqrt{\frac{\Delta(\theta)}{4}}$$

where

$$\overline{\alpha} = \frac{\alpha_1 + \alpha_2}{2}, \quad \overline{\beta} = \frac{\beta_1 + \beta_2}{2}$$

and $\theta \mapsto \Delta(\theta)$ is a continuous function with maximmum $\Delta(0) = \left(\underline{\alpha} + \underline{\beta}\right)^2$, and minimum $\Delta(\frac{\pi}{2}) = \left(\underline{\alpha} - \underline{\beta}\right)^2$ where

$$\underline{\alpha} = \frac{\alpha_1 - \alpha_2}{2}, \quad \underline{\beta} = \frac{\beta_1 - \beta_2}{2}.$$

Now taking $\lambda(\theta) := (\lambda_1(\theta), \lambda_2(\theta)) \in (\mathbb{R}^2)_\downarrow$, $\lambda(\theta)$ exactly traces out the line segment (4.4) for $\theta \in [0, \frac{\pi}{2}]$. That means there is a $\varphi \in [0, \frac{\pi}{2}]$, one has $\sigma_\downarrow(C_\varphi) = \gamma$. Hence $(A, U_\varphi B U_\varphi^*, C_\varphi)$ solves the Horn problem for (α, β, γ).

That means we have the following result:

Theorem 1. *Given $\alpha, \beta, \gamma \in (\mathbb{R}^2)_\downarrow$. Then (α, β, γ) is Horn solvable if, and only if α, β, γ satisfies Trace Condition 4.1.1 and Weyl's Inequalities 4.1.5.*

Or equivalently, we have a nice depiction of the Horn set:

Corollary 2. *Given $\alpha, \beta \in (\mathbb{R}^2)_\downarrow$. The set* $\operatorname{Horn}(\alpha, \beta)$ *is a line segment between the points*

$$(\max(\alpha_1 + \beta_2, \alpha_2 + \beta_1), \min(\alpha_1 + \beta_2, \alpha_2 + \beta_1)) \quad \text{and} \quad (\alpha_1 + \beta_1, \alpha_2 + \beta_2).$$

To given an example, take $\alpha = (1, 0)$, and $\beta = (3, -1)$. One has

$$\alpha_1 + \beta_1 = 4, \quad \alpha_1 + \beta_2 = 0, \quad \alpha_2 + \beta_1 = 3, \quad \alpha_2 + \beta_2 = -1.$$

So $\operatorname{Horn}(\alpha, \beta)$ is the line segment between $(3, 0)$ and $(4, -1)$ shown in Figure 1.

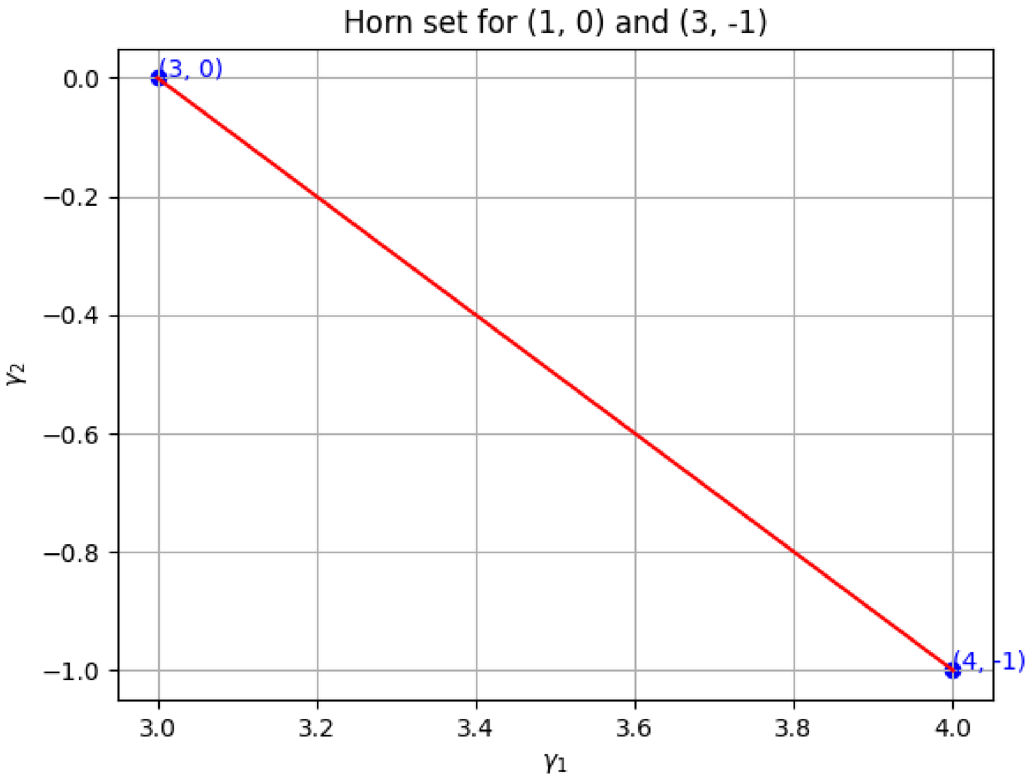


Figure 1: The set $\operatorname{Horn}((1, 0), (3, -1))$ as illustrated by the red line. The x-axis represents γ_1 and the y-axis represents γ_2.

To really drive the point home, we can have two main interpretations:

- For any γ that lies on that line segment, one can find Hermitian matrices $A, B, C \in \mathcal{H}(2)$ with compatible eigenvalues α, β, γ such that $A + B = C$.
- The set of eigenvalues as an ordered pair of $A + B$ where $A, B \in \mathcal{H}(2)$ with compatible eigenvalues α, β is that line segment; that line segment is the image of the map:

$$\mathcal{O}_\alpha \times \mathcal{O}_\beta \to \mathbb{R}^2 : (A, B) \mapsto \sigma_\downarrow(A + B).$$

Here is a Desmos model that is able to plot the Horn set for arbitrary α and β in 2 dimension.

4 – The General Case

Alas, we shall discuss the Horn problem for an arbitrary n. First we need some notation:

Fix a $r \in \{1, ..., n-1\}$ (if $n = 1$, then the sets defined below can be treated as empty). Given $I \subseteq \{1, ..., n\}$ with cardinality r. Then we write $I = \{I(1), ..., I(r)\}$ where $I(1) < I(2) < ... < I(r)$. Now define

$$S_r^n := \left\{(I, J, K) \subseteq \{1, ..., n\} : |I| = |J| = |K| = r, \quad \sum_{i=1}^{r} I(i) + J(i) - K(i) = \frac{r(r+1)}{2}\right\},$$

and

$$\begin{aligned} T_1^n &:= S_1^n \\ T_r^n &:= \left\{ (I, J, K) \in S_r^n : \sum_{i=1}^{p} I(U(i)) + J(V(i)) - K(W(i)) \leq \binom{p+1}{2} \right. \\ &\qquad \left. \text{for all } (U, V, W) \in T_p^r, p \in \{1, ..., r-1\} \right\}. \end{aligned}$$

We call the elements of T_r^n **admissible triples**.

Finally, A. Horn conjectured the following in regards to the Horn problem, which was proven by A. Klyachko [1].

Horn's Conjecture 1. *The triple (α, β, γ) (elements of $(\mathbb{R}^n)_\downarrow$) is Horn solvable if, and only if, (α, β, γ) satisfies the Trace Condition 4.1.1 and*

$$\sum_{k \in K} \gamma_k \leq \sum_{i \in I} \alpha_i + \sum_{j \in J} \beta_j \tag{4.5}$$

for all admissible triples $(I, J, K) \in T_r^n$ for all $r \in \{1, ..., n-1\}$.

By the Trace Condition 4.1.1, it shows that the set of such $\gamma \in \text{Horn}(\alpha, \beta)$ lies on the hyperplane with a normal vector of $(1, ..., 1) \in \mathbb{R}^n$ shifted by $\alpha_1 + ... + \alpha_n + \beta_1 + ... + \beta_n$. Now the inequalities (4.5) further shows that those γ's are contained in some hypercube, hence $\text{Horn}(\alpha, \beta)$ is a convex polytope in $\mathbb{R}^n$.

This result generalises the $n = 2$ case in Part 4.3. Indeed, for $n = 2$, observe that

$$S_1^2 = \{(\{1\}, \{1\}, \{1\}), (\{1\}, \{2\}, \{2\}), (\{2\}, \{1\}, \{2\})\} = T_1^2,$$

then (4.5) gives, which is the first half of Weyl's Inequalities 4.1.5:

$$\gamma_1 \leq \alpha_1 + \alpha_2, \quad \gamma_2 \leq \alpha_1 + \beta_2, \quad \gamma_2 \leq \alpha_2 + \beta_1,$$

and when combined with the Trace Condition 4.1.1, one immediately gets the entirety of Weyl's Inequalities 4.1.5.

To understand this conjecture, we shall provide an example for $n = 3$ case.

For $n = 3$, through direct computation, one has

$$\begin{gathered} S_1^3 = \{(\{1\},\{1\},\{1\}),(\{1\},\{2\},\{2\}), \\ (\{1\},\{3\},\{3\}),(\{2\},\{1\},\{2\}), \\ (\{2\},\{2\},\{3\}),(\{3\},\{1\},\{3\})\} = T_1^3 \end{gathered}$$

$$\begin{gathered} S_2^3 = \{(\{1,2\},\{1,2\},\{1,2\}),(\{1,2\},\{1,3\},\{1,3\}), \\ (\{1,2\},\{2,3\},\{2,3\}),(\{1,3\},\{1,2\},\{1,3\}), \\ (\{1,3\},\{1,3\},\{2,3\}),(\{2,3\},\{1,2\},\{2,3\})\} \end{gathered}$$

$$\begin{gathered} T_2^3 = \{(\{1,2\},\{1,2\},\{1,2\}),(\{1,2\},\{1,3\},\{1,3\}), \\ (\{1,2\},\{2,3\},\{2,3\}),(\{1,3\},\{1,2\},\{1,3\}), \\ (\{1,3\},\{1,3\},\{2,3\}),(\{2,3\},\{1,2\},\{2,3\})\}. \end{gathered}$$

Suppose $\alpha = (2,1,0)$, and $\beta = (3,0,-1)$. Then by Trace Condition 4.1.1, one has

$$\gamma_1 + \gamma_2 + \gamma_3 = 2 + 1 + 0 + 3 + 0 - 1 = 5. \tag{4.6}$$

Now T_1^3 with (4.5) gives:

$$\begin{aligned} \gamma_1 \le \alpha_1 + \beta_1 = 5 \quad \gamma_2 \le \alpha_1 + \beta_2 = 2 \quad \gamma_3 \le \alpha_1 + \beta_3 = 1 \\ \gamma_2 \le \alpha_2 + \beta_1 = 4 \quad \gamma_3 \le \alpha_2 + \beta_2 = 1 \quad \gamma_3 \le \alpha_3 + \beta_1 = 3 \end{aligned} \tag{4.7}$$

and T_2^3 with (4.5) gives:

$$\begin{aligned} \gamma_1 + \gamma_2 \le \alpha_1 + \alpha_2 + \beta_1 + \beta_2 = 6 \quad \gamma_1 + \gamma_3 \le \alpha_1 + \alpha_2 + \beta_1 + \beta_3 = 5 \\ \gamma_2 + \gamma_3 \le \alpha_1 + \alpha_2 + \beta_2 + \beta_3 = 2 \quad \gamma_1 + \gamma_3 \le \alpha_1 + \alpha_3 + \beta_1 + \beta_2 = 5 \\ \gamma_2 + \gamma_3 \le \alpha_1 + \alpha_3 + \beta_1 + \beta_3 = 4 \quad \gamma_2 + \gamma_3 \le \alpha_2 + \alpha_3 + \beta_1 + \beta_2 = 3 \end{aligned} \tag{4.8}$$

Upon simplifying (4.7), one has

$$\gamma_1 \le 5 \quad \gamma_2 \le 2 \quad \gamma_3 \le 1, \tag{4.9}$$

and upon simplifying (4.8), one has

$$\gamma_1 + \gamma_2 \le 6 \quad \gamma_1 + \gamma_3 \le 5 \quad \gamma_2 + \gamma_3 \le 2. \tag{4.10}$$

Combining (4.6) with (4.10), and including (4.7), one gets a rectangle:

$$3 \le \gamma_1 \le 5 \quad 0 \le \gamma_2 \le 2 \quad -1 \le \gamma_3 \le 1.$$

where (4.6) is bounded by, which is illustrated as a hexagon in Figure 2.

To recall, what this means is that:

- For any γ that lies on that hexagon, one can find Hermitian matrices $A, B \in \mathcal{H}(3)$ such that $A, B, A + B$ have compatible eigenvalues α, β, γ.
- The set of eigenvalues as an ordered pair of $A + B$ where $A, B \in \mathcal{H}(3)$ with compatible eigenvalues α, β is that line segment; that hexagon is the image of the map:

$$\mathcal{O}_\alpha \times \mathcal{O}_\beta \to \mathbb{R}^2 : (A, B) \mapsto \sigma_\downarrow(A + B).$$

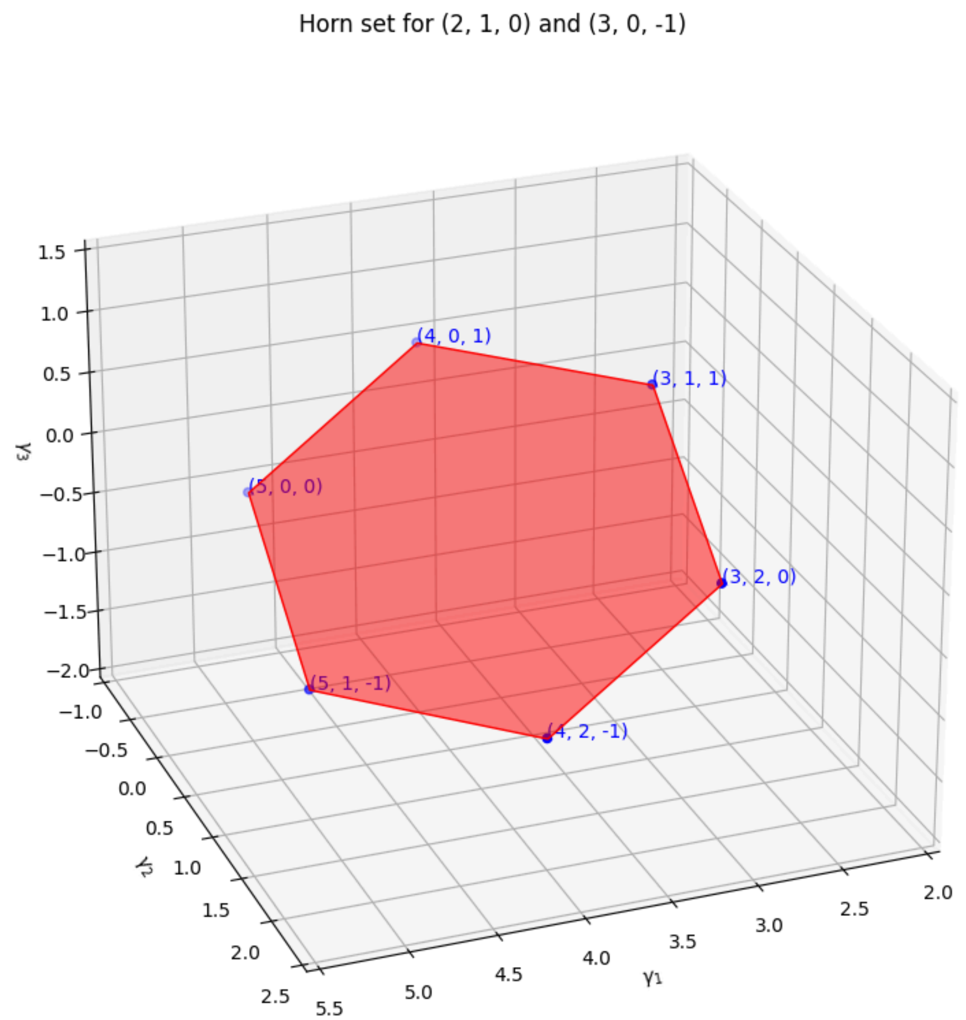


Figure 2: The set $\mathrm{Horn}((2,1,0),(3,0,-1))$ as illustrated by the red hexagon where the blue points are its vertices.

5 | Measure Theory

In this section, we shall briefly study the prerequisite knowledge required for the analysis done on the probabilistic Horn problem.

Starting from Radon measures, which will serve as the building block of Haar measures, we shall investigate the relevant properties of Radon measures, and methods of constructing Radon measures. With one of the main goal proving the famous Riesz-Markov-Kakutani representation theorem.

We will then be looking at some basic properties of measures such as supports and pushforwards, which will aid in our construction of Haar measures and developing the theory required for the probabilistic Horn problem.

1 – Radon Measures

Given a topological space X, we denote by $\mathcal{B}(X)$ the Borel σ-algebra of X. This makes $(X, \mathcal{B}(X))$ a measurable space, and any measure on that space is called a **Borel measure**. Now Borel measures can satisfy certain properties which can make them more 'compatible' with the topology on X. We say μ is

(i) **Outer regular on B** for some Borel $B \subseteq X$ if we have the following equality:

$$\mu(B) = \inf\{\mu(U) : U \supseteq B, U \subseteq X \text{ open}\}.$$

(ii) **Outer regular** if the above condition holds for all Borel $B \subseteq X$.

(iii) **Locally finite** if for each $x \in X$, there is a $N \in \mathcal{N}_X[x]$ such that $\mu(N) < \infty$.

Now assuming X is Hausdorff, hence comapct sets of X are Borel as they are closed.

(iv) **Inner regular on B** for some Borel $B \subseteq X$ if we have the following equality:

$$\mu(B) = \sup\{\mu(K) : K \subseteq B, K \subseteq X \text{ compact}\}.$$

(v) **Inner regular** if the above condition holds for all Borel $B \subseteq X$. In particular, we say μ is **weakly inner regular** if the equality above only holds for open sets.

Finally a Borel measure on a Hausdorff space is called **regular** if it is outer regular and inner regular, and **Radon** if it is outer regular, weakly inner regular, and locally finite.

We shall begin with a simple result to warm-up to these notions.

Lemma 1. *Let μ be a Borel measure on a space X. If μ is locally finite, then μ is finite on all compact subsets of X. The converse holds if X is locally compact.*

PROOF. Suppose μ is locally finite. Then given a compact $K \subseteq X$, and for each $x \in K$, there is a $N_x \in \mathcal{N}[x]$ such that $\mu(N_x) < \infty$. Now $\{N_x\}_{x \in K}$ (openly) covers K, hence by compactness, there is a $n \in \mathbb{N}$ with $\{N_i\}_{i \leq n} \subseteq \{N_x\}_{x \in K}$ such that $K \subseteq \bigcup_{i \leq n} N_i$. Finally,

$$\mu(K) \leq \mu\left(\bigcup_{i \leq n} N_i\right) \leq \sum_{i \leq n} \mu(N_i) < \infty,$$

as required.
Suppose μ is finite on all compact sets and X is locally compact. Then the result follows as each $x \in X$ has a compact neighbourhood. ■

From here on, we will see how continuous functions on locally compact Hausdorff (LCH) spaces can induce and approximate Radon measures. Here we start with some two important lemmas.

The first lemma shows how we can find nontrivial continuous functions with an arbitrarily small compact support.

Lemma 2. *Given an open U, compact $K \subseteq U$, on a LCH space X. There is a $f \in C_c(X, [0,1])$ such that $f|_K = 1$ and $\operatorname{supp}(f) \subseteq U$, and hence $1_K \leq f \leq 1_U$. In particular, $C_c(X, [0,1])$ separates the points of X.*

PROOF. For each $x \in K$, by property of LCH, there is an open $N_x \in \mathcal{N}[x]$ such that $\overline{N_x} \subseteq U$ is compact. Now $\{N_x\}_{x \in K}$ covers K, so there is a $n \in \mathbb{N}$ with $\{N_i\}_{i \leq n} \subseteq \{N_x\}_{x \in K}$ such that

$$K \subseteq \bigcup_{i \leq n} N_i \subseteq \bigcup_{i \leq n} \overline{N_i} \subseteq U.$$

Note that $M := \bigcup_{i \leq n} N_i$ is open and $\overline{M}$ is compact. Furthermore $\overline{M}$ is normal as it is also Hausdorff, so by Urysohn's lemma, there is a $g \in C\left(\overline{M}, [0,1]\right)$ such that $g|_K = 1$ and $g|_{\overline{M} \setminus M} = 0$. Now we define

$$f : X \to [0,1] : x \mapsto \begin{cases} g(x) & \text{if } x \in \overline{M} \\ 0 & \text{otherwise} \end{cases},$$

which is continuous as both $f|_M \ (= g|_M)$ and $f|_{X \setminus \overline{M}} \ (= 0)$ are continuous. It follows that $\operatorname{supp}(f) \subseteq \overline{M} \subseteq U$, as required.
To prove that $C_c(X, [0,1])$ separates the points of X, given $x, y \in X$ be distinct. Then by Hausdorff-ness, there is an open $U \in \mathcal{N}[x]$ such that $y \notin U$. Then by locally compactness, there is compact K such that $x \in K \subseteq U$, so there is a $f \in C_c(X, [0,1])$ such that $1_K \leq f \leq 1_U$. Hence $f(x) = 1 \neq 0 = f(y)$, as required. ■

The next lemma states that Radon measures on LCH spaces are uniquely determined by integration through compactly supported continuous functions.

Lemma 3. *Let μ and ν be Radon measures on a LCH space X. If*

$$\int_X f \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\nu \quad \text{for all } f \in C_c(X)$$

then $\mu = \nu$.

PROOF. Let $U \subseteq X$ be open. Then for any compact $K \subseteq U$, there is a $f \in C_c(X)$ such that $1_K \leq f \leq 1_U$ by Lemma 5.1.2. Thus

$$\mu(K) = \int_X 1_K \, \mathrm{d}\mu \leq \int_X f \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\nu \leq \int_X 1_U \, \mathrm{d}\nu = \nu(U).$$

This holds for all compact $K \subseteq U$, then by weakly inner regularity of μ, one has $\mu(U) \leq \nu(U)$. Similarly, $\nu(U) \leq \mu(U)$, hence $\mu(U) = \nu(U)$. Then as Radon measures are uniquely determined by open sets due to outer regularity, one has that $\mu = \nu$, as required. ■

Given a linear functional I on $C_c(X)$, we say I is **positive** if $I(f) \geq 0$ for all $f \in C_c(X, \mathbb{R}_{\geq 0})$. From here, we shall see how Radon measures can be constructed from linear functionals by proving the celebrated Riesz-Markov-Kakutani Representation Theorem. The proof is inspired by Folland [4, 7.2].

Riesz-Markov-Kakutani Representation Theorem 4. *Given a LCH space X and a positive linear functional I on $C_c(X)$. There is a unique Radon measure μ on X such that*

$$I(f) = \int_X f \, \mathrm{d}\mu \quad \text{for all } f \in C_c(X). \tag{5.1}$$

Furthermore, μ satisfies:

$$\mu(U) = \sup\{I(f) : f \in C_c(X, [0,1]),\ \operatorname{supp}(f) \subseteq U\} \text{ for all open } U \subseteq X, \tag{5.2}$$

and

$$\mu(K) = \inf\{I(f) : f \in C_c(X, [0,1]), 1_K \leq f\} \text{ for all compact } K \subseteq X. \tag{5.3}$$

Proof. Uniqueness follows from Lemma 5.1.3, so it suffices to prove existence.
For any open $U \subseteq X$, define

$$\mu(U) := \sup\{I(f) : f \in C_c(X, [0,1]),\ \operatorname{supp}(f) \subseteq U\}$$

We also define our potential outer measure $\mu^* : \mathcal{P}(X) \to [0, \infty]$ as

$$\mu^*(E) := \inf\{\mu(U) : U \supseteq E, U \subseteq X \text{ open}\}.$$

for any $E \subseteq X$. Note that $\mu(U) = \mu^*(U)$ for any open $U \subseteq X$.
We shall show that:
(i) μ^* is an outer measure.
(ii) Every open set is μ^*-measurable,
So following from Cathéodory's theorem, one has that $\mu = \mu^*|_{\mathcal{B}(X)}$ is an outer regular measure (which makes sense as $\mu(U) = \mu^*(U)$ for all open $U \subseteq X$). Clearly (5.2) is satisfied, then we prove that:
(iii) μ also satisfies (5.3).
Once (5.3) is proven, then μ is clearly finite on compact sets, and one has μ is also weakly inner regular. Indeed, given an open $U \subseteq X$ and $a \in \mathbb{R}$ such that $\mu(U) > a$, by (5.2), there is a $f \in C_c(X, [0,1])$ with $\operatorname{supp}(f) \subseteq U$ such that $I(f) > a$. Take $K = \operatorname{supp}(f)$, then for any $g \in C_c(X)$ such that $g \geq 1_K$, one has that $g - f \geq 0$, hence $I(g) \geq I(f) > a$, thus by (5.3), one has $\mu(K) > a$. Hence μ is weakly inner regular.
Finally, we prove that
(iv) $\int_X f \, \mathrm{d}\mu = I(f)$ for all $f \in C_c(X)$.
With (i) to (iv), our proof will be completed.

Proof of (i). Let $E \subseteq X$, $U \subseteq X$ be open, and $\{U_i\}_{i \in \mathbb{N}}$ be a collection of open sets such that $U = \bigcup_{i \in \mathbb{N}} U_i$. Let $f \in C_c(X)$ such that $K := \operatorname{supp}(f) \subseteq U$, then by compactness of K, one has that $K \subseteq \bigcup_{i \leq n} U_i$ for some $n \in \mathbb{N}$. Then there are $g_1, ..., g_n \in C_c(X, [0,1])$ such that $\operatorname{supp}(g_i) \subseteq U_i$ and $\sum_{i \leq n} g_i = 1$ on K following from Lemma 5.1.2. So $f = \sum_{i \leq n} f g_i$, but $\operatorname{supp}(f g_i) \subseteq U_i$, so by (5.2), one has

$$I(f) = \sum_{i \leq n} I(f g_i) \leq \sum_{i \leq n} \mu(U_i) \leq \sum_{i \in \mathbb{N}} \mu(U_i).$$

This holds for all such f's, so $\mu(U) \leq \sum_{i \in \mathbb{N}} \mu(U_i)$. Thus it is clear that

$$\mu^*(E) \leq \inf\left\{ \sum_{i \in \mathbb{N}} \mu(U_i) : (U_i)_{i \in \mathbb{N}} \text{ is open in } X \text{ and } E \subseteq \bigcup_{i \in \mathbb{N}} U_i \right\} =: a,$$

and if $\mu^*(E) < \infty$ then given a $\varepsilon > 0$, there is an open $U \supseteq E$ such that

$$\mu^*(E) + \varepsilon > \mu(U) \geq a.$$

Taking $\varepsilon \downarrow 0$, one has $\mu^*(E) = a$. So by [4, 1.10], μ^* defines an outer measure.

Proof of (ii). As (i) holds, it suffices to show that given an open $U \subseteq X$ and any $E \subseteq X$ with $\mu^*(E) < \infty$, one has

$$\mu^*(E) \geq \mu^*(E \cap U) + \mu^*(E \setminus U). \tag{5.4}$$

If E is open, then so is $E \cap U$. Given a $\varepsilon > 0$, by (5.2), there is a $f \in C_c(X, [0,1])$ with $\operatorname{supp}(f) \subseteq E \cap U$ such that $I(f) + \varepsilon > \mu^*(E \cap U)$. Now $E \setminus \operatorname{supp}(f)$ is also open, so again, there is a $g \in C_c(X, [0,1])$ with $\operatorname{supp}(g) \subseteq E \setminus \operatorname{supp}(f)$ such that $I(g) + \varepsilon > \mu^*(E \setminus \operatorname{supp}(f))$. As $\operatorname{supp}(f+g) \subseteq E$, one has that

$$\begin{aligned}\mu^*(E) &\geq I(f+g) = I(f) + I(g) \\ &> \mu^*(E \cap U) + \mu^*(E \setminus \operatorname{supp}(f)) - 2\varepsilon \\ &\geq \mu^*(E \cap U) + \mu^*(E \setminus U) - 2\varepsilon.\end{aligned}$$

Take $\varepsilon \downarrow 0$ and we get the (5.4) holds for open sets.

If E is not necessarily open, then by definition of μ^*, given an $\varepsilon > 0$, there is an open $V \supseteq E$ such that $\mu^*(E) + \varepsilon > \mu(V)$, i.e.

$$\begin{aligned}\mu^*(E) + \varepsilon &> \mu(V) = \mu^*(V) \\ &\geq \mu^*(V \cap U) + \mu^*(V \setminus U) \\ &\geq \mu^*(E \cap U) + \mu^*(E \setminus U).\end{aligned}$$

Take $\varepsilon \downarrow 0$, then (5.4), hence (ii), is proven.

Thus from Cathéodory, we have a outer regular Borel measure μ which satisfies (5.2) and agrees with $\mu^*|_{\mathcal{B}(X)}$.

Proof of (iii). Let $K \subseteq X$ be compact, and $f \in C_c(X, [0,1])$ with $f \geq 1_K$. Given $\varepsilon \in (0,1)$, define

$$U := \{x \in X : f(x) > 1 - \varepsilon\} \supseteq K$$

which is open, and if given a $g \in C_c(X, [0,1])$ with $\operatorname{supp}(g) \subseteq U$, one has $(1-\varepsilon)^{-1} f - g \geq 0$, hence $I(g) \leq (1-\varepsilon)^{-1} I(f)$. That means $\mu(U) \leq (1-\varepsilon)^{-1} I(f)$ by (5.2), hence

$$\mu(K) \leq \mu(U) \leq (1-\varepsilon)^{-1} I(f).$$

Take $\varepsilon \downarrow 0$, we get that $\mu(K) \leq I(f)$.

Given a $\varepsilon > 0$, by outer regularity of μ, there is an open $U \supseteq K$ such that $\mu(U) - \varepsilon < \mu(K)$, and by Lemma 5.1.2, there is a $f \in C_c(X, [0,1])$ such that $1_K \leq f \leq 1_U$, so one has $I(f) \leq \mu(U)$. Hence

$$I(f) \leq \mu(U) < \mu(K) + \varepsilon,$$

and take $\varepsilon \downarrow 0$. Thus it follows (5.3) is also satisfied.

So μ is now also weakly inner regular, hence a Radon measure.

Proof of (iv). Note that $C_c(X)$ is a linear span of $C_c(X, [0,1])$, so it suffices to show that (5.1) holds for $C_c(X, [0,1])$. Now given $f \in C_c(X, [0,1])$ and $N \in \mathbb{N}$. Let

$$K_0 := \operatorname{supp}(f) \quad \text{and} \quad K_i := \left\{x \in X : f(x) \geq \frac{i}{N}\right\},$$

and

$$f_i := \min\left\{\max\left\{f - \frac{i-1}{N}, 0\right\}, \frac{1}{N}\right\} \in C_c(X)$$

for all $i \in \{1, \dots, N\}$. Note that $K_n \subseteq \dots \subseteq K_0$. Let $i \in \{1, \dots, N\}$, then

$$f_i|_{X \setminus K_{i-1}} = 0, \quad f_i|_{K_{i-1} \setminus K_i} = f - \frac{i-1}{N}, \quad \text{and} \quad f_i|_{K_i} = \frac{1}{N}.$$

Thus

$$\frac{1}{N} 1_{K_i} \leq f_i \leq \frac{1}{N} 1_{K_{i-1}},$$

hence

$$\frac{1}{N}\mu(K_i) \leq \int_X f_i \, \mathrm{d}\mu \leq \frac{1}{N}\mu(K_{i-1}).$$

Also, as $1_{K_i} \leq Nf_i$, then by (5.3), one has $\mu(K_i) \leq NI(f_i)$. For any open $U \supseteq K_{i-1}$, one has $\operatorname{supp}(Nf_i) \subseteq U$, so by (5.2), one has $NI(f_i) \leq \mu(U)$. Using outer regularity, we get that $NI(f_i) \leq \mu(K_{i-1})$, hence

$$\frac{1}{N}\mu(K_i) \leq I(f_i) \leq \frac{1}{N}\mu(K_{i-1}).$$

Let $x \in K_{i-1} \setminus K_i$, then

$$\begin{aligned} &f_1(x) + \ldots + f_{i-1}(x) + f_i(x) + f_{i+1}(x) + \ldots + f_n(x) \\ &= (i-1)\frac{1}{N} + \left(f - \frac{i-1}{N}\right) + 0 = f \end{aligned}$$

and if $x \in X \setminus K_0$, then one has

$$f_1(x) + \ldots + f_n(x) = 0 = f(x).$$

As $X = X \setminus K_0 \cup \bigcup_{i \leq N}(K_{i-1} \setminus K_i)$, it follows that $\sum_{i \leq N} f_i = f$. So one has

$$\frac{1}{N}\sum_{i=1}^{N}\mu(K_i) \leq \int_X f \, \mathrm{d}\mu \leq \frac{1}{N}\sum_{i=0}^{N-1}\mu(K_i),$$

and

$$\frac{1}{N}\sum_{i=1}^{N}\mu(K_i) \leq I(f) \leq \frac{1}{N}\sum_{i=0}^{N-1}\mu(K_i),$$

hence

$$\left|\int_X f \, \mathrm{d}\mu - I(f)\right| \leq \frac{\mu(K_0) - \mu(K_n)}{N} \leq \frac{\mu(K_0)}{N}.$$

As $\mu(K_0) < \infty$, then (iv) is proven by taking $N \to \infty$.
The proof is thus completed. ■

Alongside with the representation theorem, we get a free result, which tells us that Radon measures on LCH spaces can also be approximated by integrals of compactly supported continuous functions.

Corollary 5. *Given a Radon measure μ on a LCH space X. Then*

$$\mu(U) = \sup\left\{\int_X f \, \mathrm{d}\mu : f \in C_c(X, [0,1]),\ \operatorname{supp}(f) \subseteq U\right\} \text{ for all open } U \subseteq X,$$

and

$$\mu(K) = \inf\left\{\int_X f \, \mathrm{d}\mu : f \in C_c(X, [0,1]), 1_K \leq f\right\} \text{ for all compact } K \subseteq X.$$

Proof. As μ is finite on compact sets, one has $C_c(X) \subseteq L^1(\mu)$, hence one has a positive linear functional $I : C_c(X) \to \mathbb{R} : f \mapsto \int_X f \, \mathrm{d}\mu$. Then the rest follows from Riesz-Markov-Kakutani Representation Theorem 5.1.4 by using existence and uniqueness. ■

2 – Regularity and Approximations

We shall further develop the theory of Radon measures here. First we present some relevant regularity results.

Lemma 1. *Let μ be a Borel measure on a Hausdorff space X. If μ is inner regular, then μ is outer regular.*

PROOF. Let $B \in \mathcal{X}$ and $\varepsilon > 0$. Then there is a compact $K \subseteq X \setminus B =: B^c$ such that

$$\mu((X \setminus B) \setminus K) = \mu(B^c \cap K^c) < \varepsilon.$$

Now $U = K^c := X \setminus K \supseteq B$ is open, and one has

$$\mu(U \setminus B) = \mu(K^c \cap B^c) < \varepsilon,$$

thus μ is outer regular. ■

Lemma 2. *Let μ be a Radon measure on a LCH space X. Then μ is inner regular on all σ-finite Borel subsets of X.*

PROOF. Let $B \in \mathcal{B}(X)$ be σ-finite. Suppose $\mu(B) < \infty$.
Let $\varepsilon > 0$, so there is an open $U \supseteq B$ such that $\mu(U) - \varepsilon < \mu(B)$. Now there is a compact $K \subseteq U$ such that $\mu(K) + \varepsilon > \mu(U)$. As $\mu(U \setminus B) < \varepsilon$, there is an open $V \supseteq U \setminus B$ such that $\mu(V) < \varepsilon$. Take $F = K \setminus V$, which is compact and $F \subseteq K \setminus (U \setminus B) \subseteq B$, and one has

$$\mu(B \setminus F) = \mu(B \setminus (K \setminus V)) \leq \mu(B \setminus K) + \mu(B \cap V) \leq \mu(U \setminus K) + \mu(V) < 2\varepsilon.$$

So it follows that μ is inner regular on B.
Suppose $\mu(B) = \infty$, so there is an increasing sequence $(B_n)_{n\in\mathbb{N}} \in \mathcal{B}(X)^{\mathbb{N}}$ of finite measure such that $B = \bigcup_{n\in\mathbb{N}} B_n$. Note that by monotone convergence theorem, one has

$$\mu(B) = \lim_{n\to\infty} \int_X 1_{B_n} \,\mathrm{d}\mu = \lim_{n\to\infty} \mu(B_n)$$

as $1_{B_n} \to 1_B$ pointwise. So for each $N \in \mathbb{N}$, there is a $n \in \mathbb{N}$ such that $\mu(B_n) > 2N$, and by inner regularity on B_n, there is a compact $K \subseteq B_n \subseteq B$ such that $\mu(K) > N$. So it follows that μ is inner regular on B, as required. ■

Corollary 3. *A σ-finite Radon measure on a LCH space is a regular measure or a Radon measure on a σ-compact LCH space is regular.*

PROOF. Immediate from the preceding lemma. ■

Proposition 4. *Let μ be a outer regular σ-finite Borel measure on a topological space X. Let $B \in \mathcal{B}(X)$ and $\varepsilon > 0$. Then there is an open $U \supseteq B$ and closed $F \subseteq B$ such that $\mu(U \setminus F) < \varepsilon$.*

PROOF. Let $B = \bigcup_{n\in\mathbb{N}} B_n$ be written as a disjoint union of μ finite sets. For each $n \in \mathbb{N}$, there is an open $U_n \supseteq B_n$ such that $\mu(U_n \setminus B_n) < \varepsilon 2^{-n-1}$. Now $U = \bigcup_{n\in\mathbb{N}} U_n \supseteq B$ is open and one has

$$\mu(U \setminus B) = \mu\left(\bigcap_{n\in\mathbb{N}} U \setminus B_n\right) = \mu\left(\bigcap_{n\in\mathbb{N}} \bigcup_{m\in\mathbb{N}} U_m \setminus B_n\right) \leq \mu\left(\bigcup_{n\in\mathbb{N}} U_n \setminus B_n\right) \leq \frac{\varepsilon}{2}.$$

Now apply the same steps as above on $X \setminus B$, one has an open $V \supseteq X \setminus B$ such that $\mu(V \setminus (X \setminus B)) \leq \varepsilon/2$. Now $F = X \setminus V \subseteq B$ is closed, and noting that

$$B \setminus F = B \cap V \subseteq V \setminus (X \setminus B),$$

one has

$$\mu(U \setminus F) = \mu(U \setminus B) + \mu(B \setminus F) \leq \varepsilon,$$

as required. ■

The following results also gives us a sufficient condition for Borel measures to be Radon. They will be the main result we will use to show Radon-ness of Borel measures.

Theorem 5. *Let X be a LCH space such that every open set is σ-compact. Then any Borel measure on X that is locally finite sets is regular, hence Radon.*

PROOF. As μ is also finite on compact sets, one has $C_c(X) \subseteq L^1(\mu)$. Then the map $f \mapsto \int_X f \, d\mu$ is a positive linear functional on $C_c(X)$, and so by the Riesz-Markov-Kakutani Representation Theorem 5.1.4, there is a unique Radon measure ν such that $\int_X f \, d\mu = \int_X f \, d\nu$ for all $f \in C_c(X)$.
Let $U \subseteq X$ be open, then $U = \bigcup_{n \in \mathbb{N}} K_n$ for some increasing sequence of compact sets $(K_n)_{n \in \mathbb{N}}$. Now by Lemma 5.1.2, there is a $f_1 \in C_c(X, [0,1])$ such that $f_1|_{K_1} = 1$ and $\operatorname{supp}(f_1) \subseteq U$, and for each $n \in \mathbb{N}$, there is a $f_n \in C_c(X, [0,1])$ such that

$$f_n|_{K_n \cup \bigcup_{i<n} \operatorname{supp}(f_i)} = 1 \quad \text{and} \quad \operatorname{supp}(f_n) \subseteq U.$$

Now one has an increasing sequence $(f_n)_{n \in \mathbb{N}}$ that converges to 1_U pointwise, so by monotone convergence theorem, one has

$$\mu(U) = \lim_{n \to \infty} \int_X f_n \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\nu = \nu(U).$$

Thus $\mu = \nu$ on open sets.
Let $B \in \mathcal{B}(X)$, and $\varepsilon > 0$. Now there is an open $U \supseteq B$ and closed $F \subseteq B$ such that $\nu(U \setminus F) < \varepsilon$. Now $U \setminus F$ is open, so $\nu(U \setminus F) = \mu(U \setminus F)$, so one has $\mu(U \setminus B) \leq \mu(U \setminus F) < \varepsilon$, so μ is outer regular.
Now F is σ-compact, so there is an increasing sequence of compact sets $(K_n)_{n \in \mathbb{N}}$ as subsets of F such that $\mu(K_n) \to \mu(F)$. Thus there is a compact $K \subseteq F \subseteq B$ such that $\mu(F \setminus K) < \varepsilon$, hence $\mu(B \setminus K) \leq \mu(B \setminus F) + \mu(F \setminus K) < 2\varepsilon$. Hence μ is inner regular. Thus μ is a regular measure, as required. ■

Corollary 6. *A locally finite Borel measure on a second-countable LCH space is Radon and regular.*

PROOF. It suffices to show that given a second-countable LCH space X, every open $U \subseteq X$ is σ-compact.
Fix an open $U \subseteq X$, and let $\mathcal{B} = \{B_n\}_{n \in \mathbb{N}}$ be a countable basis for X. Now by LCH, for each $x \in U$, there is an open $V_x \in \mathcal{N}[x]$ such that $\overline{V_x} \subseteq U$ is compact. As V_x is open, there is a $B_x \in \mathcal{B}$ such that $x \in B_x \subseteq V_x$. As $\overline{V_x}$ is compact, then so is $\overline{B_x} \subseteq \overline{V_x} \subseteq U$. So one has

$$U = \bigcup_{x \in U} \overline{B_x},$$

which is a countable union of compact sets as $\mathcal{B}$ is countable, as required. ■

Given a Radon measure μ on a LCH space X. The following result shows that $L^p(\mu)$ functions can be approximated by functions in $C_c(X)$.

Proposition 7. *Given a Radon measure μ on a LCH space X, then $C_c(X)$ is dense in $L_p(\mu)$ for all $p \in [1, \infty)$.*

PROOF. As the space of simple L_p-functions is dense in $L_p(\mu)$, then it follows that it suffices to show that 1_B can be approximated by functions in $C_c(X)$ for all μ-finite Borel $B \subseteq X$. So given a μ-finite Borel $B \subseteq X$ and $\varepsilon > 0$, there is an open $U \supseteq B$ and compact $K \subseteq B$ such that $\mu(U \setminus K) < \varepsilon^p$. Now there is a $f \in C_c(X)$ such that $1_K \leq f \leq 1_U$. So one has

$$\|f - 1_B\|_p \leq \mu(U \setminus K)^{\frac{1}{p}} < \varepsilon$$

which shows it. ∎

Finally we present one more method of constructing Radon measures from preexisting ones through density functions.

Lemma 8. *Let μ be a Radon measure on a LCH space X and $f \in L^1$. Let $\varepsilon > 0$, then there is a $\delta > 0$ such that $\mu(B) < \delta$ implies $\int_B |f| \, \mathrm{d}\mu < \varepsilon$ for all Borel $B \subseteq X$.*

PROOF. As $C_c(X)$ is dense in L^1, there is a $\varphi \in C_c(X)$ such that $\|f - \varphi\|_1 < \frac{\varepsilon}{2}$. Now if $\varphi = 0$, then the result is trivial, so suppose $\varphi \neq 0$ and let $\delta = \frac{\varepsilon}{2\|\varphi\|_\infty} > 0$. So whenever a Borel $B \subseteq X$ satisfies $\mu(B) < \delta$, one has

$$\int_B |f| \, \mathrm{d}\mu \leq \int_B |f - \varphi| \, \mathrm{d}\mu + \int_B |\varphi| \, \mathrm{d}\mu \leq \|f - \varphi\|_1 + \|\varphi\|_\infty \mu(B) < \varepsilon,$$

as required. ∎

Given a measure μ on some measurable space X. Given a measurable function $f : X \to [0, \infty]$, we denote the measure $f \, \mathrm{d}\mu$ to be the measure:

$$E \mapsto \int_E f \, \mathrm{d}\mu \quad \text{for all measurable } E \subseteq X.$$

Theorem 9. *Let X be a LCH space, μ be a Radon measure on X, and $f : X \to [0, \infty]$ be a L^1-function. Then $f \, \mathrm{d}\mu$ is a Radon regular measure on X.*

PROOF. Define $\mathrm{d}\nu = f \, \mathrm{d}\mu$, which is a finite measure as f is L^1.
Let $B \in \mathcal{B}(X)$ with finite μ measure. Let $\varepsilon > 0$, then there is a $\delta > 0$ such that if $\mu(E) < \delta$, then $\nu(E) < \varepsilon$ for all Borel $E \subseteq X$ by Lemma 5.2.8. So as μ is inner regular on σ-finite Borel sets, then there is a compact $K \subseteq B$ such that $\mu(B \setminus K) < \delta$, so one has $\nu(B \setminus K) < \varepsilon$, hence ν is inner regular on B.
Suppose $\mu(B) = \infty$. Now $E := \{x \in X : f(x) \neq 0\} = \bigcup_{n \in \mathbb{N}} E_n$ where $E_n = \left\{x \in X : f(x) > \frac{1}{n}\right\}$ for each $n \in \mathbb{N}$. Note that $(E_n)_{n \in \mathbb{N}}$ is an increasing sequence and

$$\frac{1}{n} \mu(E_n) \leq \int_X f \, \mathrm{d}\mu < \infty$$

shows that $\mu(E_n) < \infty$ for all $n \in \mathbb{N}$.

So $\mu(B \cap E_n) < \infty$ and hence there is a compact $K_n \subseteq B \cap E_n$ such that $\nu(B \cap E_n) - \nu(K_n) < 1/n$ for all $n \in \mathbb{N}$. Thus taking $n \to \infty$, we get that $\lim_{n\to\infty} \nu(K_n) = \lim_{n\to\infty} \nu(B \cap E_n) = \nu(B)$. Hence ν is inner regular.

Outer regularity follows from Lemma 5.2.1 and the rest follows. ■

3 – Pushforward Measures

A natural question is if one can make more measures out of existing measures on new measurable spaces. The answer to that question is yes and we have a nice solution too. Given measurable spaces (X, Σ_X) and (Y, Σ_Y), where μ is a measure on (X, Σ_X). Then if one has a meaurable map $f : X \to Y$, then the map

$$f_*\mu : \Sigma_Y \to [0, \infty] : B \mapsto \mu(f^{-1}(B))$$

defines a measure on (Y, Σ_Y). This is trivially verified by the properties of preimages, especially with the fact that preimages commutes with unions. We say that $f_*\mu$ is the **pushforward measure** of μ under f.

This immediately provides a broad construction of measures through measurable maps. Alongside with that, we also have a nice integral identity:

Change of Variables Formula for Pushfoward Measure 1. *Let f be a measurable map between measurable spaces (X, Σ_X) and (Y, Σ_Y), and μ be a measure on X. A measurable function $g : Y \to \mathbb{C}$ is $L^1(f_*\mu)$ if, and only if, $g \circ f$ is $L^1(\mu)$. Furthermore,*

$$\int_Y g \,\mathrm{d}f_*\mu = \int_X g \circ f \,\mathrm{d}\mu.$$

Proof. Let $B \in \Sigma_Y$, claim that $1_B \circ f = 1_{f^{-1}(B)}$.

Indeed, given $x \in X$, we have the following equivalences

$$\begin{aligned} 1_{f^{-1}(B)}(x) = 1 &\iff x \in f^{-1}(B) \\ &\iff f(x) \in B \\ &\iff (1_B \circ f)(x) = 1. \end{aligned}$$

Now one has

$$\int_Y 1_B \,\mathrm{d}f_*\mu = \mu(f^{-1}(B)) = \int_X 1_{f^{-1}(B)} \,\mathrm{d}\mu = \int_X 1_B \circ f \,\mathrm{d}\mu.$$

Thus it follows that $\int_Y \varphi \,\mathrm{d}f_*\mu = \int_X \varphi \circ f \,\mathrm{d}\mu$ for all simple functions $\varphi : Y \to \mathbb{C}$ using linearity of integrals.

Now there is a sequence of simple functions $(\varphi_n)_{n\in\mathbb{N}}$ with $\varphi_n \to g$ μ-a.e. such that $|\varphi_n| \leq |g|$ for all $n \in \mathbb{N}$, then the result follows by dominated convergence theorem. ■

The next question is how different regularity properties are preserved through pushforward of measures. Fortunately, inner regularity properties are preserved nicely.

Proposition 2. *Let f be a continuous map between topological spaces X and Y, and μ be a measure on X. If μ is weakly inner regular, then so is $f_*\mu$. Furthermore, if μ is inner regular, then so is $f_*\mu$.*

PROOF. Let $U \subseteq Y$ be open, then $f^{-1}(U)$ is open. Define:

$$A = \{\mu(K) : K \subseteq f^{-1}(U), K \text{ compact}\} \quad \text{and} \quad B = \{f_*\mu(K) : K \subseteq U, K \text{ compact}\}.$$

Let $K \subseteq f^{-1}(U)$ be compact, then $f(K) \subseteq U$ is also compact with $K \subseteq f^{-1}(f(K)))$, so one has $\mu(K) \leq \mu(f^{-1}(f(K))) = f_*\mu(f(K))$. Thus it follows that $f_*\mu(U) = \sup(A) \leq \sup(B)$. Also given a compact $K \subseteq U$, one has $f_*\mu(K) \leq f_*\mu(U)$, so $\sup(B) \leq f_*\mu(U)$, thus one has $f_*\mu(U) = \sup(B)$, as required.
If μ were to be inner regular, then we can replace U with a Borel set and the rest follows the same as above. ■

However, to also preserve outer regularity and local finiteness, one needs to impose an extra condition on f. Given a map f between topological spaces X and Y, we say f is **proper** if $f^{-1}(K)$ is compact whenever $K \subseteq Y$ is compact.

It is clear that if a measure μ on a topological space is locally finite, then so is $f_*\mu$ given that f is proper.

To prove outer regularity of $f_*\mu$, we need the following lemma.

Lemma 3. *A continuous proper map f from a topological space X to a locally compact Hausdorff space Y is a closed map.*

PROOF. Let $C \subseteq X$ be closed and let $y \in Y \setminus f(C)$. Then there is an open $U \in \mathcal{N}_Y[y]$ such that $\overline{U}$ is compact. So $f^{-1}\left(\overline{U}\right)$ is compact and define $D = C \cap f^{-1}\left(\overline{U}\right)$ which is compact. Thus $f(D)$ is compact and hence closed, and define $V = U \setminus f(D) \in \mathcal{N}_Y[y]$.
Let $z \in V \cap f(C)$, then there is a $c \in C$ such that $f(c) = z \in U$, so $c \in f^{-1}(U)$, and hence $c \in D$. Which is a contradiction as $f(c) = z \notin f(D)$. Thus V is disjoint from $f(C)$ and so $f(C)$ is closed as required. ■

Then with this, we have that Radon and regular measures are preserved through proper continuous maps.

Proposition 4. *Let f be a continuous proper map between LCH spaces X and Y. Then for any outer regular measure μ on X, $f_*\mu$ is also a outer regular measure on Y.*

PROOF.
Let $E \in \mathcal{B}(Y)$, and note that f is a closed map by preceding lemma. Define

$$A = \{\mu(U) : U \supseteq f^{-1}(E), U \text{ open}\}, \quad B = \{f_*\mu(U) : U \supseteq E, U \text{ open}\},$$

and

$$C = \{\mu(f^{-1}(U)) : f^{-1}(U) \supseteq f^{-1}(E), U \text{ open}\}.$$

We know $f_*\mu(E) = \inf(A)$, and $f_*\mu(E) \leq \inf(B)$. Now let $U \supseteq f^{-1}(E)$ be open, then $V = Y \setminus f(X \setminus U)$ is open with

$$\begin{aligned} f^{-1}(V) &= X \setminus f^{-1}(f(X \setminus U)) \supseteq X \setminus f^{-1}(f(X \setminus f^{-1}(E))) \\ &= f^{-1}(Y \setminus f(f^{-1}(Y \setminus E))) \supseteq f^{-1}(Y \setminus (Y \setminus E)) \\ &= f^{-1}(E), \end{aligned}$$

and

$$f^{-1}(V) = X \setminus f^{-1}(f(X \setminus U)) \subseteq U$$

so $f_*\mu(V) \leq \mu(U)$. Hence $\inf(C) \leq \inf(A)$.
Given an open $U \subseteq Y$ such that $f^{-1}(U) \supseteq f^{-1}(E)$, then define $V = Y \setminus f(X \setminus f^{-1}(U))$, which is open and one has

$$f^{-1}(V) = X \setminus f^{-1}(f(f^{-1}(Y \setminus U))) \subseteq f^{-1}(U) \quad \text{and} \quad V \supseteq Y \setminus f(f^{-1}(Y \setminus E)) \supseteq E$$

so $f_*\mu(V) \leq \mu(f^{-1}(U))$, and hence $\inf(B) \leq \inf(C)$. Finally, one has

$$\inf(B) \leq \inf(C) \leq \inf(A) = f_*\mu(E) \leq \inf(B)$$

hence $f_*\mu(E) = \inf(B)$ shows that $f_*\mu$ is also outer regular. ■

Corollary 5. *Let f be a continuous proper map between LCH spaces X and Y, and μ be a measure on X. One has*

(i) If μ is Radon, then so is $f_*\mu$.
(ii) If μ is regular, then so is $f_*\mu$.

4 – Measure Supports

In probability theory, it is natural to ask where a measure attains a nonzero value (probability), which is usually given by–what is called–the support of the measure.

Given a Borel measure μ on a topological space X, then the **support** of a measure μ is defined as

$$\operatorname{supp}(\mu) := \{x \in X : \mu(N) > 0 \text{ for all open } N \in \mathcal{N}_X[x]\}.$$

It follows that the support of measures are closed, hence Borel, as

$$X \setminus \operatorname{supp}(\mu) = \bigcup_{\substack{N \text{ open} \\ \mu(N)=0}} N$$

is open. We say μ has **full support** if $\operatorname{supp}(\mu) = X$.

To give some examples, the Lebesgue measure λ on $\mathbb{R}$ has full support. While a Dirac measure δ_p on some topological space X with $p \in X$ has a support of $\{p\}$.

There are nice properties of the support if the measure is Radon in particular. This is not an important result in this paper, but it is here for the sake of familiarising ourselves with this notion of measure support.

Proposition 1. *Let μ be a Radon measure on a LCH space X.*

(i) If $E \subseteq X \setminus \operatorname{supp}(\mu)$ for some $E \in \mathcal{B}(X)$, then $\mu(E) = 0$. The converse holds if E is open.
(ii) One has $x \in \operatorname{supp}(\mu)$ if, and only if, $\int_X f \, \mathrm{d}\mu > 0$ for each $f \in C_c(X, [0,1])$ with $f(x) > 0$.
(iii) For any measurable $f : X \to \mathbb{C}$, one has

$$\int_X f \, \mathrm{d}\mu = \int_{\operatorname{supp}(\mu)} f \, \mathrm{d}\mu.$$

Proof.

(i) Let $K \subseteq \bigcup_{\substack{N \text{ open} \\ \mu(N)=0}} N$ be compact, then there are $N_1, ..., N_n \subseteq X$ open with zero measure for some $n \in \mathbb{N}$ such that $K \subseteq \bigcup_{i \leq n} N_i$. Thus $\mu(K) \leq \sum_{i \leq n} \mu(N_i) = 0$. Hence by weakly inner regularity, one has that $\mu\left(\bigcup_{\substack{N \text{ open} \\ \mu(N)=0}}\right) = 0$, as required. Hence $\mu(E) = 0$. Clearly if E open and $\mu(E)$, then $E \subseteq X \setminus \operatorname{supp}(\mu)$.

(ii) Let $x \in \operatorname{supp}(\mu)$ and let $f \in C_c(X, [0,1])$ such that $f(x) > 0$. Now by continuity of f, there is a $N \in \mathcal{N}[x]$ such that $f(y) > \frac{f(x)}{2}$ for all $y \in N$. Thus one has

$$\int_X f \,\mathrm{d}\mu \geq \frac{1}{2}\mu(N)f(x) > 0,$$

as required.

If $\int_X f \,\mathrm{d}\mu > 0$ for all $f \in C_c(X, [0,1])$ with $f(x) > 0$. Let $N \in \mathcal{N}[x]$, then there is a compact $K \in \mathcal{N}[x]$ such that $K \subseteq N$. Thus there is a $f \in C_c(X, [0,1])$ such that $f|_K = 1$ and $\operatorname{supp}(f) \subseteq N$, so in particular $f(x) = 1 > 0$. Now one has

$$\mu(N) \geq \int_X f \,\mathrm{d}\mu > 0,$$

as required. So $x \in \operatorname{supp}(\mu)$.

(iii) This is clear from (i) as $\mu(X \setminus \operatorname{supp}(\mu)) = 0$. ■

There is also a nice separation regarding Borel measures with full support on LCH spaces, on its space of continuous functions.

Lemma 2. *Given a Borel measure μ with full support on a LCH space X, and let $f, g \in C(X)$ such that*

$$\int_X f\varphi \,\mathrm{d}\mu = \int_X g\varphi \,\mathrm{d}\mu$$

for all $\varphi \in C_c(X, [0,1])$. Then $f = g$.

PROOF. It suffices to show that if $\int_X f\varphi \,\mathrm{d}\mu = 0$ for all $\varphi \in C_c(X)$ implies $f = 0$.
It suffices to show that it holds for all real-valued $f \in C(X)$ as if it does, then for $f = \mathfrak{R}(f) + i\mathfrak{I}(f) \in C(X)$, one has

$$\int_X f\varphi \,\mathrm{d}\mu = \int_X \mathfrak{R}(f)\varphi \,\mathrm{d}\mu + i\int_X \mathfrak{I}(f)\varphi \,\mathrm{d}\mu = 0$$

implies $\mathfrak{R}(f) = \mathfrak{I}(f) = 0$.
So let $f \in C(X, \mathbb{R})$ and suppose $f \neq 0$. Then there is a $\xi \in X$ such that, without loss of generality, $f(\xi) > 0$, and by continuity, there is an open $U \subseteq \mathcal{N}[\xi]$ such that $f(x) > \frac{f(\xi)}{2}$ for all $x \in U$. By LCH, there is a compact $K \in \mathcal{N}[\xi]$ such that $K \subseteq U$, so there is a $\varphi \in C_c(X, [0,1])$ such that $1_K \leq \varphi \leq 1_U$. So one has

$$\int_X f\varphi \,\mathrm{d}\mu = \int_U f\varphi \,\mathrm{d}\mu > \frac{f(\xi)}{2}\int_U \varphi \,\mathrm{d}\mu \geq \frac{f(\xi)}{2}\mu(K) > 0$$

which is a contradiction. So $f = 0$, as required. ■

Now consider the probability measures

$$\mathrm{d}\mu_1 = 1_{[0,1)} \,\mathrm{d}\lambda \quad \text{and} \quad \mathrm{d}\mu_2 = \frac{1}{2}1_{(0,2]} \,\mathrm{d}\lambda$$

which immediately tells us that $\operatorname{supp}(\mu_1) = [0,1]$ and $\operatorname{supp}(\mu_2) = [0,2]$. One can ask what will be the probability distribution of the sum of $x + y$ where x and y is uniformly picked from $[0,1)$ and $(0,2]$. To calculate this distribution, we take the convolution of $1_{[0,1)}$ and $\frac{1}{2}1_{(0,2]}$:

$$f(x) := \left(1_{[0,1)} * \frac{1}{2}1_{(0,2]}\right)(x) = \frac{1}{2}\int_{\mathbb{R}} 1_{[0,1)}(x-y)1_{(0,2]}(y)\,\mathrm{d}\lambda(y)$$

$$= \frac{1}{2}\int_{\mathbb{R}} 1_{[x-1,x)\cap(0,2]}\,\mathrm{d}\lambda = \frac{1}{2}\begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leq x < 1 \\ 1 & \text{if } 1 \leq x < 2. \\ 3 - x & \text{if } 2 \leq x < 3 \\ 0 & \text{if } x \geq 3 \end{cases}$$

Hence the resulting probability distribution is given by $\mathrm{d}\mu_3 = f\,\mathrm{d}\lambda$, and one has $\operatorname{supp}(\mu_3) = [0,3]$, or equivalently

$$\operatorname{supp}(\mu_3) = \operatorname{supp}(\mu_1) + \operatorname{supp}(\mu_2).$$

Another observation is that $\operatorname{supp}(\mu_1) = \operatorname{supp}\left(1_{[0,1)}\right)$ etc., i.e. the support of those measures is equivalent to the support of their density functions. Which is in fact not a coincidence.

First, we need a definition. Given A a subset of a topological space X, we say $a \in A$ is a **neighbouring point** if for all $N \in \mathcal{N}[x]$, one has $A^\circ \cap N \neq \emptyset$. Then we say A is **neighbourful** if every point of A is a neighbouring point.

Proposition 3. *Let μ, λ be Borel measures on a topological space X and $f : X \to [0,\infty)$ be a Borel function. If $\mathrm{d}\mu = f\,\mathrm{d}\lambda$, then*

$$\operatorname{supp}(\mu) \subseteq \operatorname{supp}(f) \cap \operatorname{supp}(\lambda).$$

Furthermore, if:
- *f is continuous λ-almost everywhere (λ-a.e.);*
- *$\operatorname{supp}(f)$ is neighbourful;*
- *$\operatorname{supp}(f) \subseteq \operatorname{supp}(\lambda)$,*

then

$$\operatorname{supp}(\mu) = \operatorname{supp}(f).$$

PROOF. If $x \notin \operatorname{supp}(f)$, then there is an open $U \in \mathcal{N}[x]$ such that $f|_U = 0$, hence $\mu(U) = \int_U \mathrm{d}\lambda = 0$, so $x \notin \operatorname{supp}(\mu)$.
If $x \notin \operatorname{supp}(\lambda)$, there is an open $U \in \mathcal{N}[x]$ such that $U \subseteq X \setminus \operatorname{supp}(\lambda)$, and by Proposition 5.4.1, one has $\mu(U) = 0$, and thus $x \notin \operatorname{supp}(\mu)$.
So one has $\operatorname{supp}(\mu) \subseteq \operatorname{supp}(f) \cap \operatorname{supp}(\lambda)$. Now suppose f is continuous λ-a.e., $\operatorname{supp}(f)$ is neighbourful, and $\operatorname{supp}(f) \subseteq \operatorname{supp}(\lambda)$.
Suppose $x \in \operatorname{supp}(f)$, and let $U \in \mathcal{N}[x]$ be open. Then there is a $\xi \in A^\circ \cap U$, so there is an open $V \in \mathcal{N}[\xi]$ such that $V \subseteq \operatorname{supp}(f) \cap U$. Now as f is continuous λ-a.e. and $\lambda(V) > 0$ as $\operatorname{supp}(f) \subseteq \operatorname{supp}(\lambda)$, there must exist a $\zeta \in V$ such that f is continuous at ζ. As $f(\zeta) > 0$, then by continuity, there is an open $W \in \mathcal{N}[\zeta]$ with $W \subseteq V$ such that $f(y) > \frac{f(\zeta)}{2}$ for all $y \in W$. Finally, one has

$$\mu(U) = \int_U f\,\mathrm{d}\lambda \geq \int_W f\,\mathrm{d}\lambda > \frac{1}{2}\lambda(W)f(\zeta) > 0.$$

Hence $x \in \operatorname{supp}(\mu)$. Thus $\operatorname{supp}(\mu) = \operatorname{supp}(f)$, as required. ■

We see that f being continuous λ-a.e. with a neighbourful support is really quite important here. Consider λ being the usual Lebesgue measure on $\mathbb{R}$, and take $f = 1_{\mathbb{Q}}$. Then $\mu = 0$, but $\operatorname{supp}(f) = \operatorname{supp}(\lambda) = \mathbb{R}$. Which fails because f is not continuous λ-a.e.

Similarly, if we take $f = 1_{\{0\}}$, then $\mu = 0$ but $\operatorname{supp}(f) = \{0\}$ here. Which fails because f does not have a neighbourful support. In fact, even if the support has no isolated points or have positive λ measure, the condition will still fail with the counterexample:

$$f = 0 \quad \text{except } f = 1 \text{ on either } \left\{ \frac{1}{n} : n \in \mathbb{N} \right\} \text{ or } [2, 3].$$

This will give $\operatorname{supp}(f) = \{0\} \cup \{1/n : n \in \mathbb{N}\} \cup [2, 3]$, which has no isolated points with positive λ measure, but $\operatorname{supp}(\mu) = [2, 3]$ here.

This result shows that maps with neighbourful supports are not hard to find.

Lemma 4. *If $f : X \to \mathbb{C}$ is continuous where X is some topological space. Then $\operatorname{supp}(f)$ is neighbourful.*

PROOF. Let $x \in \operatorname{supp}(f)$, and $N \in \mathcal{N}[x]$. So there is a $y \in N$ such that $f(y) \neq 0$. Then by continuity, there is an open $U \in \mathcal{N}[y]$ with $U \subseteq N$ such that $f(z) \neq 0$ for all $z \in U$. Hence $y \in \operatorname{supp}(f)^\circ \cap N$, as required. ■

Corollary 5. *Let μ, λ be Borel measures on a topological space and $f : X \to [0, \infty)$ be a measurable function such that $\mathrm{d}\mu = f \, \mathrm{d}\lambda$. If f is continuous and $\operatorname{supp}(f) \subseteq \operatorname{supp}(\lambda)$, then*

$$\operatorname{supp}(\mu) = \operatorname{supp}(f).$$

PROOF. By the preceding lemma and Proposition 5.4.3. ■

6 | Topological Groups

In this section, we shall briefly discuss relevant properties of topological groups that will serve as a prerequisite for the next chapter.

Given a group $(G, \cdot)$ and a topology τ on G. We say τ is a **group topology** if the following group operations:

$$G \times G \to G : (x, y) \mapsto x \cdot y \quad \text{and} \quad G \to G : x \mapsto x^{-1}$$

are continuous. Thus we call the triple $(G, \cdot, \tau)$, usually just G if the context is clear, a **topological group**.

Some common topological groups include:

(i) The real numbers $(\mathbb{R}, +)$.
(ii) The positive numbers with multiplication $((0, \infty), \cdot)$.
(iii) The group of matrices under addition $(\mathbb{C}^{n \times n}, +)$.
(iv) The general linear matrix group under multiplication $(\mathrm{GL}(n, \mathbb{C}), \cdot)$.
(v) The unitary group $\mathcal{U}(n)$.

When it comes to groups, we can define group operation to sets. Given $A, B \subseteq G$, and $g \in G$, define:

$$\begin{aligned} AB &:= \{ab : a \in A, b \in B\} \\ gA &:= \{g\}A \quad \text{and} \quad Ag := A\{g\} \\ A^{-1} &:= \{a^{-1} : a \in A\}. \end{aligned}$$

Note that the group operation is associative, so we can chain these operations to yield something like gAB or AgB, which makes sense in a natural way.

A crucial note to keep in mind is that if $e \in A$, then $B \subseteq AB$.

Since that the map $x \mapsto gx$, i.e. left translation, and similarly right translations, are homeomorphisms of G onto G, and in particular, the sets gA and Ag will preserve the topological structure of A. Similarly, the map $x \mapsto x^{-1}$ is a homeomorphism, so A^{-1} will also inherit its topological properties from A.

We say that A is **symmetric** if $A = A^{-1}$.

As the group product is a continuous map, and since AB is an image of the group product under $A \times B$, then it follows that AB is compact, respectively connected, if A and B are compact, respectively connected.

We shall see that group topologies has a nice structure which will be exploited throughout this paper.

Proposition 1. *Let G be a topological group. Then the collection*

$$\{xU : x \in G, U \in \mathcal{N}[e] \text{ open and symmetric}\}$$

forms a basis for the topology of G. Also for all $U \in \mathcal{N}[e]$ and $n \in \mathbb{N}$, there is an open symmetric $V \in \mathcal{N}[e]$ such that $VV \subseteq U$.

Proof. Let $\mathcal{B}$ denote the collection in the statement. Then clearly $\mathcal{B}$ is a collection of open sets, so it suffices to prove that for any open $U \subseteq G$ and $x \in U$, there is a $V \in \mathcal{B}$ such that $x \in V \subseteq U$. Indeed, fix an open $U \subseteq G$ and $x \in U$. Then $x^{-1}U \in \mathcal{N}[e]$, and take

$$V = x^{-1}U \cap (x^{-1}U)^{-1} \in \mathcal{N}[e]$$

which is symmetric. And clearly $xV \in \mathcal{B}$ with $x \in xV \subseteq U$, as required. So $\mathcal{B}$ is basis.
By continuity of the group product at $e \in U$, then there are open $V_1, V_2 \in \mathcal{N}[e]$ such that $V_1V_2 \subseteq U$. Take $V = (V_1 \cap V_2) \cap (V_1 \cap V_2)^{-1}$, which gives an open symmetric neighbourhood of e such that $VV \subseteq U$. ∎

Just like in metric spaces we have notions of totally boundedness, in topological groups it is useful to consider an analogous result of precompactness[1]. Given a subset A of a topological group G, we say that A is **left precompact** if for all $U \in \mathcal{N}[e]$, there are finite $x_1, ..., x_n \in X$ such that

$$A \subseteq \bigcup_{i \leq n} x_i U$$

and similarly **right precompact**. A **precompact** set will be both left and right precompact.

Proposition 2. *Let A be a compact subset of a topological group G. Then A is precompact.*

Proof. Let $U \in \mathcal{N}[e]$. Then $\{xU : x \in A\}$ covers A as $x \in xU$ for each $x \in A$. Hence the rest follows from compactness. We obtain that A is left precompact and right precompactness follows similarly. ∎

[1]Some authors will use the term totally bounded in place of precompact.

7 | Haar Measures

This section will introduce measure theoretic notions to topological groups, hence allowing us to imbed notions of probability theory to the study of the Horn problem.

Given a topological group G, a Radon measure μ on G is a **left Haar measure** if $\mu \neq 0$ and $\mu(gB) = \mu(B)$ for all $g \in G$ and Borel $B \subseteq G$, we also say that μ is **left invariant** (for the latter property). Similarly for **right Haar measure** and **right invariant**. An example of a both left and right invariant Haar measure is the usual Lebesgue measure on $(\mathbb{R}^n, +)$.

When it comes to desired topological properties for topological groups for Haar measures, it turns out locally compactness and Hausdorff is what one should desire. With those properties, one can invoke the Haar's theorem:

Every LCH group has a left-invariant Haar measure that is unique up to some multiplicative constant.

We will develop the basic properties of Haar measures and prove Haar's theorem.

We shall consider the sets

$$C_c^+(G) := \{f \in C_c(G, \mathbb{R}) : f \geq 0, \|f\|_\infty > 0\}$$
$$C_c^{\geq 0}(G) := \{f \in C_c(G, \mathbb{R}) : f \geq 0\} = C_c^+(G) \cup \{0\}$$

as they are relevant to this section and introduce some relevant notations. Given a $f \in C_c(G)$ and $g \in G$, define $(gf)(x) = f(gx)$ and $(fg)(x) = f(xg)$ for all $x \in G$. Clearly gf, fg is continuous and one has a compact set

$$\text{supp}(gf) = g^{-1}\text{supp}(f),$$

so $gf \in C_c(G)$.

1 – Basic Properties

A natural question that arises is if a LCH group possesses a left Haar measure, then does it also possess a right Haar measure and vice-versa? The answer is yes, given a left Haar measure μ, we can define $\tilde{\mu}$ as its coresponding Haar measure given by $\tilde{\mu}(E) = \mu(E^{-1})$ for all Borel sets E. Hence $\tilde{\tilde{\mu}} = \mu$ and in fact, this gives us a correspondence between left and right Haar measures.

Proposition 1. *Let μ be a Radon measure on a LCH group G. Then μ is a left Haar measure if, and only if, $\tilde{\mu}$ is a right Haar measure.*

PROOF. Suppose μ is a left Haar measure, then given $E \in \mathcal{B}(G)$ and $g \in G$, one has

$$\nu(Eg) = \mu(g^{-1}E^{-1}) = \nu(E)$$

as required. Then the rest follows similarly. ■

For a nonzero Radon measure to be a left Haar measure, it suffices to have the measure to be left invariant on open sets.

Lemma 2. *Let μ be a nonzero Radon measure on a LCH group G. If $\mu(gU) = \mu(U)$ for all $g \in G$ and open $U \subseteq G$, then μ is a left Haar measure.*

PROOF. Let $E \in \mathcal{B}(G)$ and $g \in G$, by outer regularity, it suffices to show that $A = B$ where

$$A := \{\mu(U) : U \supseteq E, U \text{ open}\} \text{ and}$$
$$B := \{\mu(U) : U \supseteq gE, U \text{ open}\}.$$

Let $U \supseteq E$ be open, then $gU \supseteq gE$ is open, and

$$\mu(U) = \mu(gU) \in B,$$

hence $A \subseteq B$. Let $U \supseteq gE$ be open, then $g^{-1}U \supseteq E$ is open, and

$$\mu(U) = \mu(g^{-1}U) \in A,$$

hence $A = B$, as required. ■

We now provide fundamental properties of Haar measures with respect to integrals.

Proposition 3. *Let μ be a nonzero Radon measure on a locally compact group G. Then for any complex-valued $f \in L^1(G, \mu)$, one has*

$$\int_G gf \, \mathrm{d}\mu = \int_G f \, \mathrm{d}\mu$$

for all $g \in G$ if, and only if, μ is a left Haar measure.

PROOF. Suppose the former statement holds. By Corollary 5.1.5, it suffices to show that $A = B$ where

$$A := \left\{ \int_G f \, \mathrm{d}\mu : f \in C_c(G, [0,1]), \operatorname{supp}(f) \subseteq U \right\} \text{ and}$$
$$B := \left\{ \int_G f \, \mathrm{d}\mu : f \in C_c(G, [0,1]), \operatorname{supp}(f) \subseteq gU \right\}.$$

Let $f \in C_c(G, [0,1])$. If $\operatorname{supp}(f) \subseteq U$, then $\operatorname{supp}(g^{-1}f) = g\operatorname{supp}(f) \subseteq gU$. So one has

$$\int_G f \, \mathrm{d}\mu = \int_G g^{-1} f \, \mathrm{d}\mu \in B,$$

hence $A \subseteq B$.
If $\operatorname{supp}(f) \subseteq gU$, then $\operatorname{supp}(gf) = g^{-1}\operatorname{supp}(f) \subseteq U$, and one has

$$\int_G f \, \mathrm{d}\mu = \int_G gf \, \mathrm{d}\mu \in A,$$

hence $A = B$, as required. The rest follows by preceding lemma.
Suppose μ is a left Haar measure. Using linearity of integrals, approximation by simple functions, and the dominated convergence theorem, it suffices to show that the result holds for indicator functions. Indeed, given a Borel $B \subseteq G$ and $g \in G$, one has

$$\int_G 1_B(gx) \, \mathrm{d}\mu(x) = \mu(g^{-1}B) = \int_G 1_B \, \mathrm{d}\mu$$

as required. ■

Lemma 4. *Let μ be a left Haar measure on a LCH group G. If $U \subseteq G$ is nonempty open, then $\mu(U) > 0$, and if $f \in C_c^+(G)$, then $\int_G f \, \mathrm{d}\mu > 0$.*

PROOF. As $\mu \neq 0$ and weakly inner regular, then there is a compact $K \subseteq G$ such that $\mu(K) > 0$. Then as K is precompact, there are $x_1, ..., x_n \in G$ such that $K \subseteq \bigcup_{i \leq n} x_i U$ for some $n \in \mathbb{N}$. Thus

$$0 < \mu(K) \leq \sum_{i \leq n} \mu(x_i U) = n\mu(U)$$

which shows that $\mu(U) > 0$.
If $f \in C_c^+(G)$, then $U = \left\{x \in G : f(x) > \frac{1}{2}\|f\|_\infty\right\}$ is nonempty open, and hence

$$\int_G f \, \mathrm{d}\mu \geq \frac{1}{2}\|f\|_\infty \mu(U) > 0$$

as $\|f\|_\infty > 0$. ■

The preceding lemma shows that Haar measures on topological groups has a full support.

One nice property of Haar measures is that one has a characterization for a topological group to be compact.

Proposition 5. *Let G be a LCH group and μ be a left Haar measure on G. Then G is compact if, and only if, $\mu(G) < \infty$.*

PROOF. If G is compact, then since μ is Radon, then $\mu(G) < \infty$. Suppose G is not compact, and let $V \in \mathcal{N}[e]$ be compact. If G can covered by finite left translations of V, then any open cover of G admits a finite subcover for each left translation of V, which implies G is compact, so G cannot be covered by finite left translations of V. Take $x_1 \in V$ and define $x_{n+1} \in G \setminus \bigcup_{i \leq n} x_i V$ for each $n > 1$. Now there is a symmetric $U \in \mathcal{N}[e]$ such that $UU \subseteq V$, and claim that $x_n U$ and $x_m U$ are disjoint for $n > m$ ($n, m \in \mathbb{N}$). Indeed, if $x_n U \cap x_m U \neq \emptyset$, then there are $y_1, y_2 \in U$ such that $x_n y_1 = x_m y_2$, thus

$$x_n = x_m y_2 y_1^{-1} \in x_m UU \subseteq x_m V$$

a contradiction. So one has a disjoint sequence $\{x_n U\}_{n \in \mathbb{N}}$ in G, so by preceding lemma ($\mu(U) > 0$), one has

$$\mu(G) \geq \sum_{n \in \mathbb{N}} \mu(x_n U) = \infty\mu(U) = \infty.$$

Hence $\mu(G) = \infty$, as required. ■

2 – Haar's Theorem

Let $\mathbb{R}^{(G)}$ be the free $\mathbb{R}$-module generated by G, i.e. if $\alpha \in \mathbb{R}$, then $\alpha = \sum_{i \leq n} c_i g_i$ for some $c_1, ..., c_n \in \mathbb{R}$, $g_1, ..., g_n \in G$, and $n \in \mathbb{N}$, so it is clear that $\alpha f \in C_c(G)$. Given $x \in G$, we define

$$x\alpha = \sum_{i \leq n} c_i x g_i,$$

hence given $\beta = \sum_{j \leq m} d_j h_j$ for $d_1, ..., d_m \in \mathbb{R}$, $h_1, ..., h_m \in G$, and $m \in \mathbb{N}$, one can define

$$\alpha\beta = \sum_{i \leq n} c_i g_i \beta = \sum_{\substack{i \leq n \\ j \leq m}} c_i d_j g_i h_j.$$

We also define the 'valuation of α' to be:

$$[\alpha] = \sum_{i \leq n} c_i$$

hence one has $[\alpha + \beta] = [\alpha] + [\beta]$ and $[\alpha\beta] = [\alpha][\beta]$.

Our aim to prove the Haar's theorem is to use the Riesz-Markov-Kakutani Representation Theorem 5.1.4, where we want a positive linear functional on $C_c(G)$. Hence we should shall build our measures from the lens of integrals, hence given $f \in C_c^{\geq 0}(G)$ and $\varphi \in C_c^+(G)$, we can 'estimate' f by φ through how much φ 'covers' f by left translations:

$$(f : \varphi) := \inf\left\{[\alpha] : f \leq \alpha\varphi, \alpha \in \mathbb{R}^{(G)}\right\}$$

which will be deemed the **Haar covering number** of f by φ.

We will permanently fix a $f_0 \in C_c^+(G)$ in this subchapter, and we can 'normalize' the Haar covering number with respect to f_0, and we shall obtain a linear functional of the form

$$f \mapsto \frac{(f : \varphi)}{(f_0 : \varphi)}.$$

First, we shall discuss the well-definedness and relevant properties of the Haar covering number.

Lemma 1. *Given a LCH group G, $f \in C_c^{\geq 0}(G)$, and $\varphi \in C_c^+(G)$. The set*

$$\left\{[\alpha] : f \leq \alpha\varphi, \alpha \in \mathbb{R}^{(G)}\right\}$$

is nonempty. In fact, for each $\varepsilon \in (0, \|\varphi\|_\infty)$, there is a $n \in \mathbb{N}$ such that

$$0 \leq (f : \varphi) \leq \frac{n\|f\|_\infty}{\|\varphi\|_\infty - \varepsilon}.$$

Furthermore, if $f \in C_c^+(G)$, then $(f : \varphi) > 0$.

PROOF. Given $\alpha \in \mathbb{R}^{(G)}$, if $f \leq \alpha\varphi$, then $f \leq [\alpha]\|\varphi\|_\infty$. So

$$0 \leq \|f\|_\infty \leq [\alpha]\|\varphi\|_\infty.$$

As $\|\varphi\|_\infty > 0$, then $[\alpha] \geq 0$, hence $(f : \varphi) \geq 0$. If $f \in C_c^+(G)$, i.e. $\|f\|_\infty > 0$, then $[\alpha] \geq \frac{\|f\|_\infty}{\|\varphi\|_\infty}$, hence

$$(f : \varphi) \geq \frac{\|f\|_\infty}{\|\varphi\|_\infty} > 0.$$

Let $\varepsilon \in (0, \|\varphi\|_\infty)$, and the set

$$U_\varepsilon := \{x \in X : \varphi(x) > \|\varphi\|_\infty - \varepsilon\}$$

is open. As $\operatorname{supp}(f)$ is compact, then there is a $n \in \mathbb{N}$, and $x_1, ..., x_n \in G$ such that

$$\operatorname{supp}(f) \subseteq \bigcup_{i \leq n} x_i U_\varepsilon.$$

Let $c = \frac{\|f\|_\infty}{\|\varphi\|_\infty - \varepsilon} \geq 0$ and claim that

$$f \leq c \sum_{i \leq n} x_i^{-1}\varphi.$$

Let $x \notin \operatorname{supp}(f)$, then $f(x) = 0 \leq c\sum_{i \leq n} \varphi(x_i^{-1}x)$ holds. If $x \in \operatorname{supp}(f)$, then there is a $j \leq n$ such that $x_j^{-1}x \in U_\varepsilon$. Now one has

$$c\sum_{i \leq n} \varphi(x_i^{-1}x) \geq c\varphi(x_j^{-1}x) > c(\|\varphi\|_\infty - \varepsilon) = \|f\|_\infty \geq f(x)$$

as required.

So $nc \in \{[\alpha] : f \leq \alpha\varphi, \alpha \in \mathbb{R}^{(G)}\}$ implies $(f : \varphi) \leq \frac{n\|f\|_\infty}{\|\varphi\|_\infty - \varepsilon}$. ■

In the next lemma, we shall see that the Haar covering number posseses some linearity properties, and in fact left invariant.

Lemma 2. *Given a LCH group G, $f, g \in C_c^{\geq 0}(G)$, and $\varphi \in C_c^+(G)$. The following holds:*
(i) $(f : \varphi) = (xf : \varphi)$ *for all* $x \in G$*;*
(ii) $(cf : \varphi) = c(f : \varphi)$ *for all* $c \geq 0$*;*
(iii) $(f + g, \varphi) \leq (f : \varphi) + (g : \varphi)$*;*
(iv) $(f : \varphi) \leq (f : g)(g : \varphi)$ *if* $g \neq 0$ *(so* $g \in C_c^+(G)$*).*

PROOF. Let $\alpha, \beta \in \mathbb{R}^{(G)}$. In general, define $\{f : \varphi\} = \{[\alpha] : f \leq \alpha\varphi, \alpha \in \mathbb{R}^{(G)}\}$ here.

(i) Let $x \in G$ and suppose $[\alpha] \in \{f : \varphi\}$, then $f \leq \alpha\varphi$. Let $y \in G$, then

$$(xf)(y) = f(xy) \leq (\alpha\varphi)(xy) = (x\alpha\varphi)(y)$$

so $xf \leq x\alpha\varphi$ with $[\alpha] = [x\alpha] \in \{xf : \varphi\}$.
If $[\alpha] \in \{xf : \varphi\}$, then $xf \leq \alpha\varphi$. But that means $f \leq x^{-1}\alpha\varphi$, so $[\alpha] = [x^{-1}\alpha] \in \{f : \varphi\}$. Hence $\{f : \varphi\} = \{xf : \varphi\}$, and one has $(f : \varphi) = (xf : \varphi)$, as required.

(ii) If $c = 0$, then $\|cf\|_\infty = 0$, and hence by Lemma 7.2.1, one has $(cf : \varphi) = 0 = c(f : \varphi)$.
If $c > 0$, then one has the following equivalences:

$$\begin{aligned} [\alpha] \in \{cf : \varphi\} &\iff cf \leq \alpha\varphi \\ &\iff f \leq c^{-1}\alpha\varphi \\ &\iff c^{-1}[\alpha] = [c^{-1}\alpha] \in \{f : \varphi\} \\ &\iff [\alpha] \in c\{f : \varphi\}. \end{aligned}$$

Thus $\{cf : \varphi\} = c\{f : \varphi\}$, hence $c(f : \varphi) = (cf : \varphi)$, as required.

(iii) Suppose $[\alpha] \in \{f : \varphi\}$, and $[\beta] \in \{g : \varphi\}$. Thus $f \leq \alpha\varphi$ and $g \leq \beta\varphi$, so one has $f + g \leq (\alpha + \beta)\varphi$, i.e. $[\alpha] + [\beta] = [\alpha + \beta] \in \{f + g : \varphi\}$. Thus $\{f : \varphi\} + \{g : \varphi\} \subseteq \{f + g : \varphi\}$, thus $(f + g : \varphi) \leq (f : \varphi) + (g : \varphi)$.

(iv) If $g \neq 0$. Suppose $\alpha \in \{f : g\}$, and $\beta \in \{g : \varphi\}$, so $f \leq \alpha g$ and $g \leq \beta\varphi$. Now one has

$$f \leq \alpha g \leq \alpha(\beta\varphi) = (\alpha\beta)\varphi$$

hence $[\alpha][\beta] = [\alpha\beta] \in \{f : \varphi\}$. Thus $\{f : g\}\{g : \varphi\} \subseteq \{f : \varphi\}$. Thus $(f : \varphi) \leq (f : g)(g : \varphi)$ (note that the sets have nonnegative values, so the infimum commutes across the set product). ■

With this, for each $\varphi \in C_c^+(G)$, we can define a 'almost linear' functional on $C_c^{\geq 0}(G)$ as

$$I_\varphi : C_c^{\geq 0}(G) \to \mathbb{R} : f \mapsto \frac{(f : \varphi)}{(f_0 : \varphi)}.$$

Then due to Lemma 7.2.2, we immediately obtain the following properties.

Lemma 3. *Given a LCH G, and $\varphi \in C_c^+(G)$. Given $f, g \in C_c^{\geq 0}(G)$, the following holds:*
(i) $I_\varphi(xf) = I_\varphi(f)$ *for all* $x \in G$*;*
(ii) $I_\varphi(cf) = cI_\varphi(f)$ *for all* $c \geq 0$*;*
(iii) $I_\varphi(f + g) \leq I_\varphi(f) + I_\varphi(g)$*;*
(iv) If $f \neq 0$*, then* $(f_0 : f)^{-1} \leq I_\varphi(f) \leq (f : f_0)$*.*

In metric spaces, we have notions of uniform continuity. We can extend a similar result to topological groups. Albeit, we shall only use the definition that is relevant to this paper. Given a function f between a topological group G into $\mathbb{R}$, we say f is **left uniformly continuous** if for all $\varepsilon > 0$, there is a $U \in \mathcal{N}_G[e]$ such that $\|yf - f\|_\infty < \varepsilon$ for all $y \in U$. Similarly for **right uniformly continuity** we have $\|fy - f\|_\infty < \varepsilon$ instead.

Lemma 4. *Let $f \in C_c(G)$ for some topological group G. Then f is both left and right uniformly continuous*

PROOF. We shall prove left uniform continuity as right uniformly continuity is similar.
Note that $K := \operatorname{supp}(f)$ is compact. Let $\varepsilon > 0$, then as f is continuous, for each $x \in X$, the map

$$y \mapsto f(yx)$$

is continuous, so by continuity at e, there is a $U_x \in \mathcal{N}[e]$ such that

$$|f(yx) - f(x)| < \varepsilon$$

for each $y \in U_x$. Now there is a symmetric $V_x \in \mathcal{N}[e]$ such that $V_x V_x \subseteq U_x$. Now $\{V_x x : x \in K\}$ covers K, so there are $x_1, ..., x_n \in K$ such that $K \subseteq \bigcup_{i \leq n} V_{x_i} x_i$ for some $n \in \mathbb{N}$. Take $V = \bigcap_{i \leq n} V_{x_i} \in \mathcal{N}[e]$, which is also symmetric.
Let $y \in V$ and $x \in X$. If $x \in K$, then $xx_i^{-1} \in V_{x_i}$ for some $i \leq n$. Now $yxx_i^{-1} \in VV_{x_i} \subseteq U_{x_i}$. So $yx = y'x_i$ for some $y' \in U_{x_i}$. Note that $xx_i^{-1} \in U_{x_i}$. So one has

$$\begin{aligned}|f(yx) - f(x)| &\leq |f(yx) - f(x_i)| + |f(x_i) - f(x)| \\ &= |f(y'x_i) - f(x_i)| + \left|f(x_i) - f\left(xx_i^{-1}x_i\right)\right| \leq 2\varepsilon.\end{aligned}$$

If $x \notin K$, then $f(x) = 0$. If $yx \notin K$, then $|f(yx) - f(x)| = 0 < \varepsilon$, otherwise, $yxx_i^{-1} \in V_{x_i} \subseteq U_{x_i}$ for some $i \leq n$ ($yx \in K$). That means $xx_i^{-1} = y^{-1}yxx_i^{-1} \in VV_{x_i} \subseteq U_{x_i}$. So $x = y'x_i$ for some $y' \in U_{x_i}$, hence–just like above–one has

$$|f(yx) - f(x)| \leq |f(yx) - f(x_i)| + |f(x_i) - f(y'x_i)| \leq 2\varepsilon.$$

This shows that f is left uniformly continuous, as required. ■

The result that follows is one might expect of the usual uniform continuity in topological groups as inspired from metric spaces.

Corollary 5. *Let $f \in C_c(G)$ for some topological group G. Then for each $\varepsilon > 0$, there is a $U \in \mathcal{N}[e]$ such that*

$$|f(x) - f(y)| < \varepsilon$$

for all $xy^{-1} \in U$ or $yx^{-1} \in U$ where $x, y \in G$.

PROOF. By left uniform continuity, there is a $U \in \mathcal{N}[e]$, which can be chosen to be symmetric, such that $\|yf - f\|_\infty < \varepsilon$ for all $y \in U$. Let $x, y \in G$, if $yx^{-1} \in U$, then $y = y'x$ for some $y' \in U$, thus one has

$$|f(x) - f(y)| = |f(x) - f(y'x)| < \varepsilon.$$

The case for $xy^{-1} \in U$ is similar as U is symmetric. ■

Using our uniform continuity results, we can prove that our I_φ is approximately additive when $\operatorname{supp}(\varphi)$ is small enough.

Lemma 6. *Let G be a LCH group. Given $f_1, f_2 \in C_c^+(G)$ and $\varepsilon > 0$, there is a $U \in \mathcal{N}[e]$ such that*

$$I_\varphi(f_1 + f_2) \leq I_\varphi(f_1) + I_\varphi(f_2) \leq I_\varphi(f_1 + f_2) + \varepsilon$$

for all $\varphi \in C_c^+(G)$ such that $\operatorname{supp}(\varphi) < \varepsilon$.

PROOF. Note that $\operatorname{supp}(f_1 + f_2) = \operatorname{supp}(f_1) \cup \operatorname{supp}(f_2)$ is compact, so we can choose a $g \in C_c^+(G)$ such that $g|_{\operatorname{supp}(f_1+f_2)} = 1$. Let $\delta > 0$, consider the map $h = f_1 + f_2 + \delta g$ and $h_i = f_i / h$ (for $i \in \{1, 2\}$) where $h_i|_{G \setminus \operatorname{supp}(f_i)} = 0$. Now $h_i \in C_c^+(G)$, and by uniform continuity, there is a symmetric $U \in \mathcal{N}[e]$ such that

$$|h_i(x) - h_i(y)| < \delta \quad \text{for all } i \in \{1, 2\}$$

whenever $xy^{-1} \in U$ for $x, y \in G$.

Let $\varphi \in C_c^+(G)$ where $\operatorname{supp}(\varphi) \subseteq U$. Given $\alpha = \sum_{j \leq n} c_j x_j^{-1} \in \mathbb{R}^{(G)}$ (where $c_1, ..., c_n \in \mathbb{R}$, $x_1, ..., x_n \in G$, and $n \in \mathbb{N}$) such that $h \leq \alpha\varphi$. Fix a $i \in \{1, 2\}$. Then for each $x \in G$ one has

$$f_i(x) = h_i(x)h(x) \leq \sum_{j \leq n} c_j \varphi(x_j^{-1}x) h_i(x) \leq \sum_{j \leq n} c_j \varphi(x_j^{-1}x)(h_i(x_j) + \delta)$$

as $|h_i(x) - h_i(x_j)| < \delta$ if $x_j^{-1}x \in \operatorname{supp}(\varphi)$, or $\varphi(x_j^{-1}x) = 0$ if $x_j^{-1}x \notin \operatorname{supp}(\varphi)$. Thus

$$(f_i : \varphi) \leq \sum_{j \leq n} c_j (h_i(x_j) + \delta),$$

and as $h_1 + h_2 \leq 1$, one has

$$(f_1 : \varphi) + (f_2 : \varphi) \leq [\alpha](1 + 2\delta).$$

Choose α such that $[\alpha] = (h : \varphi)$ (like $\alpha = (h : \varphi)e$), then by Lemma 7.2.3(ii-iii), one has

$$I_\varphi(f_1 + f_2) \leq I_\varphi(f_1) + I_\varphi(f_2) \leq I_\varphi(h)(1 + 2\delta) \leq (1 + 2\delta)\big(I_\varphi(f_1 + f_2) + \delta I_\varphi(g)\big).$$

Thus, it suffices to choose $\delta > 0$ such that

$$2\delta I_\varphi(f_1 + f_2) + (1 + 2\delta)\delta I_\varphi(g) < \varepsilon$$

as required. ■

Finally, we present the Haar's Theorem which requires axiom of choice in the form of Tychonoff's theorem. The theorem proof is found in Folland's analysis textbook [4].

Haar's Theorem 7. *Let G be a locally compact group, then there is a left Haar measure μ on G that is unique up to some multiplicative constant.*

PROOF. **Existence**

Let $f \in C_c^+(G)$ and define the interval $X_f = \left[(f_0 : f)^{-1}, (f : f_0)\right]$, which is nonempty and compact.
Let $X = \prod_{f \in C_c^+(G)} X_f$, which is compact by Tychonoff.
It follows that for each $\varphi \in C_c^+(G)$, one has $I_\varphi \in X$.
For each $U \in \mathcal{N}[e]$, let

$$K(U) := \overline{\{I_\varphi : \operatorname{supp}(\varphi) \subseteq U, \varphi \in C_c^+(G)\}}.$$

Claim: $K\left(\bigcap_{i \leq n} U_i\right) \subseteq \bigcap_{i \leq n} K(U_i)$ for each $U_1, ..., U_n \in \mathcal{N}_G[e]$.
Indeed, if given $\varphi \in C_c^+(G)$ with $\operatorname{supp}(\varphi) \subseteq \bigcap_{i \leq n} U_i$, then that is equivalent to $\operatorname{supp}(\varphi) \subseteq U_i$ for each $i \leq n$. Hence one has

$$\left\{I_\varphi : \operatorname{supp}(\varphi) \subseteq \bigcap_{i \le n} U_i, \varphi \in C_c^+(G)\right\} = \bigcap_{i \le n} \left\{I_\varphi : \operatorname{supp}(\varphi) \subseteq U_i : \varphi \in C_c^+(G)\right\},$$

i.e.

$$K\left(\bigcap_{i \le n} U_i\right) \subseteq \bigcap_{i \le n} K(U_i),$$

as $\overline{\bigcap_i A_i} \subseteq \bigcap_i \overline{A_i}$ in general.
Hence claim proven.

Now given $U_1, ..., U_n \in \mathcal{N}_G[e]$, and as $K\left(\bigcap_{i \le n} U_i\right)$ is nonempty, then so is $\bigcap_{i \le n} K(U_i)$. Then as X is compact, one has $K := \bigcap_{U \in \mathcal{N}[e]} K(U)$ is nonempty (finite intersection property), and hence take a I_0 to be in that intersection.
Given $\varepsilon > 0$, any $U \in \mathcal{N}_G[e]$ and $f_1, ..., f_n \in C_c^+(G)$.

Claim $\star$: there is a $\varphi \in C_c^+(G)$ with $\operatorname{supp}(\varphi) \subseteq U$ such that $\left|I_0(f_i) - I_\varphi(f_i)\right| < \varepsilon$ for each $i \le n$.
Consider an open $N := \prod_{f \in C_c^+(G)} N_f \in \mathcal{N}_X[I_0]$ where

$$N_f = \begin{cases} (I_0(f) - \varepsilon, I_0(f) + \varepsilon) \cap X_f & \text{if } f \in \{f_1, ..., f_n\} \\ X_f & \text{otherwise} \end{cases}.$$

Now $N \cap K \neq \emptyset$, so for each $V \in \mathcal{N}[e]$, one has $N \cap \left\{I_\varphi : \operatorname{supp}(\varphi) \subseteq V\right\} \neq \emptyset$.
Hence there is a $\varphi \in C_c^+(G)$ with $\operatorname{supp}(\varphi) \subseteq U$, such that $I_\varphi \in N$. By definition of N, one has

$$\left|I_0(f_i) - I_\varphi(f_i)\right| < \varepsilon \quad \text{for all } i \le n.$$

Thus claim is proven.

Given $f \in C_c(G)$, let $f^+ := \frac{|f|+f}{2}$ and $f^- := \frac{|f|-f}{2}$, so $f^+, f^- \in C_c^{\ge 0}(G)$, and define

$$I : C_c(G) \to \mathbb{R} : f \mapsto I_0(f^+) - I_0(f^-).$$

Claim: I_0 is a left-invariant "linear" functional.
Let $f, g \in C_c^{\ge 0}(G)$, $\lambda \ge 0$, and $x \in G$. By $\star$, there is a $(\varphi_n)_{n \in \mathbb{N}} \in C_c^+(G)^{\mathbb{N}}$ such that $I_{\varphi_n}(F) \to I_0(F)$ for $F \in A := \{f, g, \lambda f, f + g, xf\}$. Now one has

$$I_0(xf) = \lim_{n \to \infty} I_{\varphi_n}(xf) = \lim_{n \to \infty} I_{\varphi_n}(f) = I_0(f).$$

So I_0 is left-invariant.
Similarly,

$$I_0(\lambda f) = \lim_{n \to \infty} I_{\varphi_n}(\lambda f) = \lambda \lim_{n \to \infty} I_{\varphi_n}(f) = \lambda I_0(f).$$

Finally, given $\varepsilon > 0$, there is a $U \in \mathcal{N}[e]$ such that $|I_\varphi(f) + I_\varphi(g) - I_\varphi(f+g)| < \varepsilon$ for $\varphi \in C_c^+(G)$ whenever $\operatorname{supp}(\varphi) \subseteq U$. Now there is such a φ with $\operatorname{supp}(\varphi) \subseteq U$ such that $|I_0(F) - I_\varphi(F)| < \varepsilon$ for all $F \in A$ by $\star$. Thus one has

$$\begin{aligned} |I_0(f) + I_0(g) - I_0(f+g)| &\le |I_0(f) - I_\varphi(f)| + |I_0(g) - I_\varphi(g)| \\ &\quad + |I_\varphi(f+g) - I_0(f+g)| + |I_\varphi(g) + I_\varphi(f) - I_\varphi(f+g)| \\ &\le 4\varepsilon, \end{aligned}$$

so one has $I_0(f + g) = I_0(f) + I_0(g)$.

Claim: I is left-invariant positive linear functional.
Let $f, g \in C_c(G)$, $\lambda \in \mathbb{R}$, and $x \in G$. Now one has

$$I(xf) = I_0(xf^+) - I_0(xf^-) = I_0(f^+) - I_0(f^-) = I(f)$$

shows that I is left-invariant.

If $\lambda \geq 0$, then

$$I(\lambda f) = I_0(\lambda f^+) - I_0(\lambda f^-) = \lambda(I_0(f^+) - I_0(f^-)) = \lambda I(f)$$

and if $\lambda \leq 0$, then

$$I(\lambda f) = I_0(-\lambda f^-) - I_0(-\lambda f^+) = \lambda(I_0(f^+) - I_0(f^-)) = \lambda I(f)$$

shows that I is homogeneous.

For additivity, note that

$$f^+ - f^- + g^+ - g^- = f + g = (f+g)^+ - (f+g)^-,$$

i.e.

$$(f+g)^- + f^+ + g^+ = (f+g)^+ + f^- + g^-.$$

So one has:

$$\begin{aligned} I(f+g) - I(f) - I(g) &= I_0((f+g)^+) - I_0((f+g)^-) - I_0(f^+) + I_0(f^-) - I_0(g^+) + I_0(g^-) \\ &= I_0((f+g)^+ + f^- + g^-) - I_0((f+g)^- + f^+ + g^+) = 0. \end{aligned}$$

Hence I is a linear functional.

Now for $f \in C_c^{\geq 0}(G)$, it is clear that $I(f) = I_0(f) \geq 0$. So I is a positive linear functional.

Now by Riesz-Markov-Kakutani Representation Theorem 5.1.4, there is a unique Radon measure $\mu : \mathcal{B}(G) \to [0, \infty]$ such that

$$I(f) = \int_G f \, \mathrm{d}\mu \quad \text{for all } f \in C_c(G).$$

By Proposition 7.1.3, μ is a left Haar measure, as required.

Uniqueness

Let μ and ν be two left Haar measures on G, and define

$$r_f = \frac{\int_G f \, \mathrm{d}\mu}{\int_G f \, \mathrm{d}\nu}$$

for all $f \in C_c^+(G)$, which is well-defined in $(0, \infty)$ by Lemma 7.1.4. Now if r_f is not dependent on f, then one has $\mu = r_f \nu$ by Lemma 5.1.3. So it remains to show that $f \mapsto r_f$ is a constant map.

Let $f, g \in C_c^+(G)$, and pick a compact symmetric $V_0 \in \mathcal{N}_G[e]$, thus the sets

$$A = V_0 \operatorname{supp}(f) \cup \operatorname{supp}(f) V_0 \quad \text{and} \quad B = V_0 \operatorname{supp}(g) \cup \operatorname{supp}(g) V_0$$

are compact. Now the maps

$$x \mapsto f(xy) - f(yx) \quad \text{and} \quad x \mapsto g(xy) - g(yx)$$

are supported on A and B respectively for all $y \in V_0$.

Let $\varepsilon > 0$, and by Lemma 7.2.4, it follows that there is a symmetric $V \in \mathcal{N}_G[e]$ with compact $\overline{V} \subseteq V_0$ such that

$$\sup_{x \in G} |f(xy) - f(yx)|, \quad \sup_{x \in G} |g(xy) - g(yx)| < \varepsilon$$

for all $y \in V$.

Choose $h \in C_c^+(G)$ such that $\operatorname{supp}(h) \subseteq V$ and $h(x) = h(x^{-1})$ for all $x \in G$ (e.g. take $h(x) = h_1(x) + h_1(x^{-1})$ for some $h_1 \in C_c^+(G)$ with $\operatorname{supp}(h_1) \subseteq V$). Now finally using Fubini's theorem (which can be obtained as f and h are compact supported, hence the restriction of the measures μ and ν on the those compact supports forms a finite measure), one has

$$\begin{aligned}\int h\,\mathrm{d}\nu \int f\,\mathrm{d}\mu &= \iint h(y)f(x)\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y) \\ &= \iint h(y)f(yx)\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y),\end{aligned}$$

and using the fact that $h(x) = h(x^{-1})$, one also has

$$\begin{aligned}&\int h\,\mathrm{d}\mu \int f\,\mathrm{d}\nu = \iint h(x)f(y)\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y) \\ &= \iint h(y^{-1}x)f(y)\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y) = \iint h(x^{-1}y)f(y)\,\mathrm{d}\nu(y)\,\mathrm{d}\mu(x) \\ &= \iint h(y)f(xy)\,\mathrm{d}\nu(y)\,\mathrm{d}\mu(x) = \iint h(y)f(xy)\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y).\end{aligned}$$

Finally,

$$\begin{aligned}\left|\int h\,\mathrm{d}\nu \int f\,\mathrm{d}\mu - \int h\,\mathrm{d}\mu \int f\,\mathrm{d}\nu\right| &= \left|\iint h(y)(f(xy)-f(yx))\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y)\right| \\ &\leq \int_V \int_A h(y)|(f(xy)-f(yx))|\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y) \\ &\leq \varepsilon\mu(A)\int_G h\,\mathrm{d}\nu\end{aligned}$$

and similarly,

$$\left|\int h\,\mathrm{d}\nu \int f\,\mathrm{d}\mu - \int h\,\mathrm{d}\mu \int f\,\mathrm{d}\nu\right| \leq \varepsilon\mu(B)\int_G h\,\mathrm{d}\nu.$$

Dividing these inequalities by $\int h\,\mathrm{d}\nu \int f\,\mathrm{d}\nu$ and $\int h\,\mathrm{d}\nu \int g\,\mathrm{d}\nu$ respectively, and adding them together to obtain

$$\left|\frac{\int f\,\mathrm{d}\mu}{\int f\,\mathrm{d}\nu} - \frac{\int g\,\mathrm{d}\mu}{\int g\,\mathrm{d}\nu}\right| \leq \varepsilon\left(\frac{\mu(A)}{\int f\,\mathrm{d}\nu} + \frac{\mu(B)}{\int g\,\mathrm{d}\nu}\right).$$

By taking $\varepsilon \downarrow 0$, we get $r_f = r_g$, as required. ■

3 – Unimodular Groups

A natural question to ask is when Haar measures are both left and right invariant? We shall answer it here.

Given a left Haar measure μ on a LCH group G, then $\mu_x : \mathcal{B}(G) \to [0,\infty] : E \mapsto \mu(Ex)$ is again a left Haar measure on G. Then by uniqueness, there is a $\Delta(x) \in (0,\infty)$ such that $\mu_x = \Delta(x)\mu$. If ν is another left Haar measure, then by uniqueness, there is a $c > 0$ such that $\nu = c\mu$, hence

$$\Delta(x)\nu = \Delta(x)c\mu = c\mu_x = \nu_x.$$

So $\Delta(x)$ is not dependent on μ, hence we have the **modular function**, $\Delta : G \to (0,\infty)$, on G. The modular function will provide us an insight between the relationship of left and right Haar measures. Clearly $\Delta(1) = 1$, but if $\Delta = 1$ (a constant function), then we say that G is **unimodular**.

Lemma 1. *Given a left Haar measure μ on a LCH group G. Then for any complex-valued $f \in L^1(G,\mu)$ and $g \in G$, one has*

$$\int_G fg \,\mathrm{d}\mu = \Delta(g^{-1}) \int_G f \,\mathrm{d}\mu.$$

PROOF. It suffices to show that the result holds for indicator functions and the rest will follow from linearity, approximation by simple functions, and the dominated convergence theorem. So given a $E \in \mathcal{B}(G)$, one has

$$\int_G 1_E(xg) \,\mathrm{d}\mu(x) = \mu(Eg^{-1}) = \Delta(g^{-1}) \int_G 1_E \,\mathrm{d}\mu,$$

as required. ■

This result shows us that how the modular function dictates how left and right Haar measures are related each other.

Corollary 2. *Given a LCH group G. A Haar measure is left and right invariant if, and only if, G is unimodular. In particular, Abelian LCH groups are unimodular.*

PROOF. Let μ be a left-right Haar measure on G. Choose any $f \in C_c^+(G)$ and by previous lemma, one has

$$\int_G f \,\mathrm{d}\mu = \int_G fg \,\mathrm{d}\mu = \Delta(g^{-1}) \int_G f \,\mathrm{d}\mu,$$

then $\Delta(g^{-1}) = 1$ for all $g \in G$. So $\Delta = 1$.
If $\Delta = 1$ and, without loss of generality, let μ be a left Haar measure, then by previous lemma, one has

$$\int_G fg \,\mathrm{d}\mu = \Delta(g^{-1}) \int_G f \,\mathrm{d}\mu = \int_G f \,\mathrm{d}\mu \quad \text{for all } f \in C_c(G) \text{ and } g \in G$$

then following from Proposition 7.1.3, one has that μ is a right Haar measure.
If G is Abelian, then given a left Haar measure μ on G. Choose any $f \in C_c^+(G)$ then one has

$$\int_G fg \,\mathrm{d}\mu = \Delta(g^{-1}) \int_G f \,\mathrm{d}\mu = \Delta(g^{-1}) \int_G gf \,\mathrm{d}\mu = \Delta(g^{-1}) \int_G fg \,\mathrm{d}\mu,$$

then $\Delta(g^{-1}) = 1$ as $fg \in C_c^+(G)$ for all $g \in G$. So $\Delta = 1$,as required. ■

We shall show that Δ is also a continuous (group) homomorphism. First we need another lemma.

Lemma 3. *If μ is a Radon measure on a LCH group G and $f \in C_c(G)$. Then the maps $x \mapsto \int_G xf \,\mathrm{d}\mu$ and $x \mapsto \int_G fx \,\mathrm{d}\mu$ on G are continuous.*

PROOF. We shall prove $x \mapsto \int_G xf \,\mathrm{d}\mu$ is continuous and the other follows similarly. Let $g \in G$ and $\varepsilon > 0$. As f is uniformly continuous, there is a symmetric relatively compact $U \in \mathcal{N}[e]$ such that $|f(a) - f(b)| < \varepsilon$ whenever $ab^{-1} \in U$. For each $x \in gU$, one has $|f(xy) - f(gy)| < \varepsilon$ for all $y \in G$ as $xy(gy)^{-1} = xg^{-1} \in U$. Let $K = \operatorname{supp}(xf) \cup \operatorname{supp}(gf$, which is compact and thus one has

$$\begin{aligned}\left|\int_G xf\,\mathrm{d}\mu - \int_G gf\,\mathrm{d}\mu\right| &\le \int_G |xf-gf|\,\mathrm{d}\mu \\ &= \int_M |xf-gf|\,\mathrm{d}\mu \\ &< \varepsilon\mu(M),\end{aligned}$$

where $\mu(M) < \infty$ as μ is Radon. Thus our maps are continuous. ■

Lemma 4. *The modular function Δ on a LCH group G is a continuous homorphism.*

PROOF. Let μ be any left Haar measure on G, and by Lemma-7-2-8, choose any $f \in C_c(G)$ such that $\int_G f\,\mathrm{d}\mu = 1$, then one has that

$$\Delta(g) = \int_X fg^{-1}\,\mathrm{d}\mu \quad \text{for all } g \in G.$$

Then by preceding lemma, Δ is continuous. Now choose any Borel $E \subseteq G$ such that $\mu(E) \in (0,\infty)$ (which exists as $\mu \neq 0$ and μ is Radon) and let $x, y \in G$, then one has

$$\Delta(xy)\mu(E) = \mu(Exy) = \Delta(y)\mu(Ex) = \Delta(y)\Delta(x)\mu(E),$$

so $\Delta(xy) = \Delta(x)\Delta(y)$. Hence Δ is a homorphism, as required. ■

Finally, we have the tools to show that the corresponding pairs of left and right Haar measures really just depend on Δ. That is given a left Haar measure μ, then $\tilde{\mu}(E) = \mu(E^{-1})$ for all Borel E are mutually continuous with a density function depending on Δ.

Theorem 5. *Let μ be a left Haar measure on a LCH group G. Then*

$$\mathrm{d}\tilde{\mu} = \frac{1}{\Delta}\,\mathrm{d}\mu.$$

PROOF. By Lemma-7-2-8 and preceding lemma, one has that

$$\int_G \frac{1}{\Delta} f\,\mathrm{d}\mu = \Delta(g)\int_G \frac{1}{\Delta(xg)} f(xg)\,\mathrm{d}\mu = \int_G \frac{1}{\Delta}(fg)\,\mathrm{d}\mu$$

for all $f \in C_c(G)$. Thus the positive linear functional $f \mapsto \int_G \frac{1}{\Delta} f\,\mathrm{d}\mu$ is right invariant, then by Riesz-Markov-Kakutani Representation Theorem 5.1.4, the linear functional has an associated Radon measure, $\frac{1}{\Delta}\,\mathrm{d}\mu$, and following from Proposition 7.1.3, $\frac{1}{\Delta}\,\mathrm{d}\mu$ is a right Haar measure. Hence by uniqueness, one has $\frac{1}{\Delta}\,\mathrm{d}\mu = c\,\mathrm{d}\tilde{\mu}$ for some $c > 0$.
If $c \neq 1$, then using continuity of $\frac{1}{\Delta}$ at 1, there is a symmetric $U \in \mathcal{N}[e]$ such that

$$\left|\frac{1}{\Delta(x)} - 1\right| < \frac{1}{2}|c-1| \quad \text{for all } x \in U.$$

By symmetry $\mu(U) = \tilde{\mu}(U)$, hence one has

$$|c-1|\mu(U) = |c\tilde{\mu}(U) - \mu(U)| = \left|\int_U \left(\frac{1}{\Delta} - 1\right)\mathrm{d}\mu\right| < \frac{1}{2}|c-1|\mu(U),$$

a contradiction. So $c = 1$, as required. ■

This result also shows that integrals in unimodular group satisfy another nice identity; they are invariant under group inversion.

Corollary 6. *Let G be a unimodular group and μ be a Haar measure on G. Then for any $f \in L^1$, one has*

$$\int_G f(x)\,\mathrm{d}\mu(x) = \int_G f(x^{-1})\,\mathrm{d}\mu(x).$$

PROOF. By the preceding theorem, $\mu = \tilde{\mu}$, and note that $1_B(x^{-1}) = 1_{B^{-1}}(x)$ for all $B \in \mathcal{B}(G)$ and $x \in G$. Hence given $B \in \mathcal{B}(G)$, one has,

$$\int_G 1_B(x)\,\mathrm{d}\mu(x) = \tilde{\mu}(B) = \mu(B^{-1}) = \int_G 1_{B^{-1}}(x) = \int_G 1_B(x^{-1})\,\mathrm{d}\mu(x).$$

Thus the statement holds for all simple functions. Then by approximation by simple functions and dominated convergence theorem, the result follows. ■

Hence the main takeaway here for unimodular groups is that Haar measures on such groups are both left and right invariant, and integrals are invariant under group inversion as stated in the preceding corollary.

Conveniently, it turns out all compact (Hausdorff) groups are unimodular.

Proposition 7. *Let G be a LCH group. If G is compact, then G is unimodular.*

PROOF. Let μ be a left Haar measure on G. Then for any $x \in G$, one has $G = Gx$, so

$$\mu(G) = \mu(Gx) = \Delta(x)\mu(G) \quad \text{implies} \quad \Delta(x) = 1$$

as $0 < \mu(G) < \infty$. Thus $\Delta = 1$, as required. ■

Thus by <u>Haar's Theorem 7.2.7</u>, as $\mathcal{U}(n)$ is compact, there is a unique probability Haar measure μ on $\mathcal{U}(n)$. By <u>Corollary 7.3.2</u>, it turns out μ is both left and right invariant. However, the existence of μ does rely on axiom of choice.

4 – Explicit Construction: Unitary Haar Measure

This part will be primarily focusing on the explicit construction of the unitary Haar measure without requiring the axiom of choice. The construction can also gives us an idea on how to compute using the unitary Haar measure.First we denote λ_n to be the Lebesgue measure on $\mathbb{C}^n$ (identified as $\mathbb{R}^{2n}$). Note that the Lebesgue measure is regular and Radon.

Let G be the Gaussian function on $\mathbb{C}^n$ defined as

$$G(z) = \frac{1}{\pi^n} e^{-|z|^2}.$$

This function is continuous and its integral, the Gaussian integral is well-known, with

$$\int_{\mathbb{C}^n} G\,\mathrm{d}\lambda_n = 1.$$

hence $G \in L^1(\lambda_n)$. Thus the **Gaussian measure** γ_n defined as

$$\mathrm{d}\gamma_n = G\lambda_n.$$

is a regular Radon probability measure by Theorem 5.2.9.

Since $\mathbb{C}^{n\times n}$ is identified as $\mathbb{C}^{n^2}$, one has that γ_{n^2} gives a regular Radon probability measure on $\mathbb{C}^{n\times n}$ through that natural identification. Now we have a nice probability measure on $\mathbb{C}^{n\times n}$, we shall construct the unitary Haar measure by considering the continuous map $F : \mathrm{GL}(n, \mathbb{C}) \to \mathcal{U}(n)$:

$$X \mapsto X(X^*X)^{-\frac{1}{2}}. \tag{7.1}$$

The proof of the properties of this map is found in Chapter 3.3.

Then taking the pushforward of γ_{n^2} under that map should give us our desired measure.

First, we shall show that almost all matrices in $\mathbb{C}^{n\times n}$ are invertible.

Lemma 1. *Let $p \in \mathbb{C}[x_1, ..., x_n]$, then the set $A = \{x \in \mathbb{C}^n : p(x) = 0\}$ is Borel and either $\lambda_n(A) = 0$ or $A = \mathbb{C}^n$.*

PROOF. Clearly A is Borel as p is continuous. If $p = 0$, then $A = \mathbb{C}^n$ holds, so suppose $p \neq 0$. If $n = 1$, then by fundamental theorem of algebra, one has $|A| \leq n$, so $\lambda(A) = 0$ holds. Consider strong induction $n > 1$, and let $p \in \mathbb{C}[x_1, ..., x_n]$, so there is a $k \in \mathbb{N}$ and $p_1, ..., p_k \in \mathbb{C}[x_1, ..., x_{n-1}]$ such that

$$p(x, x_n) = \sum_{i \leq k} p_i(x) x_n^i \quad \text{for all } (x, x_n) \in \mathbb{C}^n.$$

Define

$$B = \{(x, x_n) \in \mathbb{C}^n : p_i(x) = 0 \text{ for all } i \leq k\},$$

and

$$C = \{(x, x_n) \in \mathbb{C}^n : p(x, x_n) = 0 \text{ but } p_i(x) \neq 0 \text{ for some } i \leq k\},$$

so $A = B \cup C$. Now by inductive hypothesis,

$$\lambda(B) \leq \lambda(\{x \in \mathbb{C}^{n-1} : p_k(x) = 0\}) = 0.$$

For a fixed $x \in \mathbb{C}^{n-1}$ with $p_i(x) \neq 0$ for some i, there are finite $x_n \in \mathbb{C}$ such that $p(x, x_n) = 0$ (by fundamental theorem of algebra). Now one has

$$\begin{aligned}
\lambda_n(C) &= \int_{\mathbb{C}^n} 1_C \, \mathrm{d}\lambda_n \\
&= \int_{\mathbb{C}^{n-1}} \int_{\mathbb{C}} 1_{\{(x, x_n) \in \mathbb{C}^n : p(x_n, x) = 0\}} \, \mathrm{d}\lambda_1(x_n) \, \mathrm{d}\lambda_{n-1}(x) \\
&= \int_{\mathbb{C}^{n-1}} \lambda_1(\{(x, x_n) \in \mathbb{C}^n : p(x, x_n) = 0\}) \, \mathrm{d}\lambda_{n-1}(x) \\
&= \int_{\mathbb{C}^{n-1}} 0 \, \mathrm{d}\lambda_{n-1} = 0,
\end{aligned}$$

as required. ■

Corollary 2. *Almost all matrices in $\mathbb{C}^{n\times n}$ are invertible. Furthermore, $\gamma_{n^2}(\mathrm{GL}(n, \mathbb{C})) = 1$.*

PROOF. The map $A \mapsto \det(A)$ is a polynomial map (hence continuous), then by preceding lemma, there set of zeroes of that map has measure zero. Note that the set of zeroes corresponds to the set of singular matrices, hence almost all matrices in $\mathbb{C}^{n\times n}$ are invertible. Now as γ_{n^2} has a density with respect to λ_{n^2}, thus $\gamma_{n^2}(\mathrm{GL}(n, \mathbb{C})) = \gamma_{n^2}(\mathbb{C}^{n\times n}) = 1$, as required. ■

Now it turns out, γ_n is left and right invariant with respect to the unitary matrices.

Lemma 3. *For any $B \in \mathcal{B}(\mathbb{C}^{n\times n})$ and $U \in \mathcal{U}(n)$, one has*

$$\gamma_{n^2}(UB) = \gamma_{n^2}(BU) = \gamma_{n^2}(B).$$

PROOF. Note that the maps $X \mapsto UX$ and $X \mapsto XU$ are C^1-diffeomorphisms on $\mathbb{C}^{n\times n}\left(\cong \mathbb{C}^{n^2}\right)$ with derivative U. Thus for $U \in \mathcal{U}(n)$, note that $|\det(U)| = 1$, and one has

$$\gamma_{n^2}(UB) = \frac{1}{\pi^{n^2}} \int_{UB} e^{-\operatorname{tr}(X^*X)}\, \mathrm{d}\lambda_{n^2}(X) = \frac{1}{\pi^{n^2}} \int_{B} |\det(U)| e^{-\operatorname{tr}((U^*X)^*(U^*X))}\, \mathrm{d}\lambda_{n^2}(X) = \gamma_{n^2}(B),$$

and similarly $\gamma_{n^2}(BU) = \gamma_{n^2}(B)$, as required. ∎

Finally, let us define the our candidate unitary Haar measure

$$\omega : \mathcal{B}(\mathcal{U}(n)) \to [0,1] \quad \text{as } \omega = F_* \gamma_{n^2},$$

where F is described in (7.1). Now ω is indeed a Borel measure as F is continuous, and it is indeed a probability measure as γ_{n^2} is also one.

Note that $\mathcal{U}(n)$ is a second-countable compact space and ω is finite, then by Corollary 5.2.6, one has that ω is a Radon and regular measure. Thus it suffices to prove invariance on unitary matrices.

Theorem 4. *The ω defined above is the probability Haar measure on $\mathcal{U}(n)$.*

PROOF. We have already established that ω is a Radon probability measure, so it suffices to prove left invariance, as right invariance will follow from the fact that $\mathcal{U}(n)$ is compact through Proposition 7.3.7, similarly, uniqueness follows from Haar's Theorem 7.2.7 (which that part does not use axiom of choice).
Now given $U \in \mathcal{U}(n)$ and $B \in \mathcal{B}(\mathcal{U}(n))$, claim that $UF^{-1}(B) = F^{-1}(UB)$.
Let $X \in UF^{-1}(B)$, then there is a $Y \in F^{-1}(B)$ such that $X = UY$, now

$$F(X) = UY((UY)^*UY)^{-\frac{1}{2}} = UF(Y) \in UB$$

so $X \in F^{-1}(UB)$. Hence $UF^{-1}(B) \subseteq F^{-1}(UB)$.
Now $U^{-1}F^{-1}(UB) \subseteq F^{-1}(B)$ by above, hence $F^{-1}(UB) \subseteq UF^{-1}(B)$. Thus our claim is proven. Finally by preceding lemma,

$$\omega(UB) = \gamma_{n^2}(UF^{-1}(B)) = \omega(B)$$

shows that ω is a left Haar measure, as required. ∎

Thus for any $B \in \mathcal{B}(\mathcal{U}(n))$, by Change of Variables Formula for Pushfoward Measures 5.3.1, one has

$$\omega(B) = \gamma_{n^2}(F^{-1}(B)) = \frac{1}{\pi^{n^2}} \int_{F^{-1}(B)} e^{-|z|^2}\, \mathrm{d}\lambda_{n^2}(z).$$

8 | Fourier Analysis

In this section, we will briefly look at the Fourier transform of bounded measures on $\mathbb{C}^n$ and its immediate properties. Since we are building towards the probabilistic interpretation of the Horn problem, we will be taking convolution of such measures. In order to find the convolution, we will be observing their Fourier transforms.

Given a finite Borel measure μ on $\mathbb{C}^n$, we define the **Fourier transform**:

$$\hat{\mu} : \mathbb{C}^n \to \mathbb{C} : x \mapsto \int_{\mathbb{C}^n} e^{\langle \omega, x \rangle} \, \mathrm{d}\mu(\omega).$$

Which converges as μ is finite. It is clear that the Fourier transform here is acts linearly.

As an example, if $x \in \mathbb{C}^n$, then the Fourier transform of the Dirac delta measure δ_x is

$$\tilde{\delta_x}(t) = \int_{\mathbb{C}^n} e^{\langle \omega, t \rangle} \, \mathrm{d}\delta_x(\omega) = e^{\langle x, t \rangle} \quad \text{for all } t \in \mathbb{C}^n.$$

Given two σ-finite Borel measures μ and ν on a topological group G, we define the **convolution** of μ and ν, $\mu * \nu$, as the pushforward measure of $\mu \times \nu$ under the group action map. That is, if we define $h : G \times G \to G$ to be $h(x, y) = xy$, then one has the measure $\mu * \nu = h_*(\mu \times \nu)$. Thus for any $E \in \mathcal{B}(G)$, one has (using Tonelli-Fubini's theorem)

$$(\mu * \nu)(E) = \int_{G \times G} 1_{h^{-1}(E)} \, \mathrm{d}(\mu \times \nu) = \iint_{G \times G} 1_E(xy) \, \mathrm{d}\mu(x) \, \mathrm{d}\nu(y)$$

as $1_{h^{-1}(E)}(x, y) = 1_E(xy)$ for all $x, y \in G$. Thus using approximation by simple functions and the dominated convergence theorem, one has that

$$\int_G f \, \mathrm{d}(\mu * \nu) = \iint_{G \times G} f(xy) \, \mathrm{d}\mu(x) \, \mathrm{d}\nu(y)$$

for all $f \in L^1(\mu * \nu)$.

Note that if the group is Abelian, then $\mu * \nu = \nu * \mu$.

Let λ be a σ-finite left Haar measure on G where G is now LCH, such that $\mu \ll \lambda$ and $\nu \ll \lambda$, then by Radon-Nikodyn theorem, there are measurable $f, g : G \to [0, \infty)$ such that $\mathrm{d}\mu = f \, \mathrm{d}\lambda$ and $\mathrm{d}\nu = g \, \mathrm{d}\lambda$. Now we have another interpretation of the convolution:

$$\begin{aligned}
(\mu * \nu)(E) &= \int_G \int_G 1_E(xy) f(x) g(y) \, \mathrm{d}\lambda(y) \, \mathrm{d}\lambda(x) \\
&= \int_G \int_G 1_E(y) f(x) g(x^{-1}y) \, \mathrm{d}\lambda(y) \, \mathrm{d}\lambda(x) \\
&= \int_G \int_E f(x) g(x^{-1}y) \, \mathrm{d}\lambda(y) \, \mathrm{d}\lambda(x) \\
&= \int_E \int_G f(x) g(x^{-1}y) \, \mathrm{d}\lambda(x) \, \mathrm{d}\lambda(y) \\
&= \int_E (f * g) \, \mathrm{d}\lambda
\end{aligned}$$

where we define the convolution of $f * g : G \to [0, \infty)$ to be

$$(f * g)(y) = \int_G f(x)g(x^{-1}y)\,\mathrm{d}\lambda(x) \quad \text{for } y \in G.$$

We can define the convolution of two functions $f, g \in L^1(G, \lambda)$ as above, where λ is some measure on a topological group G. It is clear that $\|f * g\|_1 \leq \|f\|_1\|g\|_1$, so $f * g \in L^1(G, \lambda)$, and those notions convolutions generalizes the case in $\mathbb{C}^n$.

Here is a characterization of the commutativity of convolutions.

Theorem 1. *Let μ be a left Haar measure on a LCH group G. Then G is Abelian if, and only if, for all $f, g \in C_c(G)$, such that $f * g = g * f$. In particular, $f * g = g * f$ for all $f, g \in L^1$ if G is Abelian.*

PROOF. Suppose G is Abelian and $f, g \in L^1$. Note that G is also unimodular. Let $y \in G$, now one has

$$\begin{aligned}(f * g)(y) &= \int_G f(x)g(x^{-1}y)\,\mathrm{d}\mu(x)\\ &= \int_G f(xy)g((xy)^{-1}y)\,\mathrm{d}\mu(x)\\ &= \int_G f(xy)g(x^{-1})\,\mathrm{d}\mu(x)\\ &= \int_G f(x^{-1}y)g(x)\,\mathrm{d}\mu(x)\\ &= (g * f)(y)\end{aligned}$$

shows that $f * g = g * f$.

Suppose $*$ is commutative on $C_c(G)$. Then for any $f, g \in C_c(G)$ and $y \in G$, using Theorem 7.3.5, one has

$$\begin{aligned}0 = (f * g)(y) - (g * f)(y) &= \int_G f(x)g(x^{-1}y)\,\mathrm{d}\mu(x) - \int_G g(x)f(x^{-1}y)\,\mathrm{d}\mu(x)\\ &= \int_G f(x)g(x^{-1}y)\,\mathrm{d}\mu(x) - \int_G g(yx)f((yx)^{-1}y)\,\mathrm{d}\mu(x)\\ &= \int_G f(x)g(x^{-1}y)\,\mathrm{d}\mu(x) - \int_G g(yx)f(x^{-1})\,\mathrm{d}\mu(x)\\ &= \int_G f(x)g(x^{-1}y)\,\mathrm{d}\mu(x) - \int_G g(yx^{-1})f(x)\frac{1}{\Delta(x)}\,\mathrm{d}\mu(x)\\ &= \int_G f(x)\left(g(x^{-1}y) - \frac{g(yx^{-1})}{\Delta(x)}\right)\mathrm{d}\mu(x).\end{aligned}$$

Fixing g, we have the above holds for all $f \in C_c(G)$ and $y \in G$, so by Lemma 5.4.2, one has

$$\Delta(x)g(x^{-1}y) = g(yx^{-1})$$

for all $x, y \in G$. Choose $y = e$ here, and for each $x \in G$, choose any $g \in C_c^+(G, \mathbb{R})$ such that $g(x^{-1}) > 0$, hence it follows that $\Delta(x) = 1$, i.e. $\Delta = 1$. So for each $x, y \in G$, one has

$$g(x^{-1}y) = g(yx^{-1})$$

for all $g \in C_c(G)$. As $C_c(G)$ separates the points in G by Lemma 5.1.2, one has $x^{-1}y = yx^{-1}$. Replace x^{-1} with x and we have G is Abelian as required. ■

Here we have an analog of the usual convolution theorem.

Convolution Theorem 2. *Let μ, ν be finite Borel measures on $\mathbb{C}^n$, then*

$$\widehat{\mu * \nu} = \hat{\mu}\hat{\nu}.$$

PROOF. Through direct computation, one has

$$\begin{aligned}
(\widehat{\mu * \nu})(x) &= \int_G e^{\langle \omega, x \rangle} \,\mathrm{d}(\mu * \nu)(\omega) = \iint_{G \times G} e^{\langle \omega + \eta, x \rangle} \,\mathrm{d}\mu(\omega) \,\mathrm{d}\nu(\eta) \\
&= \iint_{G \times G} e^{\langle \omega, x \rangle} e^{\langle \eta, x \rangle} \,\mathrm{d}\mu(\omega) \,\mathrm{d}\nu(\eta) \\
&= \int_G e^{\langle \omega, x \rangle} \,\mathrm{d}\mu(\omega) \int_G e^{\langle \eta, x \rangle} \,\mathrm{d}\nu(\eta) = \hat{\mu}(x)\hat{\nu}(x)
\end{aligned}$$

for all $x \in \mathbb{C}^n$, as required. ■

In the case of $\mathcal{H}(n)$, which is a subspace of $\mathbb{C}^{n^2}$, given $X, Y \in \mathcal{H}(n)$ where $X = (x_{ij})_{i,j \leq n}$, $Y = (y_{ij})_{i,j \leq n}$ we take

$$X = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \quad Y = \begin{pmatrix} \zeta_1 & \cdots & \zeta_n \end{pmatrix} \quad \text{for } \xi_1, ..., \xi_n, \zeta_1, ..., \zeta_n \in \mathbb{C}^n,$$

then one has

$$\begin{aligned}
\operatorname{tr}(XY) &= \sum_{i \leq n} \langle \xi_i, \overline{\zeta_i} \rangle \\
&= x_{11}y_{11} + x_{12}y_{21} + ... + x_{1n}y_{n1} + \\
&\quad x_{21}y_{12} + x_{22}y_{22} + ... + x_{2n}y_{n2} + \\
&\quad ... + x_{n1}y_{1n} + ... + x_{nn}y_{nn} \\
&= x_{11}\overline{y_{11}} + x_{12}\overline{y_{12}} + ... + x_{1n}\overline{y_{1n}} + \\
&\quad x_{21}\overline{y_{21}} + x_{22}\overline{y_{22}} + ... + x_{2n}\overline{y_{2n}} + \\
&\quad ... + x_{n1}\overline{y_{n1}} + ... + x_{nn}\overline{y_{nn}} \\
&= \langle \Phi(X), \Phi(Y) \rangle
\end{aligned}$$

where $\Phi : \mathbb{C}^{n \times n} \to \mathbb{C}^{n^2}$ is the natural identification map. Hence given a finite measure μ on $\mathcal{H}(n)$, its Fourier transform is

$$\hat{\mu}(X) = \int_{\mathcal{H}(n)} \mathrm{e}^{\operatorname{tr}(YX)} \,\mathrm{d}\mu(Y) = \int_{\mathcal{H}(n)} \mathrm{e}^{\operatorname{tr}(XY)} \,\mathrm{d}\mu(Y) \quad \text{for } X \in \mathcal{H}(n).$$

9 | Probabilistic Horn Problem

In this section, we shall give a brief discussion on the probabilistic interpretation of the Horn problem. Recall that given $\alpha \in (\mathbb{R}^n)_\downarrow$, we defined

$$\mathcal{O}_\alpha = \{U \operatorname{diag}(\alpha) U^* : U \in \mathcal{U}(n)\}$$

which is a compact and path-connected set as an image of the continuous map

$$U \mapsto U \operatorname{diag}(\alpha) U^* \quad \text{for } U \in \mathcal{U}(n). \tag{9.1}$$

The question now asks what is the probability distribution of the eigenvalues of $X + Y$ as X varies over in $\mathcal{O}_\alpha$ and Y varies over $\mathcal{O}_\beta$ for some given $\alpha, \beta \in (\mathbb{R}^n)_\downarrow$. A large part of this study was done by Jacques Faraut, and we will be explaining on how he obtained his findings.

1 – Introduction

By Horn's Conjecture 4.4.1, we know that $\sigma_\downarrow(\mathcal{O}_\alpha + \mathcal{O}_\beta)$ is a convex polytope in $\mathbb{R}^n$.

Now let us always denote ω to the probability Haar measure on $\mathcal{U}(n)$, and one has the **orbital measure** μ_α on $\mathcal{H}(n)$ as the pushforward of ω under the mapping from $\mathcal{U}(n)$ to $\mathcal{H}(n)$ as in (9.1). Note that $\mu_\alpha(\mathcal{H}(n) \setminus \mathcal{O}_\alpha) = 0$, μ_α is also a probability measure, and μ_α satisfies

$$\int_{\mathcal{O}_\alpha} f \, \mathrm{d}\mu_\alpha = \int_{\mathcal{U}(n)} f(U \operatorname{diag}(\alpha) U^*) \, \mathrm{d}\omega(U) \quad \text{for all } f \in L^1(\mathcal{H}(n), \mu_\alpha)$$

by Change of Variables Formula for Pushfoward Measures 5.3.1.

Note that μ_α is also a Radon measure by Corollary 5.3.5 as $\mathcal{U}(n)$ is compact and one has

$$\begin{aligned}\int_{\mathcal{H}(n)} f(axb) \, \mathrm{d}\mu_\alpha(x) &= \int_{\mathcal{U}(n)} f(aU \operatorname{diag}(\alpha) U^* b) \, \mathrm{d}\omega(u) \\ &= \int_{\mathcal{U}(n)} f(U \operatorname{diag}(\alpha) U^*) \, \mathrm{d}\omega(u) = \int_{\mathcal{H}(n)} f \, \mathrm{d}\mu_\alpha\end{aligned}$$

for all $f \in L^1(\mathcal{H}(n))$, $a, b \in \mathcal{U}(n)$. So we say that μ_α is unitarily invariant, so in the sense that μ_α is 'similar' to the Haar measure ω.

In general, a measure μ on $\mathcal{H}(n)$ is said to be **unitarily invariant** if

$$\int_{\mathcal{H}(n)} f(UXU^*) \, \mathrm{d}\mu(X) = \int_{\mathcal{H}(n)} f \, \mathrm{d}\mu$$

for all $f \in L^1(\mathcal{H}(n), \mu)$, $U \in \mathcal{U}(n)$.

Note that we have the continuous surjection:

$$\mathbb{R}^n \times \mathcal{U}(n) \to \mathcal{H}(n) : (t, U) \mapsto U \operatorname{diag}(t) U^*.$$

And due to J. Faraut [5], following from Weyl's integration formula, we can decompose our integrals with respect to the mapping above. So given a measure μ on $\mathcal{H}(n)$ that is unitarity invariant, then for each $f \in L^1(\mathcal{H}(n), \mu)$, one has

$$\int_{\mathcal{H}(n)} f \, \mathrm{d}\mu = \int_{\mathbb{R}^n} \int_{\mathcal{U}(n)} f(U \operatorname{diag}(t) U^*) \, \mathrm{d}\omega(U) \, \mathrm{d}\nu(t)$$

where ν is a permutation invariant measure on $\mathbb{R}^n$. By permutation invariant, we mean that

$$\int_{\mathbb{R}^n} f \, \mathrm{d}\nu = \int_{\mathbb{R}^n} f(\sigma(x)) \, \mathrm{d}\nu(x) \tag{9.2}$$

for all $f \in L^1(\mathbb{R}^n, \nu)$ and $\sigma \in S_n$ (the symmetry group).

Thus for each μ_α, there is a unique permutation invariant measure ν_α (which is a Borel probability measure) that satisfies (9.2), called the **radial part** of μ_α, which turns out to be the eigenvalue distribution of Hermitian matrices X as X varies over $\mathcal{O}_\alpha$.

Thus, to find the eigenvalue distribution of the sum of Hermitian matrices, we will need to consider their convolution. Thus if we take the radial part of $\mu_\alpha * \mu_\beta$, which we call $\nu_{\alpha,\beta}$, we will obtain the probability distribution of eigenvalues of $X + Y$ as X varies over $\mathcal{O}_\alpha$ and Y varies over $\mathcal{O}_\beta$.[2]

Thus we have the following relation:

$$\mathrm{Horn}(\alpha, \beta) = \mathrm{supp}(\nu_{\alpha,\beta}) \cap (\mathbb{R}^n)_\downarrow.$$

Define

$$Z : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} : (x, t) \mapsto \int_{\mathcal{U}(n)} e^{\mathrm{tr}(\mathrm{diag}(x) U \mathrm{diag}(t) U^*)} \, \mathrm{d}\omega(U),$$

which is the Harish-Chandra-Itzykson-Zuber integral, which has a computable formula [6]:

$$Z(x, t) = k_n! \frac{1}{V_n(x) V_n(t)} \det\left(e^{x_i t_j}\right)_{i,j \le n} \quad x = (x_i)_{i \le n} \text{ and } t = (t_i)_{i \le n},$$

where $k_n = (n-1, n-2, ..., 1, 0)$ with $k_n! = (n-1)!(n-2)! \cdots 2!$, and $V_n(x) = \prod_{i \neq j} (x_i - x_j)$ is the Vandermonde polynomial.

Let $X = \mathrm{diag}(x)$ for some $x \in \mathbb{R}^n$, then if a bounded measure μ on $\mathcal{H}(n)$ is unitarily invariant, one has

$$\hat{\mu}(X) = \int_{\mathcal{H}(n)} e^{\mathrm{tr}(XY)} \, \mathrm{d}\mu(Y) = \int_{\mathbb{R}^n} Z(x, t) \, \mathrm{d}\nu(t)$$

where ν is the radial part of μ.

Also note that

$$\widehat{\mu_\alpha}(X) = \int_{\mathcal{H}(n)} e^{\mathrm{tr}(XY)} \, \mathrm{d}\mu_\alpha(Y) = Z(x, \alpha).$$

Hence applying the Fourier transform on $\mu = \mu_\alpha * \mu_\beta$ and using the Convolution Theorem 8.2, one obtains

$$Z(x, \alpha) Z(x, \beta) = \widehat{\mu_\alpha}(X) \widehat{\mu_\beta}(X) = \int_{\mathbb{R}^n} Z(x, t) \, \mathrm{d}\nu_{\alpha,\beta}(t). \tag{9.3}$$

Now by J. Faraut [5, 2.1], the measure $\nu_{\alpha,\beta}$ is uniquely determined by (9.3) for all $x \in \mathbb{C}^n$. Hence it suffices to find a Borel probability measure ν such that (9.3) holds for all $x \in \mathbb{C}^n$, which then gives $\nu = \nu_{\alpha,\beta}$.

[2] The convolution here is taken additively.

2 – Construction of The Probability Distribution

Define $q : \mathcal{H}(n) \to \mathbb{R}^n$ as a projection from Hermitian matrices to its diagonal elements, i.e. if $X = \left(x_{ij}\right)_{i,j\leq n} \in \mathcal{H}(n)$, then $q(X) = (x_{11}, ..., x_{nn})$. Then by the Horn convexity theorem [7], one has

$$q(\mathcal{O}_\alpha) = \operatorname{con}\{\sigma(\alpha) : \sigma \in S_n\}$$

where $\operatorname{con}(A)$ is the smallest convex set containing $A \subseteq \mathbb{R}^n$. Now we define the **Heckman measure** $M_\alpha : \mathcal{B}(\mathbb{R}^n) \to [0, 1]$ as $M_\alpha = q_*(\mu_\alpha)$. As M_α is a Borel probability measure, one has M_α is a Radon measure.

Note that given a bounded measure μ on $\mathbb{R}^n$, by the preceding chapter, the Fourier transform is defined as

$$\tilde{\mu}(x) = \int_{\mathbb{R}^n} e^{\langle \omega, x \rangle} \, \mathrm{d}\mu(\omega) \quad \text{for } x \in \mathbb{R}^n.$$

Thus given $x \in \mathbb{R}^n$ with $X = \operatorname{diag}(x)$, one has

$$\begin{aligned} \widetilde{M_\alpha}(x) = \int_{\mathbb{R}^n} e^{\langle \omega, x \rangle} \, \mathrm{d}M_\alpha(\omega) &= \int_{\mathcal{H}(n)} e^{\langle q(H), x \rangle} \, \mathrm{d}\mu_\alpha(H) \\ &= \int_{\mathcal{H}(n)} e^{\operatorname{tr}(HX)} \, \mathrm{d}\mu_\alpha(H) = Z(x, \alpha). \end{aligned}$$

Now define the skew-symmetric Borel measure on $\mathbb{R}^n$:

$$\eta_\alpha = \frac{k_n!}{V_n(\alpha)} \sum_{\sigma \in S_n} \varepsilon(\sigma) \delta_{\sigma(\alpha)}$$

where $\varepsilon(\sigma)$ is the signature of the permutation σ. Now the Fourier transform of η_α is

$$\widetilde{\eta_\alpha}(x) = \frac{k_n!}{V_n(\alpha)} \sum_{\sigma \in S_n} \varepsilon(\sigma) e^{\langle \sigma(\alpha), x \rangle} = \frac{k_n!}{V_n(\alpha)} \sum_{\sigma \in S_n} \det\left(e^{x_i \alpha_j}\right)_{i,j \leq n}$$

where the last equality is shown in J. Faraut [5], and by the Harish-Chandra-Itzykson-Zuber integral formula, one has that

$$\widetilde{\eta_\alpha}(x) = V_n(x) \widetilde{M_\alpha}(x) \quad \text{for } x \in \mathbb{R}^n.$$

Now through some more computations that is unfortunately out of the scope of this paper, J. Faraut showed that

$$\mathrm{d}\nu = \frac{1}{n!} \frac{1}{k_n!} V_n \, \mathrm{d}\left(\eta_\alpha * M_\beta\right)$$

satisfies (9.3), thereby we have obtained our desired probability distribution $\nu_{\alpha,\beta}$.

10 | References

[1] A. A. Klyachko, “Stable bundles, representation theory and Hermitian operators,” *Selecta Mathematica*, vol. 4, no. 3, pp. 419–445, Sep. 1998, doi: 10.1007/s000290050037.

[2] A. Knutson and T. Tao, “The honeycomb model of $GL_n(\mathbb C)$ tensor products I: Proof of the saturation conjecture,” *Journal of the American Mathematical Society*, vol. 12, no. 4, pp. 1055–1090, Apr. 1999, doi: 10.1090/S0894-0347-99-00299-4.

[3] R. Bhatia, “Linear Algebra to Quantum Cohomology: The Story of Alfred Horn's Inequalities,” *The American Mathematical Monthly*, vol. 108, no. 4, pp. 289–318, Apr. 2001, doi: 10.1080/00029890.2001.11919754.

[4] G. B. Folland, *Real analysis: modern techniques and their applications*, 2nd ed. in Pure and applied mathematics. Chichester Weinheim [etc.]: New York J. Wiley & sons, 1999.

[5] J. Faraut, “Horn’s problem and Fourier analysis,” *Tunisian Journal of Mathematics*, vol. 1, no. 4, pp. 585–606, 2018, Accessed: Feb. 10, 2025. [Online]. Available: https://msp.org/tunis/2019/1-4/p05.xhtml

[6] C. McSwiggen, “The Harish-Chandra integral: An introduction with examples.” Accessed: Feb. 10, 2025. [Online]. Available: http://arxiv.org/abs/1806.11155

[7] A. Horn, “Doubly Stochastic Matrices and the Diagonal of a Rotation Matrix,” *American Journal of Mathematics*, vol. 76, no. 3, p. 620, Jul. 1954, doi: 10.2307/2372705.

11 | Appendix

This section involves long-winded calculations and codes that are done throughout the report.

1 – The C_θ Matrix

The section is in reference with Chapter 4.3.

Starting with $A = \operatorname{diag}(\alpha)$, $B = \operatorname{diag}(\beta)$, and

$$U_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \quad \text{for some } \theta \in \mathbb{R}.$$

We defined $C_\theta := A + U_\theta B U_\theta^*$. Through immediate computation, we get that

$$C_\theta = \begin{pmatrix} \alpha_1 + \beta_1 \cos^2(\theta) + \beta_2 \sin^2(\theta) & (\beta_1 - \beta_2)\sin(\theta)\cos(\theta) \\ (\beta_1 - \beta_2)\sin(\theta)\cos(\theta) & \alpha_2 + \beta_2 \cos^2(\theta) + \beta_1 \sin^2(\theta) \end{pmatrix}.$$

Given $x = (x_1, x_2) \in \mathbb{R}^2$, define

$$\overline{x} = \frac{x_1 + x_2}{2} \quad \text{and} \quad \underline{x} = \frac{x_1 - x_2}{2},$$

and we also have the following identities:

$$x_1 \cos^2(\theta) + x_2 \sin^2(\theta) = \overline{x} + \underline{x}\cos(2\theta)$$
$$(x_1 - x_2)\sin(\theta)\cos(\theta) = \underline{x}\sin(2\theta).$$

So one has

$$C_\theta = \begin{pmatrix} \alpha_1 + c & d \\ d & \alpha_2 + c' \end{pmatrix}$$

where

$$c = \overline{\beta} + \underline{\beta}\cos(2\theta), \quad c' = \overline{\beta} - \underline{\beta}\cos(2\theta), \quad \text{and} \quad d = \underline{\beta}\sin(2\theta).$$

Note that

$$\operatorname{tr}(C_\theta)^2 = (\alpha_1 + c + \alpha_2 + c')^2 = (\alpha_1 + c)^2 + 2(\alpha_1 + c)(\alpha_2 + c') + (\alpha_2 + c')^2$$
$$4\det(C_\theta) = 4(\alpha_1 + c)(\alpha_2 + c') - 4d^2$$

so

$$\begin{aligned} \Delta(\theta) := \operatorname{tr}(C_\theta)^2 - 4\det(C_\theta) &= (\alpha_1 + c - \alpha_2 - c')^2 + 4d^2 \\ &= 4\left(\underline{\alpha} + \underline{\beta}\cos(2\theta)\right)^2 + 4d^2 \\ &= 4\left(\underline{\alpha}^2 + 2\underline{\alpha}\underline{\beta}\cos(2\theta) + \underline{\beta}^2\right). \end{aligned}$$

By the trace-determinant formula, the eigenvalues of C_θ are

$$\lambda_\pm(\theta) := \frac{\operatorname{tr}(C_\theta) \pm \sqrt{\Delta(\theta)}}{2} = \overline{\alpha} + \overline{\beta} \pm \sqrt{\frac{\Delta(\theta)}{4}}.$$

Now $\Delta(\theta)$ attains its minimum when $\theta = \frac{\pi}{2}$,

$$\frac{1}{4}\Delta\left(\frac{\pi}{2}\right) = \left(\underline{\alpha} - \underline{\beta}\right)^2$$

and maximum when $\theta = 0$,

$$\frac{1}{4}\Delta(0) = \left(\underline{\alpha} + \underline{\beta}\right)^2.$$

Take $\lambda_1(\theta) = \lambda_+(\theta)$, and $\lambda_2(\theta) = \lambda_-(\theta)$, so $\lambda(\theta) = (\lambda_1(\theta), \lambda_2(\theta)) \in \left(\mathbb{R}^2\right)_\downarrow$. Then as

$$\lambda_1(\theta) + \lambda_2(\theta) = 2\left(\overline{\alpha} + \overline{\beta}\right) = \alpha_1 + \alpha_2 + \beta_1 + \beta_2$$

we see that $\lambda(\theta)$ traces out a line segment as θ varies.

Observe that

$$\lambda_1(\theta) \leq \overline{\alpha} + \overline{\beta} + \underline{\alpha} + \underline{\beta} = \alpha_1 + \beta_1$$
$$\lambda_1(\theta) \geq \overline{\alpha} + \overline{\beta} + \left|\underline{\alpha} - \underline{\beta}\right| \geq \max(\alpha_1 + \beta_2, \alpha_2 + \beta_1)$$

i.e.

$$\operatorname{im}(\lambda_1) = [\max(\alpha_1 + \beta_2, \alpha_2 + \beta_1), \alpha_1 + \beta_1].$$

Similarly,

$$\operatorname{im}(\lambda_2) = [\alpha_2 + \beta_2, \min(\alpha_1 + \beta_2, \alpha_2 + \beta_1)]$$

Hence it follows that $\lambda(\theta)$ for $\theta \in \left[0, \frac{\pi}{2}\right]$ is a line segment between the points

$$(\max(\alpha_1 + \beta_2, \alpha_2 + \beta_1), \min(\alpha_1 + \beta_2, \alpha_2 + \beta_1)) \quad \text{and} \quad (\alpha_1 + \beta_1, \alpha_2 + \beta_2).$$