Volume of the set of unistochastic matrices of order 3 and the mean Jarlskog invariant

Charles Dunkl
Department of Mathematics, University of Virginia, Charlottesville, VA 22904-4137, USA

Karol Życzkowski
Institute of Physics, Jagiellonian University, Cracow and
Center for Theoretical Physics, Polish Academy of Sciences, Warsaw

November 10, 2009

Abstract

A bistochastic matrix $B$ of size $N$ is called unistochastic if there exists a unitary $U$ such that $B_{ij} = |U_{ij}|^2$ for $i, j = 1, \ldots, N$. The set $U_3$ of all unistochastic matrices of order $N = 3$ forms a proper subset of the Birkhoff polytope, which contains all bistochastic (doubly stochastic) matrices. We compute the volume of the set $U_3$ with respect to the flat (Lebesgue) measure and analytically evaluate the mean entropy of an unistochastic matrix of this order. We also analyze the Jarlskog invariant $J$, defined for any unitary matrix of order three, and derive its probability distribution for the ensemble of matrices distributed with respect to the Haar measure on $U(3)$ and for the ensemble which generates the flat measure on the set of unistochastic matrices. For both measures the probability of finding $|J|$ smaller than the value observed for the CKM matrix, which describes the violation of the CP parity, is shown to be small. Similar statistical reasoning may also be applied to the MNS matrix, which plays role in describing the neutrino oscillations. Some conjectures are made concerning analogous probability measures in the space of unitary matrices in higher dimensions.

1Email address: cfd5z@virginia.edu
2Email address: karol@tatry.if.uj.edu.pl
1 Introduction

Bistochastic matrices appear in variety of problems in different branches of science. A bistochastic matrix (also called doubly stochastic) contains real non-negative entries, the sum of which in each column and in each row is equal to unity. Thus each column and each row of such a matrix can be interpreted as a probability vector. The structure of the set $\mathcal{B}_N$ of all bistochastic matrices of order $N$ is well understood [1]. It is formed by the convex polytope of all permutation matrices of size $N$ and it is often called the Birkhoff polytope [2]. Analytical expressions for its volume for small dimensionality [3,4] and for the leading terms of the volume in the asymptotic limit [5,6] are available in the literature.

In various physical problems it is assumed that the probabilities, which form the entries of a bistochastic matrix, arise as squared modulus of an element of an (a priori unknown) unitary matrix. A bistochastic matrix $B$ which can be generated from an unitary matrix $U$ by the relation $B_{ij} = |U_{ij}|^2$ is called unistochastic (or orthostochastic).

For instance, such matrices are used in high energy physics to characterize interactions between elementary particles, which can be divided into $N$ generations. Since the Hamiltonians describing two kinds of physical interactions (usually called 'strong' and 'weak') do not commute, these two Hermitian operators determine a single unitary matrix of order $N$, which relates both eigenbases. This is the famous unitary matrix of Cabbibo-Kobayashi-Maskawa [7]. Note that the squared moduli of the CKM matrix $V_{\text{CKM}}$ form the corresponding unistochastic matrix $B$, the entries of which represent probabilities which are accessible experimentally [8–12].

According to the Standard Model of elementary particles there exist three generations of quarks, thus the case of a direct physical importance are unistochastic matrices of order $N = 3$. On the other hand, the case $N = 4$ could also become relevant in case a fourth generation of quarks should be discovered [13–15]. A similar problem in the neutrino physics is characterized by the Maki-Nakagawa-Sakata matrix (MNS matrix) [16], some parameters of which are still quite uncertain. A relation between the MNS and CKM matrices is studied in [17].

In practice, given a bistochastic matrix $B \in \mathcal{B}_N$ it is important to know, whether it belongs to the set $\mathcal{U}_N$ of unistochastic matrices. If this is the case one would like to describe the set of all unitary matrices such that $B_{ij} = |U_{ij}|^2$ for $i, j = 1, \ldots, N$. These questions are of a particular interest for research in foundations of quantum mechanics and investigation of properties of transition probabilities [18–20], scattering theory [21], quantum counterparts of Markov processes and dynamics on graphs [22,23], and the theory of quantum information processing [24].

Any bistochastic matrix of order two is unistochastic, so both sets coincide, $\mathcal{U}_2 = \mathcal{B}_2$. The situation differs already for $N = 3$. To show this fact Schur
considered the symmetric combination of cycle-three permutation matrices, $P$ and $P^{-1} = P^2$. This matrix

$$B_S = \frac{1}{2}(P + P^2) = \frac{1}{2} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

is clearly bistochastic but it is easy to see that there is no corresponding unitary matrix. Hence $U_3 \subset B_3$ and a similar relation holds for an arbitrary $N \geq 3$. Vaguely speaking, the moduli of a unitary matrix need to fulfill certain constraints, more stringent than the obvious fact that the sum of squared moduli in each row (each column) is equal to unity. As recently analyzed by P. Dit\u{a}\u{a} [25], this simple observation has important consequences for reconstruction of unitary matrices from experimental data.

A general question, if a given bistochastic matrix $B$ is unistochastic, remains open, and only partial results are available [26,27]. Necessary and sufficient conditions for unistochasticity are known for $N = 3$ [9,28,29], while the constraints for unistochasticity recently obtained by Dit\u{a}\u{a} [25] for the general case of $N \geq 4$ are formulated implicitly and do not provide a constructive solution of the problem.

Geometrical properties of the set of $U_3$ of unistochastic matrices of order 3 were studied in [27,28]. This set contains a 4-dimensional unistochastic ball centered at the flat (van der Waerden) matrix $W$, for which all entries are equal to $1/3$. The set $U_3$ is not convex but it is star shaped. This means that if $B \in U_3$ then the entire interval $BW$ belongs to the set. The volume of $U_3$ with respect to the Euclidean (Lebesgue) measure was numerically estimated by a Monte Carlo-type procedure [27].

The main aim of this work is to derive an analytical formula for the volume of the set $U_3$ of unistochastic matrices of order 3 with respect to the flat, Euclidean measure. We obtain also a compact expression allowing one to average any function of elements of $B$ over the set $U_3$. In particular we derive an explicit result for the average generalized entropies $S_q$ which in the special case $q = 1$ gives the mean Shannon entropy of the columns of unistochastic matrices. We compute also the average Jarlskog invariant $J$ proportional to the area $A$ of the unitarity triangle, which characterizes any unistochastic matrix [8,9]. Furthermore we derive higher moments $\langle J^K \rangle$ and analyze the probability distribution $P(J)$, to get more insight into properties of the CKM matrix, which describes violation of the CP symmetry. Such an approach was recently suggested by Gibbons et al. [30], who used other probability measures for this purpose.

All averages are computed with respect to the natural Euclidean measure in the set of unistochastic matrices and are compared with the averages with respect to the measure induced by the Haar measure on $U(3)$. This measure leading to the unistochastic ensemble [26], called flag–manifold measure in [30], is not uniform in the set $U_3$. Although approach presented, based on properties of intertwining operators associated to the group $S_3$ [31,32], is directly applicable to the case $N = 3$, we conjecture also some properties of the measures in the sets of unistochastic matrices in higher dimensions.
The paper is organized as follows. In the next two sections we present
the necessary definitions and review key properties of the Birkhoff polytope \( B_3 \)
and its subset \( U_3 \) containing unistochastic matrices. In section 4 we define a
family of measures in the set \( U_3 \) and compute its volume with respect to them.
Similar study of average entropies is presented in section 4 while the average
Jarlskog invariant and its distribution are investigated in section 5. The paper
is concluded with section 6, while some conjectures concerning the measures in
the set of unitary matrices for \( N \geq 4 \) are relegated to the Appendix.

2 The Birkhoff polytope \( B_3 \)

A real square matrix \( B \) of order \( N \) is called bistochastic (or doubly stochastic)
if it satisfies the following conditions

\[
\begin{align*}
    \text{i)} & \quad B_{ij} \geq 0 \\
    \text{ii)} & \quad \sum_i B_{ij} = 1 \\
    \text{iii)} & \quad \sum_j B_{ij} = 1 .
\end{align*}
\]

(2)

Such a matrix \( B \) is often used to describe discrete dynamics, \( p' = Bp \), in
the space of probability vectors. Condition i) implies that all elements of the
transformed vector \( p' \) are non-negative. Due to condition ii) its 1–norm \( \sum_i p_i \),
is preserved. A matrix satisfying two first conditions is called stochastic and
it sends the simplex of \( N \)-point probability vectors into itself. Condition iii)
implies that additionally the transposed matrix \( B^T \) is stochastic, which explains
the name.

The uniform probability vector \( p_* \) with all components equal, \( p_i = 1/N \)
stays clearly invariant with respect to any bistochastic matrix, \( BP_* = P_* \). Thus
a bistochastic matrix describes a (weak) contraction of the probability simplex
towards the uniform distribution \( p_* \).

Let \( B_N \) denote the set of all bistochastic matrices of order \( N \), called Birkhoff
polytope. This convex polytope is well known in linear programming. Since it
arises in the problem of assigning \( N \) workers to \( N \) tasks, given their efficiency
ratings for each task, it is sometimes called the assignment polytope.

The Birkhoff polytope is equivalent to the convex hull of all \( N! \) permutation
matrices of size \( N \). Hence a permutation matrix \( P \) forms an extremal point
of \( B_N \). All corners of \( B_N \) are equivalent in the sense that a given corner can
be obtained from another one by an orthogonal transformation. A bistochastic
matrix belongs to the boundary of the Birkhoff polytope if and only if at least
one of its entries is equal to zero.

There exists a unique bistochastic matrix \( W \), with all entries equal, \( W_{ij} = 1/N \).
It is also called a matrix of van der Waerden, since it saturates the van
der Waerden inequality [1] concerning the permanent of bistochastic matrices,
\( \text{per}(B) \geq N!N^{-N} \). It is easy to see that \( W \) is located symmetrically at the
center of the Birkhoff polytope.

A bistochastic matrix \( B \) can be determined by its minor of size \( (N - 1) \).
Hence the dimensionality of the Birkhoff polytope \( B_N \) equals \( (N - 1)^2 \). For
instance, the dimension of the set \( B_2 \) is equal to one, and this set forms indeed an
interval between the identity matrix $1_2$ and the 2–element permutation matrix. In other words, any bistochastic matrix of order two can be written as

$$B_2(a) = \begin{bmatrix} a & 1-a \\ 1-a & a \end{bmatrix} \text{ where } a \in [0,1]. \quad (3)$$

The length of this interval, equivalent to volume of $B_2$, is equal to unity, if we consider it as a subset of $\mathbb{R}^3$. However, we are going to consider this set as an element of $\mathbb{R}^{N^2}$ then the distance between the points $(1,0,1,0)$ and $(0,1,0,1)$ is equal to 2, so in these units one has $vol_1(B_2) = 2$.

In this paper we are going to work with the case $N = 3$, so the Birkhoff polytope is defined as a convex hull of $3! = 6$ permutation matrices,

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad P^2 = P^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad 1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (4)$$

$$P_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad P_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad P_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \quad (5)$$

We divided the permutation matrices into two triples, which belong to two totally orthogonal 2–planes. A uniform mixture in any triple produces the flat matrix $W$,

$$P + P^2 + 1 = W = P_{12} + P_{13} + P_{23}. \quad (6)$$

Working with the standard Hilbert-Schmidt distance, defined by $D^2(A,B) := \text{Tr}(A-B)(A-B)^*$, we see that both triples form equilateral triangles. To produce a sketch of them in 4 dimensions we will use the following parametrization

$$B(\vec{b}) = B(b_1, b_2, b_3, b_4) := \begin{bmatrix} b_1 & b_2 & 1-b_1-b_2 \\ b_3 & b_4 & 1-b_3-b_4 \\ 1-b_1-b_3 & 1-b_2-b_4 & \sum_{i=1}^4 b_i - 1 \end{bmatrix}. \quad (7)$$

Both triangles shown in Fig. 1, cross at their center $W$. Six permutation matrices form 15 edges, out of which all belong to the boundary of $B_3$ and all are extremal. There are six long edges, of length $\sqrt{6}$, which form two equilateral triangles, and nine short edges of length 2. If one plots all of them, as in Fig. 1b, the sketch of $B_3$ becomes complete, but not very illuminating. Another natural possibility is to look at the polytope 'from above', the direction distinguished by the vector $\vec{b} = (1/3, 1/3, 1/3, 1/3)$. The Birkhoff polytope then appears symmetrically as a regular hexagon with two inscribed equilateral triangles forming the Star of David - see Fig. 1c. Note that all the diagonals of the hexagon belong to the boundary of $B_3$, and that the distance $PP_{12}$ of the diagonal is shorter than the side $P1$ of the equilateral triangle. Any 2-d face of the polytope is formed by an isosceles triangle with two short edges and one long.
The polytope $B_3$ is defined as the convex hull of 6 corners, so one could think, it may be decomposed into two 4-d simplices, each determined by 5 points. This would be possible, if we could select 4 corners, which span the base of a simplex and then allow two other corners to play the role of an apex for two simplices, with the same base. However, this would require that the edge connecting both apexes is not extremal or it includes one of the corner from the base of the simplex. Neither of these holds for the Birkhoff polytope, so its decomposition into two simplices is not possible.

To find a decomposition of $B_3$ into three simplices take three corners of one equilateral triangle, e.g. $\triangle(P, P_{12}, P_{13})$. Out of the orthogonal triangle select a side, say the one formed by the corners $P_{12}$ and $P_{13}$. These five corners define a 4-d simplex. The same construction performed for two other sides of the $\triangle(P_{12}, P_{13}, P_{23})$ produces two other simplices. It is easy to show any point of $B_3$ belongs to one of these simplices and that the 4-d volume of any of their intersections is equal to zero. Such a triangulation of the Birkhoff polytope allows to find that its volume in $\mathbb{R}^9$ according to the Lebesgue measure is equal to $9/8$. A detailed investigation of the geometry of the Birkhoff polytope is provided in [33].

3 The set $U_3$ of unistochastic matrices

A certain class of bistochastic matrices can be generated from unitary matrices. Let $U$ denote a unitary matrix. Unitarity condition, $UU^* = \mathbb{I}$, implies that the matrix $B$ defined by

$$B = f(U), \quad \text{so that} \quad B_{ij} = |U_{ij}|^2$$

is bistochastic. Any bistochastic matrix $B \in B_N$ for which there exists unitary $U \in U(N)$ such that $B = f(U)$ is called unistochastic. The set of all unistochastic matrices of size $N$ will be denoted as $U_N$.

Note that the multiplication of $U$ by any diagonal unitary matrices $D_1$ and $D_2$ changes the phases of entries of $U$, but does not modify the corresponding
bistochastic matrix. Hence we define an equivalence relation

\[ U \approx U' = D_1 U D_2 \]  

and observe that \( B = f(U) = f(U') \).

If the unitary matrix \( U \), appearing in (8) is orthogonal, the corresponding bistochastic matrix \( B \) is called orthostochastic. This is the case for any bistochastic matrix of size 2, since writing an orthogonal matrix \( O = \begin{pmatrix} \cos \vartheta & \sin \vartheta \\ -\sin \vartheta & \cos \vartheta \end{pmatrix} \) and taking \( \vartheta = \arccos a \) we see that \( B = f(O) \) for any bistochastic matrix of order two represented in the form \( (9) \). Therefore any unistochastic matrix of order 2 is also orthostochastic. This is no longer the case for \( N \geq 3 \), as it is explicitly shown later in this section.

Interestingly this simple mathematical observation has far reaching consequences for physics. In the theory of elementary particles one defines a discrete space–time symmetry called CP, which stands for charge conjugation and parity. Such a symmetry requires that a physical process in which all particles are replaced by their antiparticles is equivalent to the mirror image of the original process.

If such a symmetry were obeyed the CKM matrix \( V_{\text{CKM}} \) would be orthogonal, (and thus invariant with respect to the complex conjugation) or it would be equivalent to an orthogonal matrix with respect to \( (9) \). For \( N = 2 \) this is the case for any unitary matrix from \( U(2) \). As the CP symmetry was discovered in 1964 to be violated in experiments on decay of neutral mesons \( K \), one could predict that the number \( N \) of generations of quarks in the theory, equal to the size of the CKM matrix, has to be greater than two.

In her first paper on the CKM matrix \[8\] Jarlskog observed that for any unitary matrix \( U \) of size 3 the number

\[ J := \text{Im}(U_{11} U_{22} U_{12}^* U_{21}^*) \]  

is invariant with respect to multiplication of the matrix \( U \) by diagonal unitary matrices and permutations. This quantity, now called the Jarlskog invariant, computed for the CKM matrix \( V_{\text{CKM}} \) can be considered as a measure of the violation of the CP symmetry.

Consider now an arbitrary bistochastic matrix \( B \) of size \( N \). To check if this matrix is unistochastic we need to know whether there exists a unitary matrix \( U \) such that \( f(U) = B \) according to eq. (8). The moduli of the unitary matrix are determined by the square roots of the entries of the bistochastic matrix, \( |U_{ij}| = \sqrt{B_{ij}} \) and one needs to find a set of phases \( \phi_{ij} \) which guarantees unitarity, where \( U_{ij} = |U_{ij}| \exp (i\phi_{ij}) \).

Choose the two first columns of \( U \), which we denote as \( |u_1 \rangle \) and \( |u_2 \rangle \). Their orthogonality relation, \( \langle u_1 | u_2 \rangle = 0 \), implies that

\[ \sum_{j=1}^{N} U_{j1} U_{j2}^* = 0. \]

\[ ^3 \text{In some papers this name is used for unistochastic matrices as well.} \]
Introducing the notation $L_j = |U_{j1}| \cdot |U_{j2}|$ and $\theta_j = \phi_{j1} - \phi_{j2}$ we may rewrite this relation as

$$\sum_{j=1}^{N} L_j \exp(i\theta_j) = 0. \quad (11)$$

This form has a nice geometric interpretation: given a set of $N$ line segments of lengths $L_1, \ldots, L_N$ we need to find the phases $\theta_j$ in such a way that the entire chain is closed. Obviously it cannot be done unless the longest link is shorter or is equal to the sum of all other links. We are free to change the order of summation in (11) and hence to relabel the links in such a way that they are ordered non-increasingly, $L_1 \geq L_2 \geq \cdots \geq L_N$. Then the chain links condition reads

$$L_1 \leq L_2 + \cdots + L_N. \quad (12)$$

If it is satisfied the chain can be closed and forms a unitarity polygon.

This relation was imposed by the assumed orthogonality of the first two columns of $U$, but analogous conditions should be fulfilled by all links corresponding to any pair of columns of $U$. Similar conditions are due to the orthogonality between any two rows of $U$. This implies the total number of $N(N - 1)$ constraints of the form (12), some of which can be dependent [26,35]. However, there is an example of a bistochastic matrix $B$ of order four, which satisfies all chain links condition for all pair of rows, but not for all pair of columns [26,36], so in practice one has to check rows and columns separately. Furthermore, for $N \geq 4$ these conditions for $B$ are only necessary but not sufficient to imply unistochasticity.

![Figure 2: Chain link condition for unistochasticity: a) 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) a bistochastic matrix $B$ not included into $U_3$ for which $A^2 < 0$.](image)

It is comforting to realize that the situation gets simpler for $N = 3$. In this case the chain links relation for the first two columns reduces to the triangle inequality,

$$|L_2 - L_3| \leq L_1 \leq L_2 + L_3, \quad (13)$$

and the first constraint is required if we relax the assumption that the links are ordered decreasingly. Although in general one should check similar relations stemming to other pairs of columns and rows of $U$, in the case $N = 3$ the last column by construction has the right moduli and does not impose any further restrictions [29]. Thus in this case the relation allowing a chain to close is sufficient for unistochasticity [9,28], and explicit formulae for the phases $\phi_{ij}$ are
provided below. If there is a bistochastic matrix $B$ such that in all relations \[ \text{equality takes place, the phases} \; \theta_j \; \text{are equal to zero or to} \; \pi. \; \text{Thus the corresponding matrix} \; U \; \text{is orthogonal, which means that} \; B \; \text{is orthostochastic. It is easy to show that a matrix} \; B \; \text{belongs to the boundary of the set} \; \mathcal{U}_3 \; \text{if and only if} \; B \; \text{is orthostochastic. The set} \; \mathcal{U}_3 \; \text{forms a 4–dimensional subset of} \; \mathcal{B}_3 \; \text{of a positive measure, while the set of orthostochastic matrices, at the boundary of} \; \mathcal{U}_3, \; \text{is three dimensional} \; [27,28].}

For any given $B$ of order three it is straightforward to check whether link conditions \[ \text{(13)} \; \text{are fulfilled, so that} \; B \; \text{is unistochastic. For instance nine short edges, (of length 2) of the Birkhoff polytope belong to} \; \mathcal{U}_3, \; \text{while the long edges (of length} \; \sqrt{6} \; \text{) do not belong to this set.}

Let us take three such edges spanned by $P, P^2$ and $I$, which form the equilateral triangle. At this plane, the set of orthostochastic matrices, for which $L_1 = L_2 + L_3$, forms a deltoid – see Fig. 3b. This figure also called 3–hypocycloid, may be obtained by sliding a circle of radius $1/3$ inside the unit circle. Thus the set of unistochastic matrices corresponds to the interior of the deltoid and is not convex. This set contains the maximal unistochastic ball of radius $r = \sqrt{2}/3$ centered at $W$, which touches the boundary at the deltoid.

Incidentally, the very same figure is related in a different way with the set of unistochastic matrices of order three. The spectra of these matrices are real or belong to the deltoid inscribed into the unit disk and stemming from the real eigenvalue equal to unity \[ [26].

Figure 3: Non convex set $\mathcal{U}_3$ of unistochastic matrices forms a proper subset $\mathcal{B}_3$. a) Deltoid obtained by the cross-section of $\mathcal{U}_3$ along the plane spanned by the equilateral triangle $\triangle(P, P^2, I)$, b) a similar cross-section along totally orthogonal plane, c) a view 'from above' as in 2c.

Consider a unistochastic matrix $B$ parametrized by \[ 7 \]. The length of the links read

\[ L_1 = \sqrt{b_1b_2}, \quad L_2 = \sqrt{b_3b_4}, \quad L_3 = \sqrt{(1 - b_1 - b_2)(1 - b_3 - b_4)}, \]

and the triangle inequality \[ 13 \] provides the direct condition for unistochasticity. Let us write down the area $A$ of this unitarity triangle with sides $L_1, L_2, \text{and} \; L_3$, and semiperimeter $p = (L_1 + L_2 + L_3)/2$. Making use of the Heron’s formula

\[ A = \sqrt{p(p - L_1)(p - L_2)(p - L_3)}, \]

\[ 15 \]
and substituting \([14]\) we arrive with a compact expression for the squared area \(A^2\).

It will be convenient to work with this quantity multiplied by sixteen,

\[
Q(b) := 4b_1b_2b_3b_4 - (b_1 + b_2 + b_3 + b_4 - 1 - b_1b_4 - b_2b_3)^2 = 16A^2. \tag{16}
\]

Here \(b = \{b_1, b_2, b_3, b_4\}\) represents a vector in \(\mathbb{R}^4\) which determines a bistochastic matrix in parametrization [7]. In fact we can form six unitarity triangles in this way, depending on what pair of columns or rows we wish to choose. Although their shapes differ due to unitarity their area \(A\) is the same [9] so the quantity \(A\) does not change under permutation of the unitary matrix \(U\).

It is easy to see that all chain-links conditions are equivalent to the single condition for unistochasticity,

\[
A^2(B) \geq 0. \tag{17}
\]

In other words \(B(b) \in \mathcal{U}_3\) if and only if \(b \in \Omega\) where

\[
\Omega := \{b \in \mathbb{R}^4 : b_1 \geq 0, b_2 \geq 0, b_1 + b_2 \leq 1, Q(b) \geq 0\}. \tag{18}
\]

Also \(\Omega\) is the closure of the connected component of \(\{b : Q(b) > 0\}\) which contains \((\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3})\) (see [31, Sect. 2]).

We relate these expressions to one of the standard parametrizations of a unitary matrix (see Dit\˘a [25, p. 11]). For any \(U' \in U(3)\) there are diagonal matrices \(D_1, D_2 \in U(3)\) such that \(D_1U'D_2 = U\)

\[
U = \begin{bmatrix}
c_{12}s_{13} & s_{12}c_{13} & s_{12}s_{13} \\
-s_{12}c_{23} - c_{12}s_{13}s_{23} - e^{i\theta}s_{13}s_{23} - e^{i\delta}c_{13}s_{23} - c_{12}s_{13}s_{23} - e^{i\delta}c_{13}s_{23}
\end{bmatrix}. \tag{19}
\]

The parameters are the angles \(\theta_{12}, \theta_{13}, \theta_{23}, \delta\) and \(c_{jk} := \cos \theta_{jk}, s_{jk} := \sin \theta_{jk}, 0 \leq \theta_{jk} \leq \frac{\pi}{2}\) for \(1 \leq j < k \leq 3\). The function [8] determines thus a unistochastic matrix \(B(b) = f(U)\) and its entries read

\[
\begin{align*}
b_1 &= c_{12}, \quad b_2 = s_{12}^2, \quad b_3 = s_{12} c_{23}, \\
b_4 &= c_{12}^2 c_{13}^2 + s_{13}^2 s_{23}^2 + 2c_{12}c_{13}s_{23}s_{13} \cos \delta.
\end{align*} \tag{20}
\]

We explain how these parameters are used to determine the phases \(\phi_{ij}\) for \(2 \leq i, j \leq 3\). We consider only the nondegenerate case in which all entries of \(U\) are nonzero. This implies \(c_{jk} > 0\) and \(s_{jk} > 0\) for \(1 \leq j < k \leq 3\). The equations [20] determine \(\delta\) up to the sign (in \(-\pi < \delta < \pi\)). The fact that \(U\) and \(U' := \left[U_{ij}ight]_{i,j=1}^3\) produce the same values for \(b_1, \ldots, b_4\) causes this ambiguity. We will adopt the normalization \(\text{Im} U_{22} > 0\). This forces \(\sin \delta < 0, \text{Im} U_{32} < 0, \text{Im} U_{23} < 0\) and \(\text{Im} U_{33} > 0\). The cosines of the phases are computed using the entries in [19]. The phases are related to the (interior) angles of the unitarity triangles derived from columns 1 and 2, and from columns 1 and 3. For the former case, using
the lengths from equation (14) and denoting the angles \( \theta_1, \theta_2, \theta_3 \) by the label on
the opposite side, we find

\[
\cos \theta_3 = \frac{L_1^2 + L_2^2 - L_3^2}{2L_1 L_2} = \frac{b_1 b_2 + b_3 b_4 - (1 - b_1 - b_3) (1 - b_2 - b_4)}{2 \sqrt{b_1 b_2 b_3 b_4}}
\]

\[
= - \cos \phi_{22},
\]

\[
\cos \theta_2 = \frac{L_1^2 + L_3^2 - L_2^2}{2L_1 L_3} = \frac{b_1 b_2 + (1 - b_1 - b_3) (1 - b_2 - b_4) - b_3 b_4}{2 \sqrt{b_1 (1 - b_1) (1 - b_2 - b_4)}},
\]

\[
= - \cos \phi_{32}.
\]

From the conditions \( \sin \phi_{22} > 0 \) and \( \sin \phi_{32} < 0 \) we obtain \( \phi_{22} = \pi - \theta_3 \) and
\( \phi_{32} = \theta_2 - \pi \) (note the interior angles satisfy \( 0 < \theta_1, \theta_2, \theta_3 < \pi \) thus \( 0 < \phi_{22} < \pi \)
and \( -\pi < \phi_{32} < 0 \). By interchanging columns 2 and 3 we find the remaining
nonzero phases (we use matrix entry notation from equation (14) for more concise
statements):

\[
\cos \phi_{23} = -\frac{b_1 B_{13} + b_3 B_{23} - B_{31} B_{33}}{2 \sqrt{b_1 b_3 B_{13} B_{23}}}, \quad \sin \phi_{23} < 0;
\]

\[
\cos \phi_{33} = -\frac{b_1 B_{13} + B_{31} B_{33} - b_3 B_{23}}{2 \sqrt{b_1 B_{13} B_{31} B_{33}}}, \quad \sin \phi_{33} > 0.
\]

For the example where each \( b_i = \frac{1}{3} \) the unitarity triangle is equilateral, each
\( \theta_i = \frac{\pi}{3} \) and the above equations give \( \phi_{22} = \frac{2\pi}{3} = \phi_{33} \) and \( \phi_{32} = -\frac{2\pi}{3} = \phi_{23} \).

We return to the consideration of the Jarlskog invariant. A straightforward
computation yields

\[
2 \text{Re} (U_{11} U_{22} U_{12}^* U_{21}^*) = 1 - b_1 - b_2 - b_3 - b_4 + b_1 b_4 + b_2 b_3,
\]

(21)

and thus the square of the Jarlskog invariant \( J \) reads

\[
[\text{Im} (U_{11} U_{22} U_{12}^* U_{21}^*)]^2 = b_1 b_3 b_4 - [\text{Re} (U_{11} U_{22} U_{12}^* U_{21}^*)]^2
\]

(22)

Substituting expression (21) into above equation and comparing the outcome with (16) we see that

\[
J^2 = \frac{1}{4} Q(b) = 4A^2
\]

(23)

Thus the squared Jarlskog invariant, proportional to the squared area of the
unitarity triangle, may also be defined as in (16) for an arbitrary bistochastic
matrix \( B \in \mathcal{B}_3 \). For simplicity we shall write according to the context \( J = J(U) \)
or \( J = J(B) = J(B(f(U))) \), as it should not lead to misunderstanding.

The squared Jarlskog invariant, \( J^2 \), is equal to zero if and only if \( B \) is
orthostochastic, so there exists an orthogonal matrix \( O \), such that \( B_{ij} = O_{ij}^2 \).

Following Haagerup [37] we shall call two unitary matrices \( U_1 \) and \( U_2 \) equivalent,
written \( U_1 \sim U_2 \), if there exist two diagonal unitary matrices \( D_1 \) and \( D_2 \)
and two permutation matrices \( P_1 \) and \( P_2 \) such that

\[
U_2 = D_1 P_1 \cdot U_1 \cdot P_2 D_2.
\]

(24)
Observe that due to permutation matrices this relation is more general than the relation (2).

Since multiplication by phases or permutations do not vary the area of the unitarity triangle we see that the squared Jarlskog invariant of two equivalent unitaries are equal, if \( U_1 \approx U_2 \) then \( J^2(U_1) = J^2(U_2) \). Going back to the set of unitary matrices \( U(3) \) we see that \( J(U) = 0 \) if \( U \) is orthogonal, or more generally, if \( U \) is equivalent to an orthogonal matrix, \( U = D_1 P_1 \cdot O \cdot P_2 D_2 \).

Thus \( J^2(U) \) measures to what extend the matrix \( U \) can be transformed into an orthogonal matrix by means of enphasing and permutations. It is easy to see that \( J \) is maximal if the unitarity triangle is equilateral, \( L_1 = L_2 = L_3 = 1/3 \) so that \( J_{\text{max}}^2 = 1/108 \) – see Appendix A. This is the case for the flat matrix \( W \) of van der Waerden, which corresponds to the unitary Fourier matrix

\[
F_3 = \frac{1}{\sqrt{3}} \begin{bmatrix}
1 & 1 & 1 \\
1 & \exp(i \cdot 2\pi/3) & \exp(i \cdot 4\pi/3) \\
1 & \exp(i \cdot 4\pi/3) & \exp(i \cdot 2\pi/3)
\end{bmatrix},
\]

which is an example of a complex Hadamard matrix of order three. Such a unitary matrix \( H \) of size \( N \) is distinguished by an extra condition that all its complex entries have the same modulus, \( |H_{ij}|^2 = 1/N \) for \( i, j = 1, \ldots, N \) [38].

Any complex Hadamard matrix \( H \) of order three is known to be equivalent to the Fourier matrix \( F_3 \) [37], which implies that \( J^2(H) = J_{\text{max}}^2 \). From this perspective the set of Hadamard matrices is maximally distant from the set of orthogonal matrices. While complex Hadamard matrices correspond to the flat bistochastic matrix \( W \), located at the center of the set of unistochastic matrices, the boundary of which is formed by the set of orthostochastic matrices.

On the other hand for any bistochastic matrix which is not unistochastic, the quantity \( J^2 = Q/4 \) defined by (13) is negative. Since \( Q = 16A^2 \) one might say that in this case the area of the unitarity triangle is imaginary, since the three segments \( L_i \) cannot be closed to form a triangle. Among all bistochastic matrices of size three the quantity \( Q \) is the smallest for the matrix \( B_S \) of Schur (1) for which \( Q = -1/16 \), see Appendix A. Indeed, looking at Fig. 3a we see that \( B_S \) is such a point of the Birkhoff polytope \( B_3 \), for which the distance to the set \( U_3 \) of unistochastic matrices, represented by the gray deltoid, is maximal.

### 4 The volume of the set of unistochastic matrices of order three.

Since unistochastic matrices of order three are used in various branches of theoretical physics, several geometric properties of the set \( U_3 \) of these matrices were studied in [27]. In particular, in that work the volume of this set was estimated numerically. In this section we shall improve these findings by deriving an analytical formula for this volume. To this end we need to introduce probability measures into the set of unistochastic matrices. A first natural choice will be:

a) the flat (Lebesgue) measure \( \mu_{3/2} \) used before in [27].
The above notation is due to the fact that this measure belongs to a one-
parameter class of measures $\mu_k$, defined below. In general any probability mea-
sure in the set $U(N)$ of unitary matrices induces by function (5) a measure into
the set $\mathcal{U}_N$. Thus we will distinguish the case
b) the measure $\mu_1$ induced by the Haar measure on $U(3)$.

This measure leads to the *unistochastic ensemble* or random unistochastic
matrices [26]. Since it is related to the unitarily invariant measure on the flag
manifold $U(3)/[U(1)^3]$ it was called *flag–manifold measure* in [30].

Suppose now that $dm$ denotes the normalized Haar measure on $U(3)$, $db$
is Lebesgue measure on $\mathbb{R}^4$, and $g$ is a continuous function on $[0, 1]^4 \subset \mathbb{R}^4$.
Due to [31, Theorem 2.1] one may relate the integrals over the space of unitary
matrices $U(3)$ and over the set $\Omega \in \mathbb{R}^4$, see eq. (18), which determines the set
of unistochastic matrices,

$$\int_{U(3)} g\left(|U_{11}|^2, |U_{12}|^2, |U_{21}|^2, |U_{22}|^2\right) dm = \frac{2}{\pi} \int_{\Omega} g(b_1, b_2, b_3, b_4) Q(b)^{-1/2} db.$$  \hspace{1cm} (26)

The function $g(b_1, b_2, b_3, b_4) = g(b)$ determines the function $g(B)$ defined on the
entire set $\mathcal{U}_3$ of unistochastic matrices.

We introduce new coordinates $(b_1, s, t, r)$ with

\begin{align*}
  b_2 &= s (1 - b_1), \\
  b_3 &= t (1 - b_1), \\
  b_4 &= (1 - s)(1 - t) + b_1 st + 2r \sqrt{b_1 st (1 - s)(1 - t)}. \quad (27)
\end{align*}

Note that $s = c^2_{21}, t = c^2_{13}, r = \cos \delta$ in terms of the matrix $U$.

Then $Q = 4b_1 (1 - b_1)^2 s (1 - s) t (1 - t) (1 - r^2)$, and $\Omega$ corresponds to

$$\left\{ (b_1, s, t, r) : (b_1, s, t) \in [0, 1]^3, -1 \leq r \leq 1 \right\}.$$ 

The Jacobian for the change-of-variables is

$$\left| \frac{\partial (b_1, b_2, b_3, b_4)}{\partial (b_1, s, t, r)} \right| = 2 (1 - b_1)^2 \sqrt{b_1 s (1 - s) t (1 - t)}.$$ 

Making use of the expressions derived in [31] we may write an explicit form
for an integral of a continuous function $g$ with respect to the measure $Q(b)^{k-3/2}$
for arbitrary $k > 1/2$. Hence it is natural to introduce a one-parameter family
of measures $\mu_k$ on $\mathcal{U}_3$ which satisfy

$$\int_{\mathcal{U}_3} g(B) d\mu_k = \int_{\Omega} g(b) Q(b)^{k-3/2} db.$$  \hspace{1cm} (28)

Thus the case $k = 3/2$ corresponds to the flat (Lebesgue) measure on $\mathcal{U}_3$,
while for $k = 1$ this expression reduces to (26), so $\mu_1$ represents the measure
induced by the Haar measure on $U(3)$. 

13
Integral at the right hand side of (28) can be rewritten as
\[
\int_\Omega g(b_2) Q(b) k^{-\frac{3}{2}} db = \int_0^1 \int_0^1 \int_0^1 \int_{-1}^1 g(b_1, s (1-b_1), t (1-b_1), b_4) (29)
\times b_2^{-k-1} (1-b_1)^{2k-1} [4s (1-s) t (1-t)]^{k-1} (1-r^2)^{k-\frac{3}{2}} dr ds dt db,
\]
where \(b_4\) is given by (27). Setting \(g = 1\) and using the standard beta integrals we determine the normalization constant \([31, \text{Prop. 3.2}]\)
\[
h_k := \int_\Omega Q(b) k^{-\frac{3}{2}} db = \frac{\pi (k)^3}{(2k-1) \Gamma (3k)}, \quad k > \frac{1}{2}, \tag{30}
\]
Suppose \(n = 1, 2, 3, \ldots\) then
\[
h_n = \frac{\pi ((n-1)!)^3}{(2n-1) (3n-1)!!}, \quad n = 1, 2, 3, \ldots
\]
\[
h_{n+1/2} = \frac{\pi^2}{2n} \left[ \left(\frac{1}{2}\right)_n \right]^2 \frac{1}{2n+1}, \quad n = 1, 2, 3, \ldots
\]
where \((x)_n := \prod_{i=1}^n (x + i - 1)\) stands for the Pochhammer symbol. In particular,
\[
h_{3/2} = \frac{\pi^2}{3 \cdot 5 \cdot 7} = \frac{\pi^2}{105}, \tag{31}
\]
gives the volume of \(U_3\) considered as a subset of \(\mathbb{R}^4\). This is multiplied by 9 to produce the volume relative to \(\mathbb{R}^9\) [5] – see Appendix [A].

Thus the ratio of the 4-dimensional volume of \(U_3\) in \(\mathbb{R}^9\) to the volume of all bistochastic matrices is
\[
\frac{\text{vol}(U_3)}{\text{vol}(B_3)} = 9 \times \frac{\pi^2}{105} / \left(\frac{9}{8}\right) = \frac{8 \pi^2}{105} = 0.751969...
\]
This is in agreement with the outcome of earlier numerical calculations [27, eqn. (24)] which were based on roughly \(10^7\) sample points and yielded 0.7520±0.0005.

For completeness let us add that the volume of the set \(U_3\) with respect to the flag–manifold measure \(\mu_1\) reads \(h_1 = \pi/2\).

5 Mean entropy of a \(N = 3\) unistochastic matrix.

The Shannon entropy of an \(N\)-point probability vector \(p = (p_1, \ldots, p_N)\) is defined by
\[
S(p) = - \sum_{i=1}^N p_i \ln p_i. \tag{33}
\]
This quantity measures to what extent the vector is mixed and varies from 0 for any pure vector \((1, 0, \ldots, 0)\) to \(\ln N\) for the maximally mixed vector \(p_* = \)


In an analogous way one defines the entropy of a bistochastic matrix
\[ S(B) = -\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} B_{ij} \ln B_{ij}, \tag{34} \]
equal to the average entropy of its rows (or columns). For any permutation matrix \( P \) this entropy is equal to zero while its maximum value \( \ln N \) is attained at the flat matrix \( W \). The mean entropy of a bistochastic matrix was considered by Slomczyński \[39\] and later analyzed in \[26,40\].

To derive an expression for the average entropy of a unistochastic matrix with respect to any probability measure on \( U_3 \) bi-invariant under the symmetric group \( S_3 \) we observe that is equal to the expected value of \(-3B_{11} \ln B_{11}\). This can be computed with respect to the probability measure
\[ h^{-1} k Q(\cdot) b^{-\frac{3}{2}} db \]
on \( \Omega \).

Let us denote the mean entropy by \( \langle S \rangle_{\Omega,k} \), where parameter \( k \) labels the measure defined in \(28\). From \(29\) specialized to functions of \( b_1 \) we obtain
\[ \langle S \rangle_{\Omega,k} = \frac{-3\Gamma(3k)}{\Gamma(k) \Gamma(2k)} \int_0^1 b_1^k \ln b_1 (1-b_1)^{2k-1} db_1 \]
\[ = \psi(3k+1) - \psi(k+1), \tag{35} \]
where the digamma function reads \( \psi(x) := \frac{\Gamma'(x)}{\Gamma(x)}, x > 0 \).

The recurrence relation \( \psi(x+1) = \psi(x) + 1/x \) is used in the following: Suppose \( n = 1, 2, 3, \ldots \) then
\[ \langle S \rangle_{\Omega,n} = \sum_{j=n+1}^{3n} \frac{1}{j}, \quad \langle S \rangle_{\Omega,n+\frac{1}{2}} = \sum_{j=n+1}^{3n+1} \frac{2}{2j+1}. \tag{36} \]

In particular the average entropy for the flag-manifold measure \( \mu_1 \) gives \( \langle S \rangle_{\Omega,1} = \frac{5}{6} = 0.833\ldots \). This result coincides with the average entropy of random complex vectors \[41,42\] which form a unitary matrix.

The mean entropy with respect to the flat measure \( \mu_{3/2} \) on \( U_3 \) reads \( \langle S \rangle_{\Omega,\frac{3}{2}} = 2 \left( \frac{1}{5} + \frac{1}{7} + \frac{1}{9} \right) = \frac{286}{315} = 0.90793651\ldots \). This quantity was approximated as 0.908 in \[27, eqn. (25)\].

These data can be compared with the maximal entropy \( S_{\text{max}} = \ln 3 \approx 1.099 \), characteristic of the flat matrix \( W \) of van der Waerden.

For comparison let us now compute the mean entropy with respect to the Lebesgue measure averaged over the entire set \( \mathcal{B}_3 \) of bistochastic matrices.

Straightforward calculation allows us to integrate functions of \( b_1, b_2 \) over \( \mathcal{B}_3 \) with respect to the flat measure \( db := \prod_{i=1}^{3} db_i \). In general, we consider an arbitrary function \( f \), integrable on the triangle with vertices \{\( (0,0), (1,0), (0,1) \}\). The result is
\[ \int_{\mathcal{B}_3} f(b_1, b_2) \, db = \int_0^1 db_1 \int_0^{1-b_1} f(b_1, b_2) (b_1 b_2 + (b_1 + b_2) (1 - b_1 - b_2)) \, db_2. \tag{37} \]
The corollary stated in the Appendix specializes this to \( \int_{0}^{1} f(b_1) \, db \) and allows us to find the volume of the Birkhoff polytope. Thus relative to \( \mathbb{R}^9 \) one has \( \text{vol}_4(B_3) = \frac{9}{8} \).

Formula (37) can also be used to compute the average entropy, equal to the expected value of \(-3b_1 \ln b_1\). The analytic result

\[
8 \int_{B_3} (-3b_1 \ln b_1) \, db = \frac{53}{60}.
\]

agrees with the numerical estimation 0.883 obtained earlier in [27]. Note that this number is smaller than the mean entropy \( \langle S \rangle_{\Omega, \frac{3}{2}} \) averaged over the set of unistochastic matrices, since these bistochastic matrices which do not belong to \( \mathcal{U}_3 \) are located close to the boundary of the Birkhoff polytope and are characterized by small entropy.

Let us now turn to the generalized entropy defined for any probability vector \( \{p_i : 1 \leq i \leq N\} \)

\[
S_q := \frac{1}{q-1} \sum_{i=1}^{N} (p_i - p_i^q).
\]

(38)

The parameter \( q \neq 1 \) is assumed to be non negative. In the limiting case the generalized entropy converges to the standard (Shannon) entropy, \( \lim_{q \to 1} S_q = -\sum_{i=1}^{N} p_i \ln p_i \).

Applying definition (38) to a unistochastic matrix \( B \in \mathcal{U}_3 \) similarly to the ordinary case we let

\[
S_q = \frac{1}{3(q-1)} \sum_{i=1}^{3} \sum_{j=1}^{3} (B_{ij} - B_{ij}^q).
\]

We compute the expected value of this expression with respect to \( h_{k}^{-1} Q(b)^{k-\frac{3}{2}} \, db \) on \( \Omega \) for \( k > \frac{1}{2} \).

It is the same as the expected value of \( \frac{3}{q-1} (B_{11} - B_{11}^q) \), indeed

\[
\langle S_q \rangle_{\Omega, k} = \frac{3\Gamma(3k)}{(q-1) \Gamma(k) \Gamma(2k)} \int_{0}^{1} (b_1 - b_1^q) b_1^{k-1} (1 - b_1)^{2k-1} \, db_1
\]

\[
= \frac{1}{q-1} \left( 1 - \frac{3\Gamma(k+q) \Gamma(3k)}{\Gamma(k) \Gamma(3k+q)} \right).
\]

When \( k \) is an integer number \( n \) or a half-integer \( k = n + \frac{1}{2} \) this expression is a rational function of \( q \) and can be expressed with help of the Pochhammer symbol \((x)_n\) defined above,

\[
\langle S_q \rangle_{\Omega, n} = \frac{1}{q-1} \left( 1 - \frac{(3n)!}{n! (q+n)_{2n}} \right),
\]

(39)

\[
\langle S_q \rangle_{\Omega, n+\frac{1}{2}} = \frac{1}{q-1} \left( 1 - \frac{3 \left( \frac{1}{2} + n \right)_{2n+1}}{(q + \frac{1}{2} + n)_{2n+1}} \right)
\]

(40)
In particular, taking \( n = 1 \) we arrive at handy expressions for the mean gene-
realized entropies averaged over the Haar measure and flat measure respectively,
which allow for an explicit partial fraction expansion,
\[
\langle S_q \rangle_{\Omega,1} = \frac{q + 4}{(q + 1)(q + 2)} = \frac{3}{q+1} - \frac{2}{q+2},
\]
\[
\langle S_q \rangle_{\Omega,3/2} = \frac{2(4q^2 + 34q + 105)}{(2q + 3)(2q + 5)(2q + 7)} = \frac{63}{4(2q + 3)} - \frac{45}{2(2q + 5)} + \frac{35}{4(2q + 7)}.
\]

For completeness we provide also an expression for the generalized entropy
averaged over the set \( B_3 \) with respect to the flat measure obtained with help of
Corollary 2
\[
\langle S_q \rangle_{B_3} = \frac{2}{q + 1} + \frac{4}{q + 2} - \frac{9}{q + 3} + \frac{4}{q + 4}.
\]

These entropies characterize well the distribution of matrices generated by
these measures. In particular, a comparison of both expressions in (42) shows
that the Haar measure on \( U(3) \) populates the region close to the boundary of \( U_3 \)
more densely then the vicinity of the flat matrix \( W \) around its center. Since the
squared Jarlskog invariant \( J^2 \) is by construction equal to zero at the boundary
of \( U_3 \), we may expect that its mean value over the flat measure \( \mu_1 \) is smaller
than the average with respect to the Haar measure \( \mu_{3/2} \). As shown in the next
section this is indeed the case.

6 Distribution of the Jarlskog invariant

The value of the Jarlskog invariant and its square at a cross-section of the set
\( U_3 \) of unistochastic matrices is shown in Fig. 4. Recent papers of Gibbons
et al. [12, 30] analyzed squared Jarlskog invariant \([8, 9]\) averaged over several
probability measures on the set of unitary matrices. In particular these authors
computed the expectation value, \( \langle J^2 \rangle \), averaged over the 'flag manifold' measure
induced by the Haar measure on \( U(3) \) and analyzed numerically the probability
distribution \( P(|J|) \) with respect to this measure. In this section we proceed one
step further and derive an analytical formula for this probability distribution.

We shall start computing the moments of the distribution of the variable
\( Q = 4J^2 \) defined in (10) as a function of a random unistochastic matrix \( B \). This
task is rather simple, since we can express the moments of \( Q \) with respect to
any measure \( \mu_k \) by the coefficients \( h_k \) defined in (30),
\[
\langle Q^n \rangle_k = \frac{h_{k+n}}{h_k} = 3^{-3n} \frac{(k - \frac{1}{2}) \Gamma (k + \frac{1}{2}) \Gamma (k + \frac{3}{4}) \Gamma (k + n)^2}{(k + n - \frac{1}{2}) \Gamma (k + n + \frac{1}{4}) \Gamma (k + n + \frac{3}{4}) \Gamma (k)^2}
\]
\[
= 3^{-3n} \frac{(k - \frac{1}{2})(k)_n^2}{(k + n - \frac{1}{2}) (k + \frac{1}{4})_n (k + \frac{3}{4})_n},
\]
where \( k \) determines the measure (28) while \( n = 0, 1, 2, \ldots \).
Figure 4: The absolute value of the Jarlskog invariant $|J|$ a), and its square $J^2$ b), at the cross-section of $\mathcal{U}_3$ along the plane formed by two permutation matrices and the identity. Dark color denotes high values of $|J|$ and $J^2$. The maximum is achieved at the van der Waerden matrix $W$ located the center of the deltoid. Note that outside the deltoid $J^2 < 0$ and the bistochastic matrix is not unistochastic.

Setting $k = 1$ and $n = 1$ we find that the mean squared Jarlskog invariant, averaged over the Haar measure reads $\langle J^2 \rangle_1 = \langle Q/4 \rangle_1 = 1/720 = 1.389 \times 10^{-3}$ in consistence with [30, eqn. (75)]. For comparison note that the average over the flat measure yields a larger value, $\langle J^2 \rangle_{3/2} = 3/1144 = 2.622 \times 10^{-3}$. In general the flat measure favors larger values of $|J|$ as it is shown in Fig.5.

Having at our disposal the complete set of the moments of $Q$ we will determine the exact distribution function $P(Q)$ in terms of hypergeometric and related functions. To avoid nuisance factors in the calculations we will consider the random variable $X := 27Q = 108J^2$ so that $X$ takes values in $[0, 1]$.

We know that the following relations hold for any $k > 1/2$ and $n = 0, 1, 2, \cdots$

$$\Gamma (\alpha + \beta) \Gamma (\alpha) \Gamma (\beta) \int_0^1 x^{n+\alpha-1} (1-x)^{\beta-1} \, dx = \frac{\Gamma (\alpha)}{(\alpha + \beta)_n}, \quad (\alpha, \beta > 0).$$

In view of the expression (43) for the moments the above relations allow us to find an alternative representation of the desired probability distribution $P(X)$.

18
Let $X_0, X_1, X_2$ be independent random variables with the densities

$$f_0(x) = \left(k - \frac{1}{2}\right)x^{k-\frac{3}{2}}, \quad (44)$$

$$f_1(x) = \frac{\Gamma\left(k + \frac{1}{3}\right)}{\Gamma(k)\Gamma\left(\frac{1}{3}\right)}x^{k-1}(1-x)^{-2/3}, \quad (45)$$

$$f_2(x) = \frac{\Gamma\left(k + \frac{2}{3}\right)}{\Gamma(k)\Gamma\left(\frac{2}{3}\right)}x^{k-1}(1-x)^{-1/3}, \quad (46)$$

respectively, each being defined on $0 \leq x \leq 1$.

Then $X$ has the same moments and the same probability distribution as the product $X_0X_1X_2$. This step, justified in Appendix B enables us to arrive at the key result of this section: an explicit expression for the probability distribution for $X = 108J^2$, where $J$ denotes the Jarlskog invariant of a random unitary matrix generated according to the measure $\mu_k$.

$$P_k(X) = c_k\left(k - \frac{1}{2}\right)X^{k-3/2} \int_0^1 x^{1/3} 2F_1\left(\frac{1}{3}, \frac{2}{3}; 1; 1-t\right)t^{-1/2}dt. \quad (47)$$

It is assumed here that $0 < X \leq 1$ and $k > 1/2$, the symbol $2F_1$ stands for the hypergeometric function, while the normalization constant reads

$$c_k := \frac{\Gamma\left(k + \frac{1}{3}\right)\Gamma\left(k + \frac{2}{3}\right)}{\Gamma(k)^2}, \quad (48)$$

In Appendix B we made use of this expression to determine an explicit expansion for the density function $P_k(X)$. These results allowed us to show the distributions $P_1(X)$ and $P_{3/2}(X)$ in Fig. 5. Observe that in the case $k = 1$ we obtain an expression for the distribution of $|J|$, considered as a function of a random unitary matrix $U$ distributed with respect to the Haar measure on $U(3)$ random variable on For small values of $|J|$ this distribution behaves as $P(|J| \sim \alpha|J|^\lambda$ with $\lambda = 0$ and $\alpha = 8\pi \approx 25.133$.

To make a direct connection with the results of [30], we reproduce here the formula for the integrated probability distribution. For any $0 < y \leq \frac{1}{6\sqrt{3}}$ the probability of finding a unitary matrix $U$ distributed according to the Haar measure on $U(3)$ with $|J(B)|$ less or equal $y$ reads,

$$P_1\{|J| \leq y\} = 8\pi y + \left\{24 \ln \left(\frac{4y^2}{y^2 + 4y^4 + 96y^6 + \cdots}\right) - 72y^2 + 128y^4 + \frac{24384}{5}y^6 + \cdots\right\}.$$  

(49)

In Appendix B we derive this formula obtained for the measure $\mu_1$ on $U_3$ as well as an analogous result for the flat measure $\mu_{3/2}$

$$P_{3/2}\{|J| \leq y\} = 420y^2 - \frac{4480}{\pi}y^3 \cdots + \frac{1680}{\pi}y^3 \ln \left(4y^2\right) + \cdots \quad (50)$$

Both cumulative distribution functions are compared in Fig. 6.
Figure 5: Probability distribution $P_k(|J|)$ of the Jarlskog invariant for random unistochastic matrices generated according to the Haar measure $\mu_1$ and the flat measure $\mu_{3/2}$.

Figure 6: Cumulative probability distribution for the functions plotted in Fig. 5.
Recent experimental data show that the observed value of the Jarlskog invariant reads \[34\]
\[ J_{\text{obs}} = 3.08 \pm 0.16, -0.18 \times 10^{-5} \] \hfill (51)

This concrete number can be now compared with the probability distribution \[39\]. A 95\% confidence interval for $|J|$ is about $[0.00202, 1/6\sqrt{3} \approx 0.096]$, while the probability of getting a value of $|J|$ outside this interval is 5\%.

Moreover, we get an explicit estimate for the probability of obtaining at random a unitary matrix, such that the absolute values of its Jarlskog invariant is smaller then the observed value,
\[ P \{|J| \leq J_{\text{obs}}\} = 7.74 \pm 0.41, -0.45 \times 10^{-4} \] \hfill (52)

The statistical hypothesis that the CKM matrix arises from the probability experiment of producing a random unitary matrix, with respect to Haar measure on $U(3)$, is rejected at the descriptive significance level of 0.08\%.

Another benchmark introduced in Gibbons et al [30], $P_{\text{flag}}'(\{|J| \leq 10^{-4}\} \approx 2.5085 \times 10^{-3}$ is consistent with numerical data obtained in Eq.(92) of that paper. Note that Gibbons et al [30] constructed several probability models for which values as small or smaller than $J_{\text{obs}}$ are more likely.

For comparison we note that the flat measure $\mu_{3/2}$ in the set of bistochastic matrices yields a smaller probability. Using this measure ($k = \frac{3}{2}$) we obtain $P_{\text{fl}}(\{|J| \leq J_{\text{obs}}\} \approx 3.98 \times 10^{-7}$. Indeed this could be viewed as statistical evidence that the transition probabilities in the CKM matrix do arise from a unitary matrix. Specifically the so-called likelihood ratio test applied to the two probability densities for $|J|$ induced by $d\mu_1$ and $\frac{8\pi^2}{105}d\mu_{3/2}$ (the factor $\frac{8\pi^2}{105}$ comes from $P\{Q < 0\} = 1 - \frac{8\pi^2}{105}$ for the flat measure on $B_3$) at $|J| = 3.08 \times 10^{-5}$ results in a factor of about 1200. This value is obtained from the formulas for $f_0(x)$ in Appendix [3]

Let us express the Jarlskog invariant \hfill (10) by the standard parameters \hfill (19) of a unitary matrix of order three,
\[ J(U) = -c_{12}c_{23}s_{13}^2s_{23}s_{13}\sin\delta \] \hfill (53)

Observed value of the Jarlskog invariant for the CKM matrix does not imply that the CP violating phase $\delta_{\text{CKM}}$ had to be small. In fact $\delta_{\text{CKM}} \in [62^\circ, 100^\circ]$ so even the value $\pi/2$ is not ruled out – see e.g. [43]. Hence the small value \hfill (51) is due to the angles $\theta_{ij}$ in \hfill (19), which determine the bistochastic matrix.

Thus we are going to conclude this section with a simple statistical statement: The CKM matrix should not be considered as a generic unitary matrix drawn at random with respect to the Haar measure on $U(3)$. Furthermore, the matrix $B_{\text{CKM}}$ of squared entries of $V_{\text{CKM}}$ is rather unlikely to be an ordinary bistochastic matrix generated at random with respect to the flat measure in this set.
7 Concluding Remarks

In this work we have analyzed the Birkhoff polytope $\mathcal{B}_3$ of $N = 3$ bistochastic matrices and its subset $\mathcal{U}_3$ of unistochastic matrices. This set contains these bistochastic matrices which arise from squared moduli of entries of a unitary matrix. We have improved the result of [27] by computing the exact volume of $\mathcal{U}_3$ with respect to the flat (Lebesgue) measure and found that it takes more than three quarters of the volume of the Birkhoff polytope $\mathcal{B}_3$.

We have introduced a one-parameter family of probability measures $\mu_k$ into the set $\mathcal{U}_3$ of unistochastic matrices. Among the measures (28) the distinguished ones are the uniform (flat) measure $\mu_{3/2}$ and the measure $\mu_1$, induced by the Haar measure on $U(3)$. Furthermore, the measure $\mu_k$ obtained in the limit $k \to 1/2$ coincides with the measure induced by the Haar measure on the orthogonal group $O(3)$.

We derived explicit formulae which allow us to compute expectation values of a smooth function of an entry of $B$ with respect to these measures. In this way we derived exact expressions for the mean entropy and the generalized entropy of a random unistochastic matrix with respect to the measures $\mu_k$. These values can serve as a reference values in studying properties of concrete unitary and bistochastic matrices of order three, used in the theory of quantum information.

In high energy physics and the theory of CP symmetry breaking one works with unitary matrices of order 3 and characterizes them by the Jarlskog invariant (10). Computing all the moments of the squared Jarlskog invariant $J^2$ with respect to the measure $\mu_k$ we could represent the probability distribution $P_k(J)$ as an integral (47) of the hypergeometric function $2 F_1$. Expanding this function in a series and integrating it term by term we arrived at an explicit representation of the desired probability distributions. In particular, working with the Haar measure $\mu_1$ we could derive analytical results on the distribution $P_1(|J|)$ consistent with the numerical results earlier obtained in Gibbons et al. [30]. Our results support then the observation, that the unitary CKM matrix $V_{CKM}$, which describes the violation of the CP symmetry, should not be regarded as a generic unitary matrix of order 3.

Acknowledgements:

It is a pleasure to thank I. Bengtsson and to W. Tadej for numerous stimulating discussions and helpful correspondence. We acknowledge financial support by the the special grant number DFG-SFB/38/2007 of Polish Ministry of Science and Higher Education and an European Research project COCOS (KŻ).

A Extreme values of the parameter $Q$

Consider a bistochastic matrix $M$ of order three parametrized by (7) and the function $Q(M)$ defined in (16). The aim of this section is to show that for any $M \in \mathcal{B}_3$ this function takes values in $[-1/16, 1/27]$. Note that if the matrix is
unistochastic, \( M \in \mathcal{U}_3 \) than \( Q \) is non negative and is proportional to the squared area of the unitarity triangle.

It is straightforward to show that \( Q \) is invariant under transposition and permutation of rows or columns (for example replacing \((b_2, b_4)\) by \((1 - b_1 - b_2, 1 - b_3 - b_4)\)). Let us first introduce parameters \( b_1, s, t, x \) with \( 0 \leq b_1, s, t \leq 1 \) and conditions on \( x \) to be determined. Motivated by the unistochastic situation let

\[
\begin{align*}
    b_2 &= (1 - b_1) s, \\
    b_3 &= (1 - b_1) t, \\
    b_4 &= (1 - s) (1 - t) + b_1 st + x.
\end{align*}
\]

Four more conditions must be satisfied for \( M (b) \) to be bistochastic (the inequalities \( M (b)_{13} \geq 0 \) and \( M (b)_{31} \geq 0 \) are already satisfied). The simultaneous inequalities \( M (b)_{23} \geq 0 \) and \( M (b)_{32} \geq 0 \) are equivalent to

\[
\begin{align*}
    x &\leq \min (u_1, u_2), \\
    u_1 &:= s (1 - t) + b_1 t (1 - s), \\
    u_2 &:= t (1 - s) + b_1 s (1 - t)
\end{align*}
\]

and \( \{ M (b)_{22} \geq 0, M (b)_{33} \geq 0 \} \) is equivalent to

\[
\begin{align*}
    -x &\leq \min (\ell_1, \ell_2), \\
    \ell_1 &:= (1 - s) (1 - t) + b_1 st, \\
    \ell_2 &:= st + b_1 (1 - s) (1 - t).
\end{align*}
\]

With these parameters

\[
Q (b) = Q' (b_1, s, t, x) := - (1 - b_1)^2 \left( x^2 - 4 b_1 st (1 - s) (1 - t) \right).
\]

For fixed \( b_1, s, t \) this is decreasing in \( x^2 \); thus the maximum value occurs at \( x = 0 \) and then maximizing over \( b_1, s, t \) we obtain \( Q (b) = \frac{1}{27} \) when \( b_1 = \frac{1}{3}, s = \frac{1}{2}, t = t \) (that is, \( M (b)_{ij} = \frac{1}{4} \) for \( 1 \leq i, j \leq 3 \)).

Next we show that the minimum value of \( Q \) on \( \mathcal{B}_3 \) is \( -1 \), achieved at the Schur matrix (1): by permutations of rows or columns, and then transposition of \( M \), if necessary, we may assume \( 0 \leq t \leq s \leq \frac{1}{2} \). In this triangle the bounds \( -\ell_2 \leq x \leq u_2 \) apply. Rather than considering \( \min (Q' (b_1, s, t, -\ell_2), Q' (b_1, s, t, u_2)) \) in this region we will minimize \( Q' (b_1, s, t, -\ell_2) \) in the triangle bounded by \( t = 0, t = s, t = 1 - s \) (vertices \((0, 0), (\frac{1}{2}, \frac{1}{2}), (1, 0)\)); this works because \( Q' (b_1, s, t, u_2 (b_1, s, t)) = Q' (b_1, 1 - s, t, -\ell_2 (b_1, 1 - s, t)) \) (writing \( \ell_2, u_2 \) as functions of \((b_1, s, t)\)). Then

\[
Q' (b_1, 1 - s, t, -\ell_2 (b_1, 1 - s, t)) = (1 - b_1)^2 (b_1 (1 - s) (1 - t) - st)^2.
\]

As a function of \((s, t)\), \( Q' \) can not have an interior minimum, so it suffices to check the edges of the triangle. On the edge \( t = 0 \) we have \( Q' = (1 - b_1)^2 b_1^2 (1 - s)^2 \) with minimum value of \( -\frac{1}{27} \) at \( b_1 = \frac{1}{3}, s = 0 \).

23
On the edge \( t = 1 - s \) we obtain \( Q' = -(1 - b_1)^4 s^2 (1 - s)^2 \) with minimum value of \(-\frac{1}{16}\) at \( b_1 = 0, s = \frac{1}{2} = t \) (the Schur matrix for \( x = -\frac{1}{4} \)).

On the edge \( s = t \), \( Q' = -(1 - b_1)^2 (b_1 (1 - s)^2 - s^2)^2 \). This function has no interior minimum on the interval \( 0 \leq s \leq \frac{1}{2} \). The endpoints \( s = 0 \) and \( s = \frac{1}{2} \) have already been considered.

There is a neat formula for the integral of functions of \((b_1, b_2)\) over \( B_3 \) with respect to the flat measure \( db := \prod_{i=1}^4 db_i \). The derivation of the formula involves adding over the four regions formed in the unit square by the lines \( s = t, s + t = 1 \).

**Proposition 1** Let \( f \) be integrable on \( \{(x, y) : x, y \geq 0, x + y \leq 1\} \), then

\[
\int_{B_3} f(b_1, b_2) \, db = \int_0^1 db_1 \int_0^{1-b_1} f(b_1, b_2) (b_1 b_2 + (b_1 + b_2) (1 - b_1 - b_2)) \, db_2.
\]

Observe that the integral kernel is an elementary symmetric function of \((b_1, b_2, 1 - b_1 - b_2)\).

**Corollary 2** Let \( f \) be integrable on \([0, 1]\) then

\[
\int_{B_3} f(b_1) \, db = \frac{1}{6} \int_0^1 f(b_1) (1 + 5b_1) (1 - b_1)^2 \, db_1.
\]

Thus \( \int_{B_3} 1 \, db = \frac{1}{6} \).

When considering \( B_3 \) as a subset of \( \mathbb{R}^9 \) the element of volume \( db = \prod_{i=1}^4 db_i \) is multiplied by 9. The map \( B : \mathbb{R}^4 \to \mathbb{R}^9 \), defined in equation 7 is affine onto a 4-dimensional linear manifold (translate of a subspace) and its Jacobian equals 9, calculated as the square root of the determinant of the Gram matrix of the images of the unit vectors relative to \( B \left( \vec{0} \right) \). For example

\[
b_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix} + b_1 \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} ;
\]

the Gram matrix is

\[
\begin{bmatrix} 4 & 2 & 2 & 1 \\ 2 & 4 & 1 & 2 \\ 2 & 1 & 4 & 2 \\ 1 & 2 & 2 & 4 \end{bmatrix}
\]

and its determinant equals 81.

In this way we may obtain the average entropy and directly derive expression (42) for the generalized entropy. Furthermore, \( 8 \int_{B_3} Q(b) \, db = \frac{1}{168} \), \( 8 \int_{B_3} Q(b)^2 \, db = \frac{1}{90} \), thus the standard deviation reads \( \sigma_Q = 0.01153 \ldots \).

Since the set \( U_3 \) of unistochastic matrices is the subset of \( B_3 \) for which \( Q \geq 0 \), equivalently

\[
|x| \leq R := \left\{ 4b_1 st (1 - s) (1 - t) \right\}^{1/2} ,
\]

\[
24
\]
we see that
\[
\int_{t_b} f(b_1) \, db = \int_0^1 f(b_1) (1 - b_1)^2 \, db_1 \int_0^1 ds \int_0^1 dt \int_{-R}^R dx = \frac{\pi^2}{16} \int_0^1 f(b_1) b_1^{1/2} (1 - b_1)^2 \, db_1,
\]
and \(\int_{t_b} 1 \, db = \pi^2/105\) in agreement with eq. \(61\). Observe that \(R \leq \min(l_1, l_2, u_1, u_2)\) by the (well-known) inequality \(2\sqrt{x+y} \leq x + y\) for \(x, y \geq 0\); this is the reason that the integral extends over \(0 \leq s, t \leq 1\).

**B Jarlskog invariant as a product of three random variables**

In this appendix we show that the product of three random variables \(X_0 X_1 X_2\) introduced in \(44\) has the same probability distribution as the rescaled squared Jarlskog invariant \(X := 27Q = 108J^2\) of random unistochastic matrices generated with respect to the measure \(\mu_k\) defined in \(23\). We shall start by quoting the lemma on probability distribution of a product of two independent random variables.

**Lemma 3** Suppose \(Y_1, Y_2\) are random variables on \([0,1]\) with densities \(g_i\) and c.d.f.’s \(G_i, i = 1, 2\) (that is, \(G_i(x) = \int_0^x g_i(t) \, dt = P\{Y_i \leq x\}, 0 \leq x \leq 1\)). Then the density for \(Y_1 Y_2\) is

\[
\left(k - \frac{1}{2}\right) x^{k-\frac{1}{2}} \int_0^1 g_1(t) t^{1/2-k} \, dt, 0 < x < 1.
\]

Let us apply this lemma, which can be proved by direct integration, to a random variable \(Y_2 = X_0\) distributed as in \(44\).

**Corollary 4** If \(g_2(t) = (k - \frac{1}{2}) t^{k-\frac{3}{2}}\) with \(k > \frac{1}{2}\) then the density of \(Y_1 Y_2\) is

\[
\left(k - \frac{1}{2}\right) x^{k-\frac{3}{2}} \int_0^1 g_1(t) t^{1/2-k} \, dt, 0 < x < 1.
\]

Making use of the normalization constant \(c_k\) introduced in \(48\) we can write down explicit form for the density of the product \(X_1 X_2\).

**Proposition 5** The density \(f_{12}\) of \(X_1 X_2\) is \(c_k x^{k-1} 2F_1\left(\frac{1}{2}, \frac{k}{2}; 1 - x\right), 0 < x \leq 1\).

**Proof.** By Lemma 3 the density is

\[
f_{12}(x) = \frac{c_k}{\Gamma\left(\frac{k}{2}\right) \Gamma\left(\frac{k}{2}\right)} \int_x^1 t^{k-1} (1-t)^{-1/3} \left(\frac{x}{t}\right)^{k-1} \left(\frac{1-x}{t}\right)^{-2/3} \frac{dt}{t} = \frac{c_k x^{k-1}}{\Gamma\left(\frac{k}{2}\right) \Gamma\left(\frac{k}{2}\right)} \int_x^1 (1-t)^{-1/3} (t-x)^{-2/3} t^{-1/3} dt = \frac{c_k x^{k-1}}{\Gamma\left(\frac{k}{2}\right) \Gamma\left(\frac{k}{2}\right)} \int_0^1 (1-s)^{-1/3} s^{-2/3} (x+s(1-x))^{-1/3} ds,
\]

25
using the substitution \( t = x + s(1 - x) \). Now restrict \( x \) to \( \frac{1}{2} < x \leq 1 \) then
\[
0 \leq \frac{1 - x}{x} < 1 \quad \text{and we can expand}
\]
\[
(x + s (1 - x))^{-1/3} = x^{-1/3} \left(1 - 1 \frac{x - 1}{x}\right)^{-1/3}
\]
\[
= x^{-1/3} \sum_{n=0}^{\infty} \frac{\left(\frac{1}{2}\right)_n}{n!} s^n \left(\frac{x - 1}{x}\right)^n.
\]
Integrating term-by-term we obtain
\[
f_{12}(x) = \frac{c_k x^{k-1}}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{3}{2}\right)} \sum_{n=0}^{\infty} \frac{\left(\frac{1}{2}\right)_n \Gamma\left(\frac{1}{2} + n\right) \Gamma\left(\frac{3}{2}\right)}{n! \Gamma(n + 1)} (x - 1)^n x^{-n-1/3}
\]
\[
= c_k x^{k-1} \sum_{n=0}^{\infty} \frac{\left(\frac{1}{2}\right)_n (\frac{1}{2})_n}{n!} (-1)^n \sum_{j=0}^{\infty} \frac{\left(\frac{1}{2} + n\right)_j}{j!} (1 - x)^{n+j}
\]
\[
= c_k x^{k-1} \sum_{m=0}^{\infty} \frac{\left(\frac{1}{2}\right)_m}{m!} (1 - x)^m \sum_{n=0}^{m} \frac{\left(\frac{1}{2}\right)_n m! (-1)^n}{n! (1)_n (m - n)!}
\]
where the summation variables are changed to \( n \) and \( m = n + j \). The inner sum is evaluated with the Chu-Vandermonde sum
\[
\sum_{n=0}^{m} \frac{\left(\frac{1}{2}\right)_n m! (-1)^n}{n! (1)_n (m - n)!} = \sum_{n=0}^{m} \frac{\left(\frac{1}{2}\right)_n (-m)_n}{n! (1)_n} = \frac{(1 - \frac{1}{2})_m}{(1)_m} = \frac{(\frac{3}{2})_m}{(1)_m}.
\]
This shows that \( f_{12}(x) = c_k x^{k-1} \sum_{m=0}^{\infty} \frac{\left(\frac{1}{2}\right)_m m! (-1)^n}{n! (1)_n (m - n)!} \) at least for \( \frac{1}{2} < x \leq 1 \), but both sides are analytic on \( 0 < x < 2 \) so the equality holds for \( 0 < x \leq 1 \). ■

To derive an expression for the density of the triple product \( X = X_0 X_1 X_2 \) we need to combine lemma 3 with corollary 4. Hence we can write \( f(x) = (k - \frac{1}{2}) x^{k-3/2} \int_0^1 f_{12}(t) t^{1/2 - k} dt \) and
\[
f_{12}(t) = c_k t^{k-1} \sum_{m=0}^{\infty} \frac{\left(\frac{1}{2}\right)_m m! (-1)^n}{n! (1)_n (m - n)!}
\]
\[
\int_0^1 t^{1/2 - k} dt = t^{-1/2}.
\]
This completes the proof of formula (47) for the distribution of the rescaled squared Jarlskog invariant \( X := 27Q = 108 J^2 \).

Following Gibbons et al [30] we shall now concentrate on the probability distribution for the absolute value of the Jarlskog invariant, \( |J| \), equal to \( (X/108)^{1/2} \). Let \( f_0(x) \) denote the density function of \( X^{1/2} \), thus \( f_0(x) = 2xf(x^2) \). It is not hard to compute a series for \( f_0(x) \) when \( x \) is near 1. We change the variable of integration \( t = (1 - s)^2 \) and obtain
\[
f_0(x) = 4c_k \left(k - \frac{1}{2}\right) x^{2k-2} \int_0^{1-x} 2F_1 \left(\frac{1}{3}, \frac{2}{3}; 1; s(2 - s)\right) ds.
\]
We expand \((s \cdot 2 - s)^j\) for \(j = 1, 2, \ldots\), collect the coefficients of \(s^m\) and integrate term-by-term to get

\[
\int_0^{1-x} 2F1 \left( \frac{1}{3}; \frac{2}{3}; 1; s(2-s) \right) \, ds = \sum_{m=0}^{\infty} \frac{(1-x)^{m+1}}{m+1} \sum_{j=0}^{m/2} \frac{(-1)^j \left( \frac{2}{3} \right)_m \left( \frac{2}{3} \right)_{m-j}}{(m-j)! (m-2j)! j!} ;
\]

there is no nice formula for the inner \(j\)-sum. Thus

\[
f_0(x) = 4ck \left( k - \frac{1}{2} \right) x^{2k-2} (1-x) \left\{ 1 + \frac{2}{9} (1-x) + \frac{22}{243} (1-x)^2 + \frac{310}{6561} (1-x)^3 + \cdots \right\},
\]

for \(x\) near 1 (that is, not too close to zero). When \(k = 1\) we have \(c_1 = \Gamma \left( \frac{4}{3} \right) \Gamma \left( \frac{5}{3} \right) = \frac{4\pi\sqrt{3}}{27}\) and

\[
f_0(x) = \frac{8\pi\sqrt{3}}{27} (1-x) \left( 1 + \frac{2}{9} (1-x) + \frac{22}{243} (1-x)^2 + \frac{310}{6561} (1-x)^3 + \cdots \right).
\]

**Lemma 6**  
\[ \int_0^1 2F1 \left( \frac{1}{3}; \frac{2}{3}; 1; 1-t \right) t^{-1/2} dt = 3. \]

**Proof.** Indeed

\[
\int_0^1 2F1 \left( \frac{1}{3}; \frac{2}{3}; 1; 1-t \right) t^{-1/2} dt = \sum_{n=0}^{\infty} \frac{\left( \frac{2}{3} \right)_n \left( \frac{2}{3} \right)_n}{n! n!} \int_0^1 (1-t)^n t^{-1/2} dt
\]

\[
= \sum_{n=0}^{\infty} \frac{\left( \frac{2}{3} \right)_n \left( \frac{2}{3} \right)_n}{n! n!} \frac{\Gamma (n+1) \Gamma \left( \frac{2}{3} \right)}{\Gamma (n+\frac{4}{3})} = \frac{\Gamma \left( \frac{2}{3} \right) \Gamma \left( \frac{2}{3} \right) \Gamma \left( \frac{2}{3} \right)}{\Gamma \left( \frac{2}{3} \right) \Gamma \left( \frac{2}{3} \right) \Gamma \left( \frac{2}{3} \right)} = \frac{\pi}{\frac{\pi}{6}} = 6 \sin \frac{\pi}{6} = 3.
\]

We used the Gauss sum \(\sum_{n=0}^{\infty} \frac{\left( \frac{2}{3} \right)_n \left( \frac{2}{3} \right)_n}{n! n!} (A_n - \ln t)^n\) for \(c > a + b\), and the equation \(\Gamma (1-t) \Gamma (t) = \frac{\pi}{\sin \pi t}\). 

Thus \(f(x) = 4ck \left( k - \frac{1}{2} \right) x^{2k-3/2} (3 - \int_0^x 2F1 \left( \frac{1}{3}; \frac{2}{3}; 1; 1-t \right) t^{-1/2} dt)\). To analyze the behavior for \(x\) near zero we use the classical formulas for the hypergeometric series \(2F1 (a; b; c; t)\) at the singular point \(t = 1\). The special case \(c = a + b\) is more complicated (see [44, p. 257]):

\[
2F1 \left( \frac{1}{3}; \frac{2}{3}; 1; 1-t \right) = \frac{\Gamma \left( \frac{1}{3} \right) \Gamma \left( \frac{2}{3} \right) \Gamma \left( \frac{2}{3} \right)}{\Gamma \left( \frac{2}{3} \right) \Gamma \left( \frac{2}{3} \right) \Gamma \left( \frac{2}{3} \right)} (A_n - \ln t)^n,
\]

\[
A_n := 2\psi (n+1) - \psi \left( n + \frac{1}{3} \right) - \psi \left( n + \frac{2}{3} \right).
\]

By the triplication formula for the \(\psi\)-function (recall \(\psi (t) = \frac{\text{d}}{\text{d}t} \Gamma (t) / \Gamma (t)\)),

\[
\psi \left( n + \frac{1}{3} \right) + \psi \left( n + \frac{2}{3} \right) = 3\psi (3n) - \psi (n) - 3 \ln 3
\]

\[
= 3\psi (3n + 1) - \psi (n + 1) - 3 \ln 3,
\]

27
because $\psi(t + 1) = \psi(t) = \frac{1}{t}$ for $t > 0$; the latter formula is valid for $n \geq 0$. Thus

$$A_n = 3(\psi(n + 1) - \psi(3n + 1)) + 3 \ln 3$$

$$= - \sum_{j=n+1}^{3n} \frac{3}{j} + 3 \ln 3.$$

Also

$$\int_0^x (A_n - \ln t) t^{n-1} dt = \frac{2x^{n+1/2}}{2n+1} \left( - \ln x + A_n + \frac{2}{2n+1} \right).$$

Thus

$$f(x) = c_k \left( k - \frac{1}{2} \right) x^{k-3/2} \left\{ 3 - \frac{\sqrt{3}}{2\pi} \sum_{n=0}^{\infty} \left( \frac{1}{n!} \right)_n x^{2n+1} \right\} \left( - \ln x + A_n + \frac{2}{2n+1} \right),$$

$$f_0(x) = 2c_k \left( k - \frac{1}{2} \right) x^{2k-2} \left\{ 3 - \frac{\sqrt{3}}{2\pi} \sum_{n=0}^{\infty} \left( \frac{1}{n!} \right)_n x^{2n+1} \right\} \left( - 2 \ln x + A_n + \frac{2}{2n+1} \right);$$

and we have found the density function of $x = \sqrt{X} = 6\sqrt{3}|J|$ exhibiting the behavior for $x$ near zero.

In the expression for $f_0$ the first few terms inside the braces $\{\}$ are

$$3 - \frac{\sqrt{3}}{2\pi} \left[ - \ln \left( \frac{x^2}{27} \right) \left( 2x + \frac{4}{27}x^3 + \frac{4}{81}x^5 \right) + 4x - \frac{22}{81}x^3 - \frac{49}{405}x^5 \right].$$

The cumulative distribution function $F_0(x) = P\left\{ \sqrt{X} < x \right\} = P\left\{ |J| < x/6\sqrt{3} \right\}$ is

$$F_0(x) = 2c_k \left( k - \frac{1}{2} \right) x^{2k-1}$$

$$\times \left\{ 3 - \frac{\sqrt{3}}{2\pi} \sum_{n=0}^{\infty} \left( \frac{1}{n!} \right)_n x^{2n+1} \right\} \left( - \ln \frac{x^2}{27} - \sum_{j=n+1}^{3n} \frac{3}{j} + \frac{2}{2n+1} + \frac{1}{n+k} \right).$$

The important cases are:

1. $k = 1$, Haar measure on $U(3)$,

$$F_0(x) = \frac{4\pi\sqrt{3}}{9} x^2 - \frac{2}{9} \left( - \ln \left( \frac{x^2}{27} \right) \left( x^2 + \frac{1}{27}x^4 + \frac{2}{243}x^6 + \cdots \right) + 3x^2 - \frac{4}{81}x^4 - \frac{127}{7290}x^6 + \cdots \right);$$

2. $k = \frac{3}{2}$, flat measure,

$$F_0(x) = \frac{70}{27} x^2 \left\{ \frac{3}{2} - \frac{\sqrt{3}}{2\pi} \left( - \ln \left( \frac{x^2}{27} \right) \left( \frac{2}{3}x + \frac{4}{135}x^3 + \cdots \right) + \frac{16}{9}x - \frac{86}{2025}x^3 + \cdots \right) \right\}$$

Introducing a new variable, $y = x/6\sqrt{3}$, we arrive at expressions [19] and [20] presented in Sec. 4.
C Conjectures of measures in higher dimensions

The method used in Section 4 relies on one of the authors’ [31] construction of a linear operator commuting with the action of $S_3$, mapping homogeneous polynomials in three variables to homogeneous polynomials of the same degree, and depending on a parameter $k$ (a particular case of the “Dunkl intertwining operator”). This operator is realized as an integral over $U(3)$. The case $k = 1$ is based on a formula of Harish-Chandra (see Helgason [45, p. 328])

$$\int_{U(N)} \exp \left( \text{Tr} \left( D(x)U D(y)U^* \right) \right) dm(U) = \frac{c_N}{a(x)a(y)} \sum_{w \in S_N} \det(w) \exp \left( \langle xw, y \rangle \right),$$

where the symmetric group on $N$ objects is identified with the set of permutation matrices in $O(N)$, for $x, y \in \mathbb{R}^N$ the inner product is $\langle x, y \rangle := \sum_{j=1}^{N} x_j y_j$, $D(x)$ is the diagonal matrix with $D(x)_{jj} = x_j$, $a(x) := \prod_{1 \leq i < j \leq N} (x_i - x_j)$, $dm$ is Haar measure and $c_N$ is a constant. The relation to unistochastic matrices follows from the equation

$$\text{Tr}(D(x)U D(y)U^*) = \sum_{i,j=1}^{N} x_i |U_{ij}|^2 y_j.$$

The aforementioned linear operator is known algebraically, that is, with some computational effort one can determine the action on any (low-degree!) polynomial. For $N = 3$ we were able to find a parametrized family of measures to implement the operator, roughly

$$p(x) \mapsto \int_{U(3)} p(xf(U)) d\mu_k(U),$$

where $f(U)_{ij} := |U_{ij}|^2$ as in (5) for any polynomial $p$. It is known that

$$\int_{U(N)} g \left(|U_{ij}|^2\right) dm(U) = \frac{N}{N-1} \int_0^1 g(t) (1-t)^{N-2} dt,$$

for any continuous function $g$ and any matrix entry $U_{ij}$. As in Section 5 we can compute the average (generalized) entropy for the entries $|U_{ij}|^2$ with respect to Haar measure ($q \neq 1$):

$$\langle S_q \rangle_{\text{Haar}} = \frac{1}{N(q-1)} \int_{U(N)} \sum_{i,j=1}^{N} \left(|U_{ij}|^2 - |U_{ij}|^{2q} \right) dm(U)$$

$$= \frac{N(N-1)}{q-1} \int_0^1 (t-t^q) (1-t)^{N-2} dt$$

$$= \frac{1}{q-1} \left(1 - \frac{N!}{(q+1)_{N-1}}\right)$$

$$= N! \sum_{i=0}^{N-2} \frac{(-1)^i}{i! (N-2-i)! (i+2) (q+i+1)};$$

29
the last equation is the partial fraction decomposition. Also \( \lim_{q \to 1} \langle S_q \rangle = \psi(N + 1) - \psi(2) = \sum_{j=2}^{N} \frac{1}{j} \), in agreement with the known results for the mean entropy of random complex vectors distributed according to the unitarily invariant measure \([41,42]\).

We are thus tempted to speculate that there exists a measure \( \mu_k \) on \( U(N) \) such that

\[
\int_{U(N)} g \left( |U_{ij}|^2 \right) d\mu_k (U) = \frac{\Gamma(Nk)}{\Gamma(k) \Gamma((N-1)k)} \int_0^1 g(t) t^{k-1} (1-t)^{(N-1)k-1} dt,
\]

with \( g, |U_{ij}|^2 \) as above. However we must emphasize that there is an important difference between \( U_3 \) and \( U_N \) with \( N \geq 4 \). There is a single inequality characterizing \( U_3 \) inside \( B_3 \) (the condition is \( Q(b) \geq 0 \)) while \((N - 2)^2 \) inequalities occur in general \([25]\). Another difference is that the elements of \( U \) can not necessarily be determined from the values \( \{ |U_{ij}|^2 : 1 \leq i, j \leq N \} \) (that is, up to left and right multiplication by diagonal unitary matrices and permutation of rows or columns). For instance for the flat matrix \( W_4 \) of van der Waerden, with all entries equal to \( 1/4 \) there exists a one parameter family of unitary matrices \( U(\alpha) \) (rescaled complex Hadamard matrices \([38]\)) such that

\[
||U(\alpha)||_{ij}^2 = |W_4|_{ij} = 1/4.
\]

It would be interesting to be able to fit the “flat” measure on \( U_N \) (inherited from \( B_N \)) into the \( \mu_k \) framework suggested above, but this appears to be a sizable research problem in itself.

References

[1] A. W. Marshall and I. Olkin, Inequalities: Theory of Majorization and its Applications, Academic Press, New York, 1979.
[2] G. Birkhoff, Tres observaciones sobre el algebra lineal, Univ. Nac. Tucumán Rev. A5 147 (1946).
[3] C.S. Chan and D. P. Robbins, On the volume of the polytope of doubly stochastic matrices, Exp. Math. 8, 291–300 (1999).
[4] M. Beck and D. Pixton, The Ehrhart Polynomial of the Birkhoff Polytope, Discrete & Computational Geometry 30, 623–637 (2003).
[5] E. R. Canfield and B. D. McKay, The asymptotic volume of the Birkhoff polytope, preprint arXiv:0705.2422v1 [math.CO], 2007.
[6] V. Cappellini, H.-J. Sommers, W. Bruzda, and K. Życzkowski, Random bistochastic matrices, J. Phys. A 42, 365209 (2009).
[7] M. Kobayashi and T. Maskawa, CP-violation in the renormalizable theory of weak interaction, Prog. Theor. Phys. 49 652-657 (1973).
[8] C. Jarlskog, Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Nonconservation, Phys. Rev. Lett. 55, 1039 (1985).

[9] C. Jarlskog and R. Stora, Phys. Lett. B208, 268 (1988).

[10] P. Diță, Global fits to the Cabibbo-Kobayashi-Maskawa matrix: unitarity condition method versus standard unitarity triangles approach Mod. Phys. Lett. A20 1709-1722 (2005).

[11] P. Diță, CKM matrix unitarity and a novel type of global fits preprint arXiv:0804.3282 (2008).

[12] G. W. Gibbons, S. Gielen, C. N. Pope, and N. Turok, Naturalness of CP Violation in the Standard Model, Phys. Rev. Lett. 102, 121802 (2009).

[13] J. D. Bjorken and I. Dunietz, Rephasing-invariant parametrizations of generalized Kobayashi-Maskawa matrices, Phys. Rev. D 36, 2109-2118 (1987).

[14] L. Lavoura, Parametrization of the four-generation quark mixing by the moduli of its matrix elements, Phys. Rev. D 40 2440-2448 (1989).

[15] G. Auberson, A. Martin and G. Mennessier, On the reconstruction of a unitary matrix from its moduli, Commun. Math. Phys. 140 417-437 (1991).

[16] Z. Maki, M. Nakagawa and S. Sakata, Prog. Theor. Phys. 28 870 (1962).

[17] M.-C. Chen and K.T. Mahanthappa, From CKM matrix to MNS matrix, Phys. Rev. D 62 113007 (2000).

[18] A. Landé, From Dualism to Unity in Quantum Physics, Cambridge U. P. (1960).

[19] C. Rovelli, Int. J. of Theor. Phys. 35 (1996) 1637.

[20] J.D. Louck, Doubly stochastic matrices in quantum mechanics, Found. Phys. 27, 1085 (1997).

[21] G. Mennessier and J. Nyuts, J. Math. Phys. 15 1525 (1974).

[22] G. Tanner, J. Phys. A34 8485 (2001).

[23] P. Pakoński, G. Tanner and K. Życzkowski, Families of line-graphs and their quantization, J. Stat. Phys. 111, 1331 (2003).

[24] R. F. Werner, All Teleportation and Dense Coding Schemes, J. Phys. A 34 7081-94 (2001).

[25] P. Diță, Separation of unistochastic matrices from the double stochastic ones. Recovery of a 3 x 3 unitary matrix from experimental data, J. Math. Phys. 47 083510 (2006).
[26] K. Žyczkowski, M. Kuś, W. Słomczyński and H.-J. Sommers, Random unistochastic matrices, *J. Phys.* A 36 3425 (2003).

[27] I. Bengtsson, A. Ericsson, M. Kuś, W. Tadej, and K. Žyczkowski, Birkhoff’s polytope and unistochastic matrices, N=3 and N=4, *Comm. Math. Phys.* 259, 307-324 (2005).

[28] Y.-H. Au-Yeung and Y.-T. Poon, 3 × 3 Orthostochastic Matrices and the Convexity of Generalized Numerical ranges, *Lin. Alg. Appl.* 27 69 (1979).

[29] H. Nakazato, Set of 3 × 3 Orthostochastic Matrices, *Nihonkai Math. J.* 7 83-100 (1996).

[30] G. W. Gibbons, S. Gielen, C. N. Pope, and N. Turok, Measures on Mixing Angles *Phys. Rev.* D 79, 013009 (2009).

[31] C. F. Dunkl, Intertwining operators associated to the group $S_3$, *Trans. Amer. Math. Soc.* 345 (1995), 3347-3374.

[32] C. F. Dunkl, Intertwining operators and polynomials associated with the symmetric group, *Monatsh. Math.* 126 (1998), 181-209.

[33] R. A. Brualdi and P. M. Gibson, *J. Comb. Theory* A 22, 194 (1977).

[34] CKMfitter Group (J. Charles et al.), *Eur. Phys. J.* C 41, 1-131 (2005); updated results and plots available at: http://ckmfitter.in2p3.fr

[35] P. Pakoński, K. Žyczkowski and M. Kuś, Classical 1D maps, quantum graphs and ensembles of unitary matrices, *J. Phys.* A 34, 9303 (2001).

[36] P. Pakoński, Ph.D. Thesis, Jagiellonian University, Cracow (unpublished) 2002

[37] U. Haagerup, Orthogonal maximal abelian $*$-subalgebras of the $n \times n$ matrices and cyclic $n$–rots, Operator Algebras and Quantum Field Theory (Rome), 1996 (Cambridge, MA: International Press) pp 296-322

[38] W. Tadej and K. Žyczkowski, A concise guide to complex Hadamard matrices, *Open Syst. Inf. Dyn.* 13, 133-177 (2006) and an online updated version at http://chaos.if.uj.edu.pl/~karol/hadamard

[39] W. Słomczyński, *Open Sys. Inf. Dyn.* 9, 201 (2002).

[40] H. G. Gadiyar, K. M. S. Maini, R. Padma and H. S. Sharatchandra, *J. Phys.* A 36, L109 (2003).

[41] K.R.W. Jones, *J. Phys.* A 23, L1247 (1990).

[42] I. Bengtsson and K. Žyczkowski, *Geometry of Quantum States*, Cambridge University Press, 2006.
[43] Y. Koide and H. Nishiura, Maximal CP violation hypothesis and a lepton mixing matrix, Phys. Rev. D 79, 093005 (2009)

[44] N. N. Lebedev, *Special Functions and Applications*, Eng. edition. translated by R. Silverman, Dover, New York, 1972.

[45] S. Helgason, *Groups and Geometric Analysis*, Academic Press, New York, 1984.