Birkhoff polytope
and its subset of unistochastic matrices

Karol Życzkowski

in collaboration with

Charles Dunkl (University of Virginia)

J. Math. Phys. **50**, 123521 (2009).

Uniwersytet Warmińsko–Mazurski, Olsztyn

June 21, 2012

Stochastic matrices & Markov chains

Stochastic matrices

Classical states: N-point probability distribution, $\mathbf{p} = \{p_1, \dots p_N\}$,
where $p_i \geq 0$ and $\sum_{i=1}^{N} p_i = 1$

Discrete dynamics – a Markov chain: $p'_i = S_{ij} p_j$, where S is a **stochastic matrix** of size N
and maps the simplex of classical states into itself, $S: \ \Delta_{N-1} \to \Delta_{N-1}$.

Frobenius–Perron theorem

Let S be a **stochastic matrix**:

a) $S_{ij} \geq 0$ for $i, j = 1, \dots, N$,

b) $\sum_{i=1}^{N} S_{ij} \ = \ 1$ for all $j = 1, \dots, N$.

Then

i) the spectrum $\{z_i\}_{i=1}^{N}$ of S belongs to the unit disk,

ii) the leading eigenvalue equals unity, $z_1 = 1$,

iii) the corresponding eigenstate forms a probability vector $\mathbf{p}_{\rm inv}$,
which is invariant, $S\mathbf{p}_{\rm inv} = \mathbf{p}_{\rm inv}$.

Spectra of stochastic matrices

Let Σ_N denotes a subset of the unit disk which supports the spectra of **stochastic** matrices of size N.

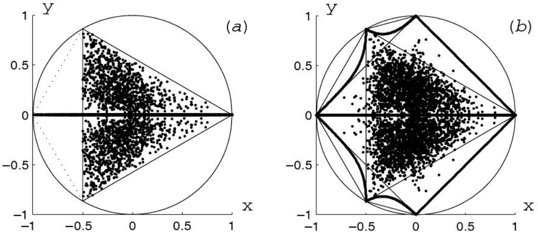
Let Z_k be a regular polygon centered at 0 with a corner at $z = 1$.

a) $N = 2$: the spectrum of S is real, $\Sigma_2 = [-1, 1] = Z_2$

b) $N = 3$: the spectrum contains an interval and a triangle, $\Sigma_3 = Z_2 \cup Z_3$

c) $N = 4$: the spectrum contains an interval, a triangle, and a square $\Sigma_4 \supset Z_2 \cup Z_3 \cup Z_4$ but it is contained in the convex hull of this set.

$N = 3$



$N = 4$:

The boundary of the non–convex set Σ_N was derived by **Karpelevich (1951)**, a simplified proof given by **Djokovic** in 1990.

Set $\mathcal{B}_N$ of bistochastic matrices of size N

The Birkhoff polytope

A square matrix B is called **bistochastic** (doubly stochastic) if
– it has positive elements $B_{ij} \geq 0$,
– the sum in each column and each row is equal to unity,

$$\sum_i B_{ij} = \sum_j B_{ij} = 1.$$

Birkhoff theorem. Every bistochastic matrix can be written as a convex combination of permutation matrices P_k, $\quad B = \sum_k q_k P_k$.

Thus the set $\mathcal{B}_N$ is called the **Birkhoff polytope**

In general a matrix $B \in \mathcal{B}_N$ is described by $(N-1)^2$ parameters, so the **Birkhoff polytope** $\mathcal{B}_N \subset \mathbb{R}^{(N-1)^2}$.

Bistochastic matrices for $N = 2$

$$B_2 = (a) = \begin{bmatrix} a & 1-a \\ 1-a & a \end{bmatrix} = a\mathbb{1} + (1-a)P_{12}, \text{ for } a \in [0,1].$$

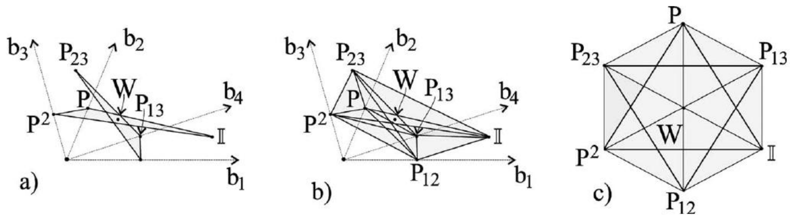
Thus $N = 2$ **Birkhoff polytope** is equivalent to unit interval, $\mathcal{B}_2 = [0,1]$.

Set $\mathcal{B}_N$ of bistochastic matrices of size N

Bistochastic matrices for $N = 3$, for which $\mathcal{B}_3 \subset \mathbb{R}^4$

$$B_3(b_1, b_2, b_3, b_4) \ := \begin{bmatrix} b_1 & b_2 & 1 - b_1 - b_2 \\ b_3 & b_4 & 1 - b_3 - b_4 \\ 1 - b_1 - b_3 & 1 - b_2 - b_4 & \sum_{i=1}^{4} b_i - 1 \end{bmatrix} \in \mathcal{B}_3$$

The set $\mathcal{B}_3$ is the convex hull of $3! = 6$ permutation matrices, $\{\mathbb{1},\ P = P_{123},\ P^2 = P_{132},\ P_{12},\ P_{13},\ P_{23}\}$.



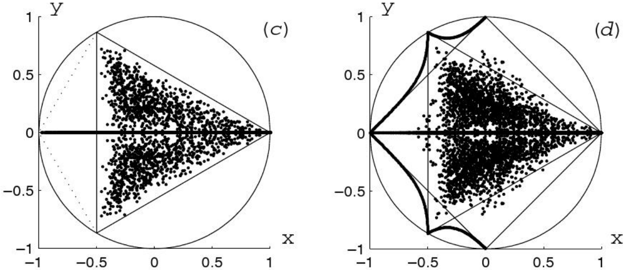
W denotes the **flat bistochastic** matrix, $W = \frac{1}{3}[1, 1, 1;\ 1, 1, 1;\ 1, 1, 1]$, located at the center of the **Birkhoff polytope** $\mathcal{B}_3$.

Spectra of bistochastic matrices

Let Σ'_N denotes a set which supports the spectra of **bistochastic** matrices of size N.

Since any **bistochastic** matrix is stochastic,
the support Σ'_N is contained in Σ_N

Are both sets **equal** for each N ??

$N = 3$  $N = 4$:

Superimposed spectra of 3000 random bistochastic matrices of size $N = 3$ and $N = 4$.

Markov chains and (underlying) unitary dynamics

Consider discrete dynamics described by a **Markov chain**, $p'_i = B_{ij} p_j$, represented by a **bistochastic** matrix $B_{ij} \geq 0$, sum in each column (each row) equal to unity, $\sum_i B_{ij} = \sum_j B_{ij} = 1$.

In physical problems, such a dynamics is often governed by

a **unitary matrix** V,

such that the measurable **transition probabilities** read $B_{ij} = |V_{ij}|^2$, for $i, j = 1, \ldots, N$.

A natural mathematical question arises:

Given a **bistochastic** matrix $\mathbf{B}$ find out if there exists a corresponding **unitary** matrix V such that $|V_{ij}|^2 = \mathbf{B_{ij}}$

and check, whether such a unitary V is orthogonal.

Unistochastic and orthostochastic matrices

Definitions

A bistochastic matrix $B \in \mathcal{B}_N$ is called **unistochastic** if there exists a unitary $\mathbf{U} \in U(N)$ such that

$$B_{ij} = |\mathbf{U_{ij}}|^2, \text{ written } \ B = f(U).$$

A bistochastic matrix $B \in \mathcal{B}_N$ is called **orthostochastic** if there exists an orthogonal $\mathbf{O} \in O(N)$ such that

$$B_{ij} = \mathbf{O_{ij}}^2, \text{ written } \ B = f(O).$$

Let $\mathcal{U}_N$ and $\mathcal{O}_N$ denote the sets of **unistochastic** and **orthostochastic** matrices of size N, respectively.

By definition the following inclusion relation hold

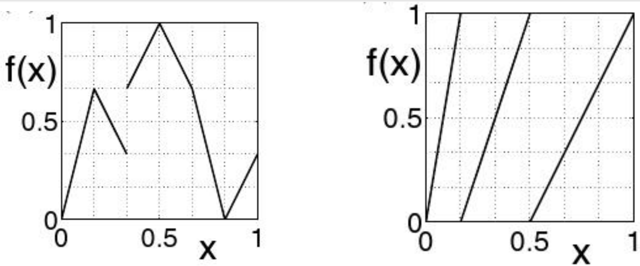
$$\mathcal{O}_N \subset \mathcal{U}_N \subset \mathcal{B}_N.$$

Quantized 1-d dynamical systems

Quantum evolution (of a closed system!) is **unitary**, $|\psi'\rangle = U|\psi\rangle$,
and it is reversible, $|\psi\rangle = U^*|\psi'\rangle$.

To find a **quantum analogue** of a dynamical system $g : \mathbb{R} \to \mathbb{R}$ one
a) finds its **Markov partition** and transition matrix B and verifies, whether it is bistochastic.
b) if it is so one cheks if B is unistochastic,
i.e. there exists unitary U such that $B_{ij} = |\mathbf{U_{ij}}|^2$.

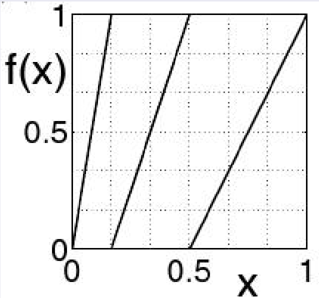
The matrix U describes a **quantum analogue** of the *classical system g*.



Examples: a) quantizable, b) non-quantizable classical system

Non-quantizable 1-d dynamical system

A counterexample (for quantization)

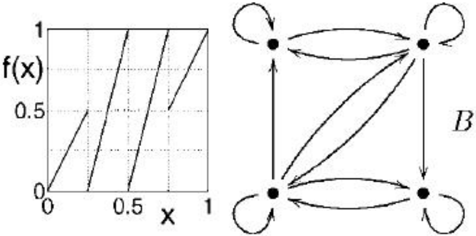


$$B = \frac{1}{6}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 & 2 \\ 3 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 3 \end{bmatrix}$$

The transition matrix B is not **unistochastic** !

There is no **quantum analogue** – no corresponding unitary matrix U...

Example 3: **Four legs map** and its quantization
Pakoński, Kuś, K.Ż. (2001)



$$B^{(4)} = \frac{1}{4}\begin{pmatrix} 2 & 2 & . & . \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ . & . & 2 & 2 \end{pmatrix}$$

$$U^{(4)} = \frac{1}{2}\begin{pmatrix} \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{pmatrix}$$

$$B^{(8)} = \frac{1}{4}\begin{pmatrix} 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 \end{pmatrix}$$

$$U^{(8)} = \frac{1}{2}\begin{pmatrix} \sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2} & -\sqrt{2} \end{pmatrix}$$

Uni- and ortho-stochastic matrices for $N = 2$

Proposition For $N = 2$ all three sets coincide, $\mathcal{O}_2 = \mathcal{U}_2 = \mathcal{B}_2 = [0, 1]$.

Proof. Take any $a \in [0, 1]$ and set $B = \begin{bmatrix} a & 1-a \\ 1-a & a \end{bmatrix}$.

The corresponding orthogonal matrix reads

$$O = \begin{bmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{bmatrix}, \quad \text{where } a = \cos^2\vartheta.$$

In other words

every $N = 2$ **bistochastic** matrix is **orthostochastic**

(and thus also **unistochastic**).

Uni- and ortho-stochastic matrices for $N = 3$

Proposition. For $N = 3$ both inclusion relations
$\mathcal{O}_3 \subset \mathcal{U}_3 \subset \mathcal{B}_3$ are proper.
a) b)

Proof by demonstration.

a) **Fourier** matrix $F_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{bmatrix}$ with $\omega = e^{2\pi i/3}$ is unitary and corresponds to the flat bistochastic matrix $W = f(F_3)$, as $W_{ij} = 1/3$.
Thus W is **unistochastic** but not **orthostochastic**, since for $N = 3$ there are no **Hadamard matrices**.

b) Example of **Schur**: the bistochastic matrix B_S,

$B_S = \frac{P+P^2}{2} = \frac{1}{2} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ is **bistochastic** but not **unistochastic**.

Unistochasticity and chain–link condition for $N = 3$

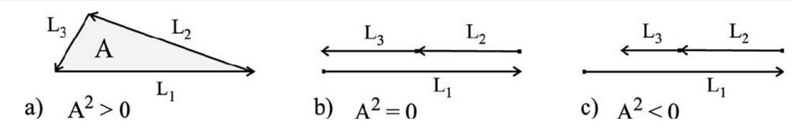
Unitarity condition $UU^\dagger = \mathbb{1}$ can be written as $\langle u_\mu | u_\nu \rangle = \delta_{\mu,\nu}$

Diagonal elements $\langle u_\mu | u_\mu \rangle = 1$, impose bistochasticity, $\sum_i B_{i\mu} = 1$ while orthogonality relation $\langle u_\mu | u_\nu \rangle = 0$ imposes further **constraints** for elements of $B = f(U)$!

What **constraints** for **unistochasticity**?

As the sum $\langle u_1 | u_2 \rangle = \sum_{j=1}^{3} U_{j1} U_{j2}^* = L_1 e^{i\chi_1} + L_2 e^{i\chi_2} + L_3 e^{i\chi_3}$ of three complex numbers should vanish, their (ordered) moduli $L_1 \geq L_2 \geq L_3$ satisfy the following **chain link condition** (*triangle inequality*)

$$L_1 \;\leq\; L_2 + L_3 \qquad \text{with } L_k := \sqrt{B_{1k} B_{2k}}.$$



a) **unistochastic** matrix with a positive area of the unitarity triangle, $A^2 > 0$, b) limiting case: an **orthostochastic** matrix with $A^2 = 0$, c) **bistochastic** matrix B not included into $\mathcal{U}_3$ for which $A^2 < 0$

Unitarity triangle formed by links L_1, L_2, L_3

The length of the links of the **unitarity triangle** read

$$L_1 = \sqrt{b_1 b_2}, \quad L_2 = \sqrt{b_3 b_4}, \quad L_3 = \sqrt{(1 - b_1 - b_2)(1 - b_3 - b_4)}, \quad (1)$$

Let $p = (L_1 + L_2 + L_3)/2$ denotes its **semiperimeter**.
Making use of the **Heron's formula** for the area of the triangle

$$A \;=\; \sqrt{p(p - L_1)(p - L_2)(p - L_3)}\ , \quad (2)$$

we arrive with a compact expression for the squared area A^2,

$$A^2 = [4 b_1 b_2 b_3 b_4 - (b_1 + b_2 + b_3 + b_4 - 1 - b_1 b_4 - b_2 b_3)^2]/16. \quad (3)$$

The chain–links conditions

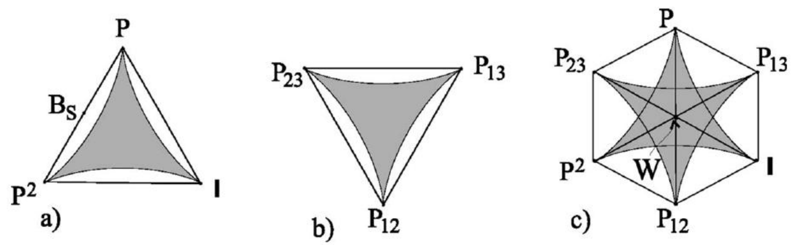
are equivalent to a single condition for **unistochasticity**:

$$A^2(B) \geq 0$$

(if a triangle exists its area is real and positive !)

The set $\mathcal{U}_3$ of unistochastic matrices of size $N = 3$

cross-sections of $\mathcal{U}_3$ (implied by $A^2(B) \geq 0$)



Nonconvex 3-Hypocycloid obtained by the cross-section of $\mathcal{U}_3$ along the plane spanned by the equilateral triangle $\triangle(P, P^2, \mathbb{1})$, b) a similar cross-section along totally orthogonal plane, c) a view 'from above'.

The set $\mathcal{O}_3$ of orthostochastic matrices

Proposition. For $N = 3$ the set $\mathcal{O}_3$ of **orthostochastic** matrices forms the boundary of the $4D$ set $\mathcal{U}_3$ of **unistochastic** matrices.

Unistochastic matrices

are useful for quantizing classical dynamical systems (which lead to bistochastic transition matrices).

Prot Pakoński, Ph.D. Thesis 2002,
Pakoński, Życzkowski, Kuś, 2001,

The set $\mathcal{U}_3$ of $N = 3$ unistochastic matrices

was investigated in Bentgsson, Ericsson, Kuś, Tadej, Życzkowski, *Commun. Math. Phys.* (2005).

The set $\mathcal{U}_3$ of **unistochastic** matrices of size $N = 3$ occupies (with respect to the Lebesgue measure) more than $3/4$ of the corresponding Birkhoff polytope $\mathcal{B}_3$,

$$\frac{\mathrm{vol}(\mathcal{U}_3)}{\mathrm{vol}(\mathcal{B}_3)} = \frac{8\pi^2}{105} = 0.751969... \qquad (4)$$

Dunkl, Życzkowski, 2009

Unitarity triangle and Jarlskog invariant (for $N = 3$)

Jarlskog invariant

For any unitary $U \in U(3)$ define the number $J(U) := \mathbf{Im}(U_{11} U_{22} U_{12}^* U_{21}^*)$ called **Jarlskog invariant**.

Equivalent unitary matrices

Two unitary matrices U and U' are called **equivalent** if there exist two diagonal unitary matrices, D_A and D_B, and two permutations P_A and P_B such that

$$U \sim U' \;=\; D_A P_A U P_B D_B \tag{5}$$

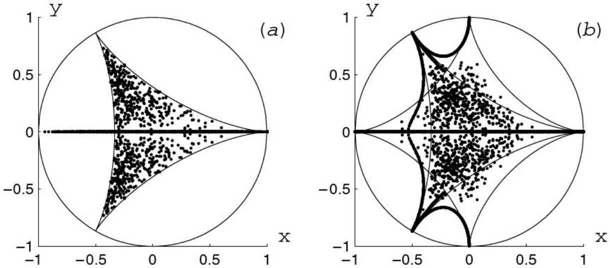
The following relation holds: if $U \sim U'$ then $J(U) = J(U')$, **Jarlskog** 1985

Simple calculation shows that the **Jarlskog invariant** is related to the **area** of the **unitarity triangle**,

$$J^2(U) = 4A^2(B), \qquad \text{where } B = f(U).$$

Spectra of unistochastic matrices

Since any **unistochastic** matrix is **bistochastic**,
the support of the spectra of matrices from $\mathcal{U}_N$
is contained in the support Σ'_N of spectra of bistochastic matrices.



$N = 3$ $N = 4$:

Superimposed spectra of 3000 **Haar random** unistochastic matrices of size $N = 3$ and $N = 4$.

***N*–hypocycloids** again...

Speculations on the set of unistochastic matrices

The set $\mathcal{B}_N$ of Bistochastic matrices (Birkhof Polytope)

$\mathcal{B}_N =$ **convex hull** of the set of $N!$ permutation matrices

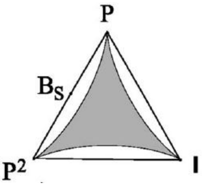
wilde speculation:

The set $\mathcal{U}_N$ of Unistochastic matrices

perhaps

$\mathcal{U}_N =$ a **"special, non-convex" hull**
of the set of $N!$ permutation matrices

example $N = 3$:



What kind of **"special, non-convex" hull** ??

Speculation 1. **Cayley–convex set**

Cayley transform

Let S be a skew hermitian matrix, $S = -S^\dagger$.

Then its **Cayley transform** is unitary,

$$C(S) \ = \ \frac{\mathbb{1} - S}{\mathbb{1} + S} \ = \ U.$$

The **inverse Cayley transform** sends a unitary U into skew hermitian S:

$$C^{-1}(U) \ = \ \frac{\mathbb{1} - U}{\mathbb{1} + U} \ = \ S$$

Cayley combination of two unitaries, U and W

$$V(a) = C\left[aC^{-1}(U) + (1-a)C^{-1}(W)\right] = \frac{\mathbb{1} - a\frac{\mathbb{1}-U}{\mathbb{1}+U} - (1-a)\frac{\mathbb{1}-W}{\mathbb{1}+W}}{\mathbb{1} + a\frac{\mathbb{1}-U}{\mathbb{1}+U} + (1-a)\frac{\mathbb{1}-W}{\mathbb{1}+W}}$$

is **unitary !**

Speculation 2. **Log–convex set**

Logarithm of a unitary matrix

Any unitary matrix U can be diagonalized, $U = WDW^\dagger$.
Define the **logarithm** $L = \log U = W^\dagger(\log D)W$ such that $U = \exp(L)$.

technial assumption: the spectrum D does not contain -1

Log–convex combination of two unitaries, U and W

$$W' = U^a V^{1-a}$$

or

$$W = \exp[a \log U + (1-a) \log V]$$

is **unitary !**

Ando mean of

Geometric mean of two matrices of a full rank reads

$$A\#B = A^{1/2}\,(A^{-1/2}BA^{-1/2})^{1/2}\,A^{1/2},$$

see **(Ando 1978)** but also **Pusz and Woronowicz (1975)**

Ando–convex combination of two unitaries, U and W

$$U\#_t W = U^{1/2}\exp\Big(t\log(U^{-1/2}WU^{-1/2})\Big)U^{1/2}.$$

is **unitary !**

Is the set $\mathcal{U}_N$ of **unistochastic matrices** related to Cayley/log/Ando-combinations of permutation matrices ??

Some open question

- What is the set of Cayley/log/Ando-combinations of all permutation matrices of order N ?
- What is the (minimal) set of unitary matrices such that their Cayley/log/Ando-combinations form the entire set of unitary matrices
- Are bistochastic matrices obtained from Cayley/log/Ando-combinations of permutation matrices at the boundary of the set $\mathcal{U}_N$ of unistochastic matrices of size N?

  Consider, for instance the **Cayley combination** of matrices.
  Is the following implication true:

$$B = \sum_i a_i P_i \in \ \partial \mathcal{B}_N \ \Rightarrow \ f\Big(C \big[\sum_{i=1}^{M} a_i C^{-1}(U_i) \big] \Big) \in \partial \mathcal{U}_N .$$

Concluding Remarks

- A **bistochastic** matrix B corresponds to a unitary matrix if it is **unistochastic**, $B = f(U)$ so that $B_{ij} = |U_{ij}|^2$.
- for $N = 2$ every bistochastic matrix is **orthostochastic**.
- The set $\mathcal{U}_3$ of **unistochastic** matrices of size $N = 3$ is explicitely characterized by the **unitarity triangle** condition:
$$B \in \mathcal{U}_3 \Leftrightarrow A^2(B) \geq 0.$$
- For $N = 3$ the boundary of the set $\mathcal{U}_3$ consists of **orthostochastic** matrices, for which $A^2(B) = 0$.
Thus a generic unistochastic matrix of is **not** orthostochastic
- For $N = 3$ we computed the volume of the set $\mathcal{U}_3$ and the average value of the **Jarlskog invariant** J for a random Haar unitary matrix $U \in U(3)$.
- For $N \geq 4$ the unistochasticity problem **remains open** !