Boolean Inner-Product Spaces and Boolean Matrices

Stan Gudder

Department of Mathematics, University of Denver, Denver CO 80208

Frédéric Latrémolière

Department of Mathematics, University of Denver, Denver CO 80208

Abstract

This article discusses the concept of Boolean spaces endowed with a Boolean valued inner product and their matrices. A natural inner product structure for the space of Boolean n-tuples is introduced. Stochastic boolean vectors and stochastic and unitary Boolean matrices are studied. A dimension theorem for orthonormal bases of a Boolean space is proven. We characterize the invariant stochastic Boolean vectors for a Boolean stochastic matrix and show that they can be used to reduce a unitary matrix. Finally, we obtain a result on powers of stochastic and unitary matrices.

Key words: Boolean Vector Spaces, Boolean matrices, Boolean inner product.
1991 MSC: 15A03, 15A51, 06E99

1 Introduction

A Boolean space $\mathcal{L}_n(\mathcal{B})$ is the set of all n-tuples of elements of a fixed Boolean algebra $\mathcal{B}$. The elements of $\mathcal{L}_n(\mathcal{B})$ are called Boolean vectors and they possess a natural linear space-like structure. Moreover, we can define on $\mathcal{L}_n(\mathcal{B})$ an operation which is analogous to an inner product. By using this "inner product" we can also define a $\mathcal{B}$-valued norm and orthogonality relations for Boolean vectors.

Email addresses: sgudder@math.du.edu (Stan Gudder),
frederic@math.du.edu (Frédéric Latrémolière).

A Boolean matrix is a matrix whose entries are elements of a Boolean algebra $\mathcal{B}$. With the natural choice of matrix multiplication defined in terms of the lattice operations of $\mathcal{B}$, such matrices become the linear mappings between Boolean linear spaces. The study of Boolean matrices is a fascinating blend of linear algebra and boolean algebra which finds many applications, and was undertaken in [1–3,6–12,14–20].

An important concept in our work is that of a stochastic vector. These are Boolean vectors of norm one whose components are mutually disjoint. In particular, a finite partition of the universe of a Boolean algebra would correspond to a stochastic Boolean vector. We define an orthonormal basis of $\mathcal{L}_n(\mathcal{B})$ the usual way and it turns out it must be made of stochastic vectors. Our first main result is that all orthonormal bases for $\mathcal{L}_n(\mathcal{B})$ have cardinality n and conversely, any orthonormal set of stochastic vectors with cardinality n is a basis for $\mathcal{L}_n(\mathcal{B})$. Our next main result states that any orthonormal set of stochastic vectors in $\mathcal{L}_n(\mathcal{B})$ can be extended to an orthonormal basis for $\mathcal{L}_n(\mathcal{B})$. In order to prove this result, we introduce a notion of linear subspace of $\mathcal{L}_n(\mathcal{B})$.

We define stochastic and unitary Boolean matrices in terms of properties of their product with their adjoint matrices. We then show that stochastic Boolean matrices are precisely those whose columns are stochastic vectors and unitary matrices are precisely those whose rows and columns are stochastic.

We next characterize the invariant stochastic Boolean vectors for stochastic Boolean matrices and show that they can be employed to reduce unitary Boolean matrices. As mentioned in Section 2, stochastic Boolean matrices may be used to describe a dynamics analogous to a Markov chain. It is thus of interest to consider powers of stochastic Boolean matrices because they correspond to iterations in the dynamics. Our last result concerns such powers. The paper includes examples that illustrate various points which we wish to emphasize.

As a matter of notations, we shall write $\mathbb{N}$ as the set of nonzero natural numbers.

2   Definitions and Motivation

Throughout this article, $\mathcal{B}$ will denote a Boolean algebra. We denote the smallest and largest element of $\mathcal{B}$ respectively by 0 and 1. For any $a \in \mathcal{B}$, we denote by a^c its complement. For $a, b \in \mathcal{B}$, we denote the infimum of a and b by ab (instead of $a \wedge b$). We denote by $a \backslash b = a(b^c)$. The supremum of a, b is denoted by $a \vee b$.

For all $n \in \mathbb{N}$ we denote by $\mathcal{L}_n(\mathcal{B})$ the set of all n-tuples of elements in $\mathcal{B}$. We endow $\mathcal{L}_n(\mathcal{B})$ with the following operations: if $\underline{a} = (a_1, \ldots, a_n)$ and $\underline{b} = (b_1, \ldots, b_n)$ are in $\mathcal{L}_n(\mathcal{B})$, and $c \in \mathcal{B}$ then

$$\underline{a} + \underline{b} = (a_1 \vee b_1, \ldots, a_n \vee b_n)$$

and

$$c\underline{a} = (ca_1, \ldots, ca_n).$$

Then $\mathcal{L}_n(\mathcal{B})$ has the usual properties of a linear space except for the lack of additive inverses. In particular, our structure differs from the notion of Boolean vector space introduced in [14–16] which assumes an underlying additive group and is best modelled by the action of a Boolean space on a regular vector space by means of a (finitely additive) measure.

We call the elements of $\mathcal{L}_n(\mathcal{B})$ *Boolean vectors* and call $\mathcal{L}_n(\mathcal{B})$ a *Boolean (linear) space.* We will use the following definitions throughout this paper

Definition 2.1 *A Boolean vector* $\underline{a} = (a_1, \ldots, a_n)$ *is an orthovector when* $a_i a_j = 0$ *for* $i, j \in \{1, \ldots, n\}$ *and* $i \neq j$.

Definition 2.2 *An orthovector* $\underline{a} = (a_1, \ldots, a_n)$ *is a stochastic vector when* $\bigvee_{i=1}^{n} a_i = 1$.

The Boolean space $\mathcal{L}_n(\mathcal{B})$ is endowed with a natural inner product.

Definition 2.3 *Let* $\underline{a} = (a_1, \ldots, a_n)$ *and* $\underline{b} = (b_1, \ldots, b_n)$ *in* $\mathcal{L}_n(\mathcal{B})$. *Then we define the* $\mathcal{B}$*-valued inner product of these two vectors by*

$$\langle \underline{a}, \underline{b} \rangle = \bigvee_{i=1}^{n} a_i b_i.$$

The norm of $\underline{a}$ *is defined by* $\|\underline{a}\| = \langle \underline{a}, \underline{a} \rangle$.

The Boolean inner product shares most of the usual properties of the Euclidian inner product, if we replace scalar sums and products by the supremum and infimum in $\mathcal{B}$. Thus given $\underline{a}, \underline{b}, \underline{c} \in \mathcal{L}_n(\mathcal{B})$ and $\alpha \in \mathcal{B}$ then

- $\langle \alpha \underline{a} + \underline{b}, \underline{c} \rangle = \alpha \langle \underline{a}, \underline{c} \rangle \vee \langle \underline{b}, \underline{c} \rangle$,
- $\langle \underline{a}, \underline{b} \rangle = \langle \underline{b}, \underline{a} \rangle$,
- $\langle \alpha \underline{a}, \underline{c} \rangle = \langle \underline{a}, \alpha \underline{c} \rangle$,
- $\langle \underline{a}, \underline{a} \rangle = 0$ if and only if $\underline{a} = (0, \ldots, 0) = \underline{0}$.

We now give some properties of the norm.

Theorem 2.4 *Let $\underline{a}, \underline{b} \in \mathcal{L}_n(\mathcal{B})$ and $c \in \mathcal{B}$. Then*

(1) $\|c\underline{a}\| = c \|\underline{a}\|$,
(2) $\|\underline{a} + \underline{b}\| = \|\underline{a}\| \vee \|\underline{b}\|$,
(3) $\langle \underline{a}, \underline{b} \rangle \leq \|\underline{a}\| \|\underline{b}\|$,
(4) *If $\underline{a}$ and $\underline{b}$ are orthovectors and $\|\underline{a}\| = \|\underline{b}\|$ then $\langle \underline{a}, \underline{b} \rangle = \|\underline{a}\| \|\underline{b}\|$ if and only if $\underline{a} = \underline{b}$.*

PROOF. We have

$$\|c\underline{a}\| = \langle c\underline{a}, c\underline{a} \rangle = c \langle \underline{a}, c\underline{a} \rangle = c \langle \underline{a}, \underline{a} \rangle = c \|\underline{a}\|$$

and, denoting $\underline{a} = (a_1, \ldots, a_n)$ and $\underline{b} = (b_1, \ldots, b_n)$, we have

$$\|\underline{a} + \underline{b}\| = \bigvee_{i=1}^{n} (a_i \vee b_i) = \left(\bigvee_{i=1}^{n} a_i \right) \vee \left(\bigvee_{i=1}^{n} b_i \right) = \|\underline{a}\| \vee \|\underline{b}\|$$

while

$$\langle \underline{a}, \underline{b} \rangle = \bigvee_{i=1}^{n} a_i b_i \leq \bigvee_{i,j=1}^{n} a_i b_j = \left(\bigvee_{i=1}^{n} a_i \right) \left(\bigvee_{j=1}^{n} b_j \right) = \|\underline{a}\| \|\underline{b}\|.$$

Now let us assume that $\underline{a}$ and $\underline{b}$ are orthovectors and $\|\underline{a}\| = \|\underline{b}\|$ and that $\langle \underline{a}, \underline{b} \rangle = \|\underline{a}\| \|\underline{b}\|$. Hence, $\langle \underline{a}, \underline{b} \rangle = \|\underline{a}\|$ so $\bigvee_{i=1}^{n} a_i b_i = \bigvee_{i=1}^{n} a_i$. Hence, for all $j \in \{1, \ldots, n\}$

$$a_j b_j = \left(\bigvee_{i=1}^{n} a_i b_i \right) b_j = a_j b_j \vee \left(\bigvee_{i \neq j} a_i b_j \right).$$

Hence $\bigvee_{i \neq j} a_i b_j \leq a_j b_j$ yet $\bigvee_{i \neq j} a_i b_j \leq a_j^c b_j$ since $\underline{a}$ is an orthovector, so $\bigvee_{i \neq j} a_i b_j = 0$ and thus $a_i b_j = 0$. Therefore

$$a_j (\|\underline{a}\| \setminus b_j) = a_j (\|\underline{b}\| \setminus b_j) = a_j \left(\bigvee_{i \neq j} b_i \right) = \bigvee_{i \neq j} a_j b_i = 0.$$

Hence, using again that a is an orthovector, $a_j = a_j \|\underline{a}\| = a_j b_j \leq b_j$. Symmetrically, $b_j \leq a_j$ so $a_j = b_j$ for all $j \in \{1, \ldots, n\}$. Hence $\underline{a} = \underline{b}$. $\square$

Note that the condition $\|\underline{a}\| = \|\underline{b}\|$ in the last statement of Theorem (2.4) is necessary. If we let $\underline{a} = (a, 0, \ldots, 0)$ and $\underline{b} = (b, 0, \ldots, 0)$ with $a, b \in \mathcal{B}$ and $a \neq b$ then $\underline{a}, \underline{b}$ are orthovectors of different norms, and yet trivially $\langle \underline{a}, \underline{b} \rangle = \|\underline{a}\| \|\underline{b}\|$. Also, the condition that $\underline{a}$ and $\underline{b}$ are orthovectors is necessary since if $\underline{a} = (1, a)$ and $\underline{b} = (1, b)$ for $a, b \in \mathcal{B}$ with $a \neq b$ then $\|\underline{a}\| = \|\underline{b}\| = 1$ and $\langle \underline{a}, \underline{b} \rangle = 1$.

Corollary 2.5 *If $\underline{a}$ and $\underline{b}$ are stochastic Boolean vectors then $\langle \underline{a}, \underline{b} \rangle = 1$ if and only if $\underline{a} = \underline{b}$.*

PROOF. By assumption, $\underline{a}$ and $\underline{b}$ are orthovectors with $\|\underline{a}\| = \|\underline{b}\| = 1$ so the result follows from Theorem (2.4). $\square$

We now introduce the following standard notions:

Definition 2.6 *Two vectors $\underline{a}$ and $\underline{b}$ in $\mathcal{L}_n(\mathcal{B})$ are orthogonal when $\langle \underline{a}, \underline{b} \rangle = 0$, in which case we shall write $\underline{a} \perp \underline{b}$. The vector $\underline{a}$ is a unit vector when $\|\underline{a}\| = 1$.*

Definition 2.7 *An orthogonal set in $\mathcal{L}_n(\mathcal{B})$ is a subset E of $\mathcal{L}_n(\mathcal{B})$ such that for all $\underline{e}, \underline{f} \in E$ we have $\underline{e} \neq \underline{f} \implies \langle \underline{e}, \underline{f} \rangle = 0$. An orthonormal subset of $\mathcal{L}_n(\mathcal{B})$ is an orthogonal set whose elements all have norm 1.*

The next section of this paper will address the concept of dimension for a Boolean vector space. It will be based on the notion of basis. We now introduce:

Definition 2.8 *Let $\mathcal{A}$ be a subset of $\mathcal{L}_n(\mathcal{B})$. A vector $\underline{b} \in \mathcal{L}_n(\mathcal{B})$ is a linear combination of elements in $\mathcal{A}$ when there exists a finite subset $\{\underline{a_1}, \ldots, \underline{a_m}\}$ of $\mathcal{A}$ and $b_1, \ldots, b_m \in \mathcal{B}$ such that $\underline{b} = \sum_{i=1}^{m} b_i \underline{a_i}$.*

A subset $\mathcal{A}$ of $\mathcal{L}_n(\mathcal{B})$ is a generating subset of $\mathcal{L}_n(\mathcal{B})$ when all vectors in $\mathcal{L}_n(\mathcal{B})$ are linear combinations of elements in $\mathcal{A}$.

A subset $\mathcal{A}$ is free when for any $b_i, d_j \in \mathcal{B} \setminus \{0\}$ and $\underline{a_i}, \underline{c_j} \in \mathcal{A}$ with $i = 1, \ldots, m$ and $j = 1, \ldots k$ such that $\sum_{i=1}^{m} b_i \underline{a_i} = \sum_{j=1}^{k} d_j \underline{c_j}$ we have:

$$m = k, \quad \{b_1, \ldots, b_m\} = \{d_1, \ldots, d_m\} \quad and \quad \{\underline{a_1}, \ldots, \underline{a_m}\} = \{\underline{c_1}, \ldots, \underline{c_m}\}.$$

Thus a set $\mathcal{A}$ is free whenever a linear combination of elements in $\mathcal{A}$ has unique nonzero coefficients and associated vectors of $\mathcal{A}$. We naturally introduce:

Definition 2.9 *A subset $\mathcal{A}$ of $\mathcal{L}_n(\mathcal{B})$ is a basis of $\mathcal{L}_n(\mathcal{B})$ when every element of $\mathcal{L}_n(\mathcal{B})$ can be written as a unique linear combination of elements of $\mathcal{A}$ with nonzero coefficients, i.e. when $\mathcal{A}$ is generating and free.*

A first easy observation is that a basis must be made of unit vectors.

Lemma 2.10 *Let $\mathcal{A}$ be a basis of $\mathcal{L}_n(\mathcal{B})$. If $\underline{a} \in \mathcal{A}$ then $\|\underline{a}\| = 1$.*

PROOF. Note first that, if $\underline{0} = (0, \ldots, 0) \in \mathcal{L}_n(\mathcal{B})$ were in $\mathcal{A}$ and $\underline{1} =$

$(1, \ldots, 1) \in \mathcal{L}_n(\mathcal{B})$ then $\underline{1} = \underline{1} = \underline{1} + \underline{0}$, so $\underline{1}$ could be written as two distinct linear combinations of elements in $\mathcal{A}$ with coefficients 1. This is a contradiction so $\underline{0} \notin \mathcal{A}$. Let $\underline{a} \in \mathcal{A}$. Then $\underline{a} = 1\underline{a} = \|\underline{a}\| \, \underline{a}$. Hence if $\|\underline{a}\| \neq 1$ then $\underline{a}$ can be written as two distinct linear combinations of elements in $\mathcal{A}$ with nonzero coefficients (since $\underline{a} \neq \underline{0}$ so $\|\underline{a}\| \neq 0$) which contradicts the definition of a basis. $\square$

A second easy observation is:

Lemma 2.11 *Let $\mathcal{A}$ be an orthonormal set in $\mathcal{L}_n(\mathcal{B})$. Then $\mathcal{A}$ is free.*

PROOF. Let $\underline{e} = \sum_{i=1}^{m} b_i \underline{a_i} = \sum_{i=1}^{k} d_i \underline{c_i}$ with $\underline{a_1}, \ldots, \underline{a_m}, \underline{c_1}, \ldots, \underline{c_k} \in \mathcal{A}$ and $b_1, \ldots, b_m, d_1, \ldots, d_k \in \mathcal{B} \backslash \{0\}$. Note that $d_i = \langle \underline{e}, \underline{c_i} \rangle$ for $i = 1, \ldots, k$. Now if $\underline{c_j} \notin \{ \underline{a_1}, \ldots, \underline{a_m} \}$ for some $j \in \{1, \ldots, k\}$ then $d_j = \langle \underline{c_j}, \underline{e} \rangle = \langle \underline{c_j}, \sum_{i=1}^{m} b_i \underline{a_i} \rangle = 0$ which is a contradiction. Hence $\{ \underline{c_1}, \ldots, \underline{c_k} \} \subseteq \{ \underline{a_1}, \ldots, \underline{a_m} \}$. The reverse inclusion is obtained by symmetry. Then for all $i = 1, \ldots, m$ there exists $j \in \{1, \ldots, m\}$ such that $b_i = \langle \underline{e}, \underline{a_i} \rangle = \langle \underline{e}, \underline{c_j} \rangle = d_j$, concluding this proof. $\square$

We thus can set:

Definition 2.12 *A subset $\mathcal{A}$ of $\mathcal{L}_n(\mathcal{B})$ is an orthonormal basis of $\mathcal{L}_n(\mathcal{B})$ when it is an orthonormal generating subset of $\mathcal{L}_n(\mathcal{B})$.*

An orthonormal basis is thus a generating set which, by Lemma (2.11), is also free, so it is basis, so that our vocabulary is consistent.

There always exist orthonormal bases of $\mathcal{L}_n(\mathcal{B})$ and we now give some examples. First, the *canonical basis* or *standard basis* of $\mathcal{L}_n(\mathcal{B})$ is defined as the basis $(\underline{\delta_i})_{i=1,\ldots,n}$ with $\underline{\delta_1} = (1, 0, \ldots, 0)$, $\underline{\delta_2} = (0, 1, 0, \ldots, 0)$, $\ldots$, $\underline{\delta_n} = (0, \ldots, 0, 1)$. More generally, we have:

Example 2.13 *Let $\underline{a} = (a_1, \ldots, a_n)$ be a stochastic vector. Let*

$$\underline{e_i} = (a_i, a_{i+1}, \ldots, a_n, a_1, \ldots, a_{i-1})$$

for all $i \in \{1, \ldots, n\}$. Then by construction, $(\underline{e_i})_{i=1,\ldots,n}$ is an orthonormal subset of $\mathcal{L}_n(\mathcal{B})$. Moreover

$$\underline{\delta_1} = a_1\underline{e_1} + a_2\underline{e_2} + \ldots + a_n\underline{e_n}$$
$$\underline{\delta_2} = a_2\underline{e_1} + a_3\underline{e_2} + \ldots + a_n\underline{e_{n-1}} + a_1\underline{e_n}$$
$$\vdots$$
$$\underline{\delta_n} = a_n\underline{e_1} + a_1\underline{e_2} + \ldots + a_{n-1}\underline{e_n}$$

so $\left(\underline{e_i}\right)_{i=1,\ldots,n}$ is a generating set and thus an orthonormal basis for $\mathcal{L}_n\left(\mathcal{B}\right)$.

Let us observe that in general, linear independence in $\mathcal{L}_n\left(\mathcal{B}\right)$ is not an easy concept. We propose in this paper to use orthogonality as a substitute. Indeed, if $\{v_1, \ldots, v_k\}$ is a generating subset of $\mathcal{L}_n\left(\mathcal{B}\right)$ made of pairwise orthogonal, nonzero vectors, then it is a minimal generating set, in the sense that any strict subset is not generating (since, say, v_i is not a linear combination of the vectors in $\{v_1, \ldots, v_k\} \setminus \{v_i\}$ as all such combinations are orthogonal to v_i, the inner product is definite yet $v_i \neq 0$). However, orthogonality still allows for some pathologies. For instance, assume there exists $a \in \mathcal{B}$ such that a is neither 0 or 1. Then $(a, 0)$, $(a^c, 0)$ and $(0, 1)$ are three nonzero orthogonal vectors generating $\mathcal{L}_2\left(\mathcal{B}\right)$. It is a minimal generating set, yet its cardinality is not minimal among all generating families (since the canonical basis of $\mathcal{L}_2\left(\mathcal{B}\right)$ has cardinal 2). If $\mathcal{B}$ is large enough, we can even build on the same model infinite orthogonal generating families of nonzero vectors, which are therefore minimal! We shall prove in the next section that these pathologies are avoided when one restricts one's attention to orthonormal bases. We shall also see that the concept of a basis, i.e. a free generating subset, is in fact identical to the concept of an orthormal basis.

The natural maps for our structure are:

Definition 2.14 *A map $T : \mathcal{L}_n\left(\mathcal{B}\right) \longrightarrow \mathcal{L}_m\left(\mathcal{B}\right)$ is linear when for all $a \in \mathcal{B}, \underline{b}, \underline{c} \in \mathcal{L}_n\left(\mathcal{B}\right)$ we have $T(a\underline{b} + \underline{c}) = aT(\underline{b}) + T(\underline{c})$.*

As usual, $T(0) = 0$ when T is linear. When T is linear from $\mathcal{L}_n\left(\mathcal{B}\right)$ into $\mathcal{L}_n\left(\mathcal{B}\right)$, we call T an operator on $\mathcal{L}_n\left(\mathcal{B}\right)$. An operator T on $\mathcal{L}_n\left(\mathcal{B}\right)$ is invertible when there exists an operator S such that $S \circ T = T \circ S = I$ where $I : x \in \mathcal{L}_n\left(\mathcal{B}\right) \mapsto x$ is the identity operator. In the usual way, one can check that T is an invertible operator if and only if T is a linear bijection, and the inverse is a unique operator and is denoted by T^{-1}.

We shall denote by $\mathcal{B}^n$ the Boolean algebra product of $\mathcal{B}$ with itself n times. Of course, the elements of $\mathcal{B}^n$ are the same as the elements of $\mathcal{L}_n\left(\mathcal{B}\right)$, but the algebraic structures are different.

Lemma 2.15 *If T is an invertible operator on $\mathcal{L}_n\left(\mathcal{B}\right)$ then T is a Boolean algebra automorphism on $\mathcal{B}^n$.*

PROOF. Note that the supremum operation $\vee$ on $\mathcal{B}^n$ agrees with the addition on $\mathcal{L}_n\left(\mathcal{B}\right)$ by definition. So for any operator L on $\mathcal{L}_n\left(\mathcal{B}\right)$ we have $L\left(\underline{a} \vee \underline{b}\right) = L(\underline{a}) \vee L(\underline{b})$ and L preserves the order $\leq$ on $\mathcal{B}^n$. Hence, T and T^{-1} both preserve the order. Consequently, $\underline{a} \leq \underline{b}$ if and only if $T(\underline{a}) \leq T(\underline{b})$. Hence T is a lattice morphism, i.e. it also preserves the infimum. Also note that this implies that $T(1,\dots,1) = (1,\dots,1)$ – since $(1,\dots,1)$ is the largest element of $\mathcal{B}^n$, we deduce that T preserves the complement operation as well. This concludes the proof. $\square$

The converse of Lemma (2.15) does not hold, namely: if $T : \mathcal{B}^n \longrightarrow \mathcal{B}^n$ is a Boolean algebra automorphism then $T : \mathcal{L}_n\left(\mathcal{B}\right) \longrightarrow \mathcal{L}_n\left(\mathcal{B}\right)$ need not be linear. For example, let $\mathcal{B} = \{0, 1, \omega, \omega^c\}$ and consider the Boolean algebra $\mathcal{B}^2$. Define the automorphism S on $\mathcal{B}$ by $S(\omega) = \omega^c$ (so that $S(0) = 0$, $S(1) = 1$ and $S(\omega^c) = \omega$). Then $T = S \times S$ is an automorphism of $\mathcal{B}^2$. Yet, seen as a map on $\mathcal{L}_2\left(\mathcal{B}\right)$ we have

$$T\left(\omega\left(1,0\right)\right) = T(\omega, 0) = (\omega^c, 0)$$

and yet

$$\omega T\left(1, 0\right) = (\omega, 0)$$

and thus T is not linear.

We now show that if $\mathcal{B}$ is a finite Boolean algebra, then any orthonormal basis for $\mathcal{L}_n\left(\mathcal{B}\right)$ has cardinality n. Indeed, let $\left\{\underline{e_1}, \dots, \underline{e_m}\right\}$ be an orthonormal basis for $\mathcal{L}_n\left(\mathcal{B}\right)$. Define $T : \mathcal{L}_n\left(\mathcal{B}\right) \longrightarrow \mathcal{L}_m\left(\mathcal{B}\right)$ by

$$T\left(\underline{a}\right) = \left(\left\langle \underline{a}, \underline{e_1}\right\rangle, \dots, \left\langle \underline{a}, \underline{e_m}\right\rangle\right).$$

Then T is a bijection from $\mathcal{B}^n$ onto $\mathcal{B}^m$ by definition of orthonormal basis. Hence $n = m$ since $\mathcal{B}$ is finite. As previously mentioned, we shall show in the next section that this result holds for any Boolean algebra $\mathcal{B}$. Also, notice that T thus defined is an invertible operator on $\mathcal{L}_n\left(\mathcal{B}\right)$, hence a Boolean algebra automorphism of $\mathcal{B}^n$ by Lemma (2.15).

As in traditional linear algebra, the study of linear maps is facilitated by introducing matrices. A *Boolean matrix* A is a $n \times m$ matrix with entries in $\mathcal{B}$. We then write $A = [a_{ij}]$ with $a_{ij} \in \mathcal{B}$ for $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$. If A is an $n \times m$ Boolean matrix and if B is an $m \times k$ Boolean matrix,

then we define the product AB as the $n \times k$ matrix whose (i, j) entry is given by $\vee_{p=1}^{m} a_{ip} b_{pj}$. In particular, we see elements of $\mathcal{L}_n(\mathcal{B})$ as $n \times 1$ matrices (i.e. column vectors). Boolean matrices, and a generalization to distributive lattices have a considerable literature of investigation [1–3,6–12,17–20].. These matrices provide useful tools in various fields such as switching nets, automata theory and finite graph theory. Notice that permutation matrices are a special case of (invertible) Boolean matrices.

Our main motivation for studying Boolean matrices comes from an analogy of a Markov chain [5,4,13]. Let G be a finite directed graph whose vertices are labelled $1, 2, \ldots, n$ and let $\mathcal{B}$ be a fixed Boolean algebra. We think of the vertices of G as sites that a physical system can occupy. The edges of G designate the allowable transitions between sites. If there is an edge from vertex i to vertex j, we label it by an element a_{ji} of $\mathcal{B}$. We think of a_{ji} as the event, or proposition that the system evolves from site i to site j in one time-step. If there is no edge between i and j then we set $a_{ji} = 0$. The Boolean matrix $A = [a_{ij}]$ is the transition matrix in one-time-step for the physical system. The transition matrix for m-time-steps is then naturally given by A^m.

Assuming that the system evolves from a site i to some specific site j in one-time-step, we postulate that $a_{ji} a_{ki} = 0$ for $j \neq k$ and $\vee_{j=1}^{n} a_{ij} = 1$ for all $i = 1, \ldots, n$. Thus each column of A is a stochastic vector. In the next section, we will refer to such matrices as stochastic matrices. Suppose that b_i is the event that the system is in the site i initially. We would then have that the vector $\underline{b} = (b_1, \ldots, b_n)$ is a stochastic vector and $A\underline{b}$ describes the system location after one-time-step. As we shall see, $A\underline{b}$ is again a stochastic vector and in a natural way, $(A\underline{b})_i = \bigvee_{j=1}^{n} a_{ij} b_j$ is the event that the system is at site i at one time-step. Thus, $m \in \mathbb{N} \mapsto A^m$ describes the dynamics of the system and this is analogous to a traditional Markov chain. If in addition, we impose the condition that for every site i there is a specific site j from which the system evolved in one time-step, then we would have $a_{ij} a_{ik} = 0$ and $\bigvee_{i=1}^{n} a_{ij} = 1$. Such matrices are called unitary and will be studied from Section 4 onward.

In general, if G is a directed graph with n vertices and A is an $n \times n$ stochastic matrix corresponding to the edges of G, we call (G, A) a Boolean Markov chains. In section 6, we study the powers of A which are important for the description of the dynamics of (G, A).

3 The Dimension Theorem

An orthonormal set is said to be *stochastic* if all of its elements are stochastic. In this section, we show that all orthonormal bases of $\mathcal{L}_n(\mathcal{B})$ have cardinality n. Conversely, we show that any stochastic orthonormal set with cardinality n is a basis for $\mathcal{L}_n(\mathcal{B})$.

We shall use the following notations. Given a set $\mathcal{A} = \{\underline{a}_1, \ldots, \underline{a}_m\}$ of m vectors, we use the notation $\underline{a}_j = (a_{1j}, \ldots, a_{nj})$ with $a_{ij} \in \mathcal{B}$ ($i = 1, \ldots, n$ and $j = 1, \ldots, m$). Thus, we often think about a set $\{\underline{a}_1, \ldots, \underline{a}_m\}$ as a matrix $[a_{ij}]_{n \times m}$ whose columns are the elements of the set. By abuse of notation, we denote this matrix by $\mathcal{A}$ again.

We first establish that orthonormal bases possess a duality property

Theorem 3.1 *Let* $\mathcal{A} = \{\underline{a}_1, \ldots, \underline{a}_m\}$ *be an orthonormal subset of* $\mathcal{L}_n(\mathcal{B})$. *Then* $\mathcal{A}$ *is an orthonormal basis for* $\mathcal{L}_n(\mathcal{B})$ *if and only if the set* $\mathcal{A}^*$ *of columns of* $[a_{ji}]_{m \times n}$ *is an orthonormal subset of* $\mathcal{L}_m(\mathcal{B})$.

PROOF. For all $j \in \{1, \ldots, m\}$ we denote $a_j = (a_{1j}, \ldots, a_{nj})$. Assume that $\mathcal{A}$ is an orthonormal basis for $\mathcal{L}_n(\mathcal{B})$. Then there exists $b_1, \ldots, b_m \in \mathcal{B}$ such that $\underline{\delta_1} = \sum_{j=1}^m b_j \underline{a}_j$. In particular, $0 = \bigvee_{j=1}^m b_j a_{ij}$ for $i \neq 1$ so $b_j a_{ij} = 0$ for all $i \in \{2, \ldots, n\}$ and all $j \in \{1, \ldots, m\}$. Hence, $b_j a_{1j} = b_j \left(\bigvee_{i=1}^n a_{ij} \right) = b_j$ since $\vee_{i=1}^n a_{ij} = 1$. Hence $b_j \leq a_{1j}$ for all $j \in \{1, \ldots, m\}$. On the other hand, $1 = \bigvee_{j=1}^m b_j a_{1j}$ and a_{1j} and a_{1k} are disjoint for $j \neq k$, so we must have $b_j a_{1j} = a_{1j}$ for all $j \in \{1, \ldots, m\}$. Consequently, $\bigvee_{j=1}^m a_{1j} = 1$. Moreover since $b_j a_{ij} = 0$ for $i \neq 1$, we conclude that $a_{1j} a_{ij} = 0$ for $i \neq 1$.

Replacing $\underline{\delta_1}$ by $\underline{\delta_k}$ for $k \in \{1, \ldots, n\}$ we see similarly that $\bigvee_{j=1}^m a_{kj} = 1$ and $a_{kj} a_{ij} = 0$ for $i \neq k$ and for all $j \in \{1, \ldots, m\}$. Hence, the set of columns of $[a_{ji}]_{m \times n}$ is indeed an orthonormal subset of $\mathcal{L}_m(\mathcal{B})$.

Conversely, assume that $\mathcal{A}^*$ is an orthonormal subset of $\mathcal{L}_m(\mathcal{B})$. This means by definition, and using the same notations as before, that $\bigvee_{j=1}^m a_{ij} = 1$ for all $i = 1, \ldots, n$ and $a_{kj} a_{ij} = 0$ for all $i \neq k$ between 1 and n and $j = 1, \ldots, m$. It

10

follows that

$$\bigvee_{j=1}^{m} a_{kj}a_{ij} = \delta_{ij} \quad (k, i = 1, \ldots, n) \tag{3.1}$$

where δ_{ij} is $1 \in \mathcal{B}$ if $i = j$ and $0 \in \mathcal{B}$ otherwise. Now (3.1) is equivalent to

$$\underline{\delta_k} = \bigvee_{j=1}^{m} a_{kj}\underline{a_j}$$

for $k = 1, \ldots, n$ and thus $\{\underline{a_1}, \ldots, \underline{a_m}\}$ generates $\mathcal{L}_n(\mathcal{B})$ and, since it is an orthonormal set by assumption, it is an orthonormal basis of $\mathcal{L}_n(\mathcal{B})$. $\quad \square$

Corollary 3.2 *An orthonormal basis is stochastic.*

Corollary 3.3 *If $\{\underline{a_1}, \ldots, \underline{a_n}\}$ is a stochastic orthonormal subset of $\mathcal{L}_n(\mathcal{B})$ then it is a basis.*

PROOF. Let $a = \left(\bigvee_{j=1}^{n} a_{1j}\right)^c$ and assume $a \neq 0$. By Stone's Theorem, there exists a set Ω, a Boolean algebra of subsets of Ω and a Boolean algebra isomorphism $\mathcal{B} \longrightarrow \mathcal{B}_\Omega$. We identify $\mathcal{B}$ and $\mathcal{B}_\Omega$ in this proof and thus regard the elements of $\mathcal{B}$ as subsets of Ω, with 0 identified with $\emptyset$ and 1 with Ω.

Let $\omega \in a$. Then $\omega \notin a_{1j}$ for $j = 1, \ldots, n$. Since $\mathcal{A}$ is stochastic and orthonormal, we must have that $\omega \in a_{i_1 1}$, $\omega \in a_{i_2 2}$, $\ldots$, $\omega \in a_{i_{n-1} n-1}$ for some $i_1, \ldots, i_{n-1}$ with $i_r \neq 1$ and $i_r \neq i_s$ for $r, s = 1, \ldots, n-1$. Now, suppose $\omega \in a_{kn}$ for some $k \in \{1, \ldots, n\}$. Then $k \neq 1$ (since $\omega \in a$) and $k \neq i_r$ for $r = 1, \ldots, n-1$ (orthogonality). But this is a contradiction since this precludes n values for k which can only take n values. Hence $\omega \notin a_{kn}$ for all $k \in \{1, \ldots, n\}$. This contradicts, in turn, that $\underline{a_n}$ is a unit vector, i.e. form a partition of Ω. Hence, $a = 0$.

The same reasoning applies to show that $\bigvee_{j=1}^{n} a_{kj} = 1$ for all $k \in \{1, \ldots, n\}$. Hence $\mathcal{A}^*$ is an orthonormal subset of $\mathcal{L}_n(\mathcal{B})$ and thus by Theorem (3.1), $\mathcal{A}$ is an orthonormal basis for $\mathcal{L}_n(\mathcal{B})$. $\quad \square$

By symmetry, we can restate Theorem (3.1) by stating that $\mathcal{A}$ is an orthonormal basis for $\mathcal{L}_n(\mathcal{B})$ if and only if $\mathcal{A}^*$ is an orthonormal basis for $\mathcal{L}_m(\mathcal{B})$. We call $\mathcal{A}^*$ the dual basis for $\mathcal{A}$. For example, if $a_1, a_2, a_3 \in \mathcal{B}$ with $a_1 \vee a_2 \vee a_3 = 1$ and $a_i a_j = 0$ for $i \neq j$ in $\{1, 2, 3\}$, then the columns of the

11

following matrix:

$$\mathcal{A} = \begin{bmatrix} a_1 & a_3 & a_2 \\ a_2 & 0 & a_2^c \\ a_3 & a_3^c & 0 \end{bmatrix}$$

form an orthonormal basis for $\mathcal{L}_3(\mathcal{B})$. The rows form the corresponding dual basis. Notice that $\mathcal{A}$ need not be symmetric. Such a matrix $\mathcal{A}$ is what we shall call a unitary matrix in section 4.

We now establish a core result concerning the construction of stochastic vectors.

Theorem 3.4 *Let $n > 1$. Let $\underline{a} = (a_1, \ldots, a_n)$ and $\underline{b} = (b_1, \ldots, b_n)$ be two stochastic vectors in $\mathcal{L}_n(\mathcal{B})$. Then $\underline{a} \perp \underline{b}$ if and only if there exists a stochastic vector $\underline{c} = (c_1, \ldots, c_{n-1})$ in $\mathcal{L}_{n-1}(\mathcal{B})$ such that $b_i = c_i a_i^c$ for $i = 1, \ldots, n-1$. If $\underline{a} \perp \underline{b}$ then we can always choose $\underline{c}$ with $c_i = b_n a_i \vee b_i$ for $i = 1, \ldots, n-1$.*

PROOF. Suppose that $\underline{a} \perp \underline{b}$. Let $i \in \{1, \ldots, n-1\}$. We set $c_i = b_n a_i \vee b_i$. Since $\underline{a} \perp \underline{b}$, we have $b_i \leq a_i^c$. Hence

$$c_i a_i^c = (b_n a_i \vee b_i) a_i^c = b_i a_i^c = b_i.$$

Now, since $\underline{a}$ and $\underline{b}$ are stochastic vectors, we conclude that for all $j \in \{1, \ldots, n\}$ and $j \neq i$ we have

$$\begin{aligned} c_i c_j &= (b_n a_i \vee b_i)(b_n a_j \vee b_j) \\ &= b_n a_i a_j \vee b_n b_j a_i \vee b_i b_n a_j \vee b_i b_j = 0. \end{aligned}$$

Finally, we have

$$\begin{aligned} \bigvee_{i=1}^{n-1} c_i &= \bigvee_{i=1}^{n-1} (b_n a_i \vee b_i) = \left(b_n \bigvee_{i=1}^{n-1} a_i \right) \vee \bigvee_{i=1}^{n-1} b_i \\ &= b_n a_n^c \vee b_n^c = b_n \vee b_n^c = 1. \end{aligned}$$

We conclude that $\underline{c} = (c_1, \ldots, c_{n-1})$ is a stochastic vector, and it obviously has the desired property.

Conversely, suppose that there exists a stochastic vector $\underline{c}$ in $\mathcal{L}_{n-1}(\mathcal{B})$ such that $b_i = c_i a_i^c$ for $i = 1, \ldots, n-1$. Then by construction $a_i b_i = 0$ for $i = 1, \ldots, n-1$. Moreover

$$a_n b_n = a_n \left(\bigvee_{i=1}^{n-1} b_i \right)^c = a_n \left(\bigvee_{i=1}^{n-1} c_i a_i^c \right)^c$$

$$= a_n \bigwedge_{i=1}^{n-1} (a_i \vee c_i^c) = a_n \bigwedge_{i=1}^{n-1} c_i^c = a_n \left(\bigvee_{i=1}^{n-1} c_i \right)^c = 0.$$

It follows that $\underline{a} \perp \underline{b}$. $\square$

We can now show:

Lemma 3.5 *If* $\mathcal{A} = \left\{ \underline{a_1}, \ldots, \underline{a_m} \right\}$ *is a stochastic orthonormal set in* $\mathcal{L}_n(\mathcal{B})$ *then* $m \leq n$.

PROOF. We proceed by induction on $n \in \mathbb{N}$. For $n = 1$ the only orthonormal set is $\{1\}$ so the result holds trivially. Now we assume the results holds for some $n \in \mathbb{N}$. Let $\mathcal{A} = [a_{ij}]_{(n+1) \times m}$ be a stochastic orthonormal set in $\mathcal{L}_{n+1}(\mathcal{B})$. By Theorem (3.4), for each $j = 2, \ldots, m$ there exists a stochastic vector $c_j = (c_{1j}, \ldots, c_{nj})$ in $\mathcal{L}_n(\mathcal{B})$ such that $a_{ij} = c_{ij} a_{i1}^c$ for all $i = 1, \ldots, n$ and $\overline{j} = 2, \ldots, m$. Let $j, k \in \{2, \ldots, m\}$ with $j \neq k$ and $i \in \{1, \ldots, n\}$. Recall from Theorem (3.4) that $c_{ij} = a_{i1} a_{nj} \vee a_{ij}$, and since $\mathcal{A}$ is orthonormal

$$
\begin{aligned}
c_{ij} c_{ik} &= (a_{i1} a_{nj} \vee a_{ij})(a_{i1} a_{nk} \vee a_{ik}) \\
&= a_{i1} a_{nj} a_{nk} \vee a_{i1} a_{nj} a_{ik} \vee a_{i1} a_{ij} a_{nk} \vee a_{ij} a_{ik} \\
&= 0.
\end{aligned}
$$

Hence $\left\{ \underline{c_2}, \ldots, \underline{c_m} \right\}$ is a stochastic orthonormal set in $\mathcal{L}_n(\mathcal{B})$. By our induction hypothesis, $m - 1 \leq n$ and thus $m \leq n + 1$, which completes our proof by induction. $\square$

The main result of this section is:

Theorem 3.6 *If* $\mathcal{A}$ *is an orthonormal basis for* $\mathcal{L}_n(\mathcal{B})$ *then the cardinality of* $\mathcal{A}$ *is* n.

PROOF. We proceed by induction on n. The result is trivial for $n = 1$. Assume that for some $n \in \mathbb{N}$, if $\mathcal{A}_0$ is an orthonormal basis for $\mathcal{L}_k(\mathcal{B})$ with $k \leq n$ then $\mathcal{A}_0$ contains exactly k vectors. Let $\mathcal{A}$ be an orthonormal basis of $\mathcal{L}_{n+1}(\mathcal{B})$. By Corollary (3.2), $\mathcal{A}$ is stochastic. Applying Lemma (3.5), we deduce that the cardinality m of $\mathcal{A}$ satisfies $m \leq n+1$. Assume that $m < n+1$. By Theorem (3.1), $\mathcal{A}^*$ is an orthonormal basis for $\mathcal{L}_m(\mathcal{B})$. Since $m \leq n$, we

conclude by our induction hypothesis that the cardinality of $\mathcal{A}^*$ is m. But by construction, the cardinality of $\mathcal{A}^*$ is $n + 1$, which is a contradiction. Hence $m = n + 1$ which completes our proof by induction. $\quad\square$

Combining Theorem (3.6) and Corollary (3.3) we obtain the following result:

Corollary 3.7 *A stochastic orthonormal set $\mathcal{A}$ is a basis for $\mathcal{L}_n(\mathcal{B})$ if and only if the cardinality of $\mathcal{A}$ is n.*

To be fully satisfactory, we shall now check that the orthonormal families of $\mathcal{L}_n(\mathcal{B})$ of cardinality n are in fact basis. We shall use the following:

Lemma 3.8 *If $\underline{a} = (a_1, \ldots, a_n) \in \mathcal{L}_n(\mathcal{B})$ is a unit vector, then there exists a stochastic vector $\underline{b} = (b_1, \ldots, b_n)$ with $b_i \le a_i$ for all $i = 1, \ldots, n$.*

PROOF. For $i = 1, \ldots, n$ we set $b_i = a_i \left(a_1^c a_2^c \ldots a_{i-1}^c \right) \le a_i$. Then $b_i b_j = 0$ for $i, j = 1, \ldots, n$ and $i \ne j$, and $\bigvee\limits_{i=1}^{n} b_i = \bigvee\limits_{i=1}^{n} a_i = 1$ so $\underline{b}$ is a stochastic vector. $\quad\square$

Now, we can state:

Corollary 3.9 *An orthonormal set of $\mathcal{L}_n(\mathcal{B})$ is a basis for $\mathcal{L}_n(\mathcal{B})$ if and only if it has cardinality n.*

PROOF. Let $\mathcal{A} = \left\{ \underline{a_1}, \ldots, \underline{a_n} \right\}$ be an orthonormal set. Using Lemma (3.8), there exists a set of stochastic vectors $\underline{b_1}, \ldots, \underline{b_n}$ such that $b_{ij} \le a_{ij}$. Therefore, $\left\{ \underline{b_1}, \ldots, \underline{b_n} \right\}$ is a stochastic orthogonal set of size n and thus it is a basis for $\mathcal{L}_n(\mathcal{B})$ by Corollary (3.7). Now, let $i, j, k, l = 1, \ldots, n$ with $i > j$. Let $\underline{v} = a_{ik} a_{jk} \underline{\delta_i}$. Then, using the construction of Lemma (3.8), we have

$$a_{ik} a_{jk} b_{il} = a_{ik} a_{jk} a_{il} a_{1l}^c \ldots a_{i-1,l}^c = 0$$

since either $l = k$ and then $a_{ik} a_{jk} b_{il} \le a_{jk} a_{jk}^c = 0$ since $i > j$, or $l \ne k$ and $a_{ik} a_{il} = 0$ since $\mathcal{A}$ is orthogonal. Hence the vector $\underline{v}$ is orthogonal to $\underline{b_1}, \ldots, \underline{b_n}$, thus $\underline{v} = 0$. Hence, $\mathcal{A}$ is stochastic. By Corollary (3.7), it is an orthonormal basis of $\mathcal{L}_n(\mathcal{B})$.

The converse is Theorem (3.6). $\quad\square$

In view of Corollary (3.9), we call n the *dimension* of the Boolean linear space $\mathcal{L}_n(\mathcal{B})$. We now consider the following question: can any stochastic or-

thonormal subset $\mathcal{A}$ of $\mathcal{L}_n(\mathcal{B})$ be extended to an orthonormal basis for $\mathcal{L}_n(\mathcal{B})$? By Lemma (3.5), $\mathcal{A}$ can not have more than n vectors. Of course, if the cardinality of $\mathcal{A}$ is n then it is already a basis by Corollary (3.3). Moreover, Example (2.13) shows that if $\mathcal{A}$ is reduced to a unique stochastic vector, then there is an orthonormal basis of $\mathcal{L}_n(\mathcal{B})$ containing $\mathcal{A}$ so the answer is affirmative. We shall now prove that the answer is affirmative in general.

We shall use the following concept:

Definition 3.10 *A subset $\mathcal{M} \subseteq \mathcal{L}_n(\mathcal{B})$ is a subspace if it is generated by an orthonormal set $\mathcal{A} = \left\{ \underline{a_1}, \ldots, \underline{a_m} \right\}$, i.e.*

$$\mathcal{M} = \left\{ \sum_{i=1}^{m} b_i \underline{a_i} : b_1, \ldots, b_m \in \mathcal{B} \right\}.$$

Any orthonormal set $\mathcal{A}$ generating $\mathcal{M}$ is called an orthonormal basis for $\mathcal{M}$.

We emphasize that we do not require orthonormal bases of subspaces to be stochastic. In fact, a subspace may not contain any stochastic orthonormal basis: for example, if there exists $a \in \mathcal{B}$ such that $a \notin \{0,1\}$ then the subset $E = \{b(1,a) : b \in \mathcal{B}\}$ is a subspace with basis $(1,a)$. Since any orthonormal set of two vectors generates $\mathcal{L}_2(\mathcal{B}) \neq E$, any orthonormal basis for E is necessarily reduced to one vector. If this vector is stochastic, then it is of the form (b, b^c) for some $b \in \mathcal{B}$. It is then easy to check that $(1,a)$ can not be of the form (cb, cb^c) and thus E has no stochastic vector basis. Thus, we will sometimes use:

Definition 3.11 *A subspace with a stochastic orthonormal basis is called a stochastic subspace.*

Linear maps generalize trivially to linear maps between two subspaces. Of special interest to us will be:

Definition 3.12 *A linear map $T : \mathcal{M} \longrightarrow \mathcal{N}$ between two subspaces $\mathcal{M}$ and $\mathcal{N}$ of, respectively, $\mathcal{L}_n(\mathcal{B})$ and $\mathcal{L}_m(\mathcal{B})$, is called an isometry when for all $\underline{a}, \underline{b} \in \mathcal{M}$ we have $\langle T(\underline{a}), T(\underline{b}) \rangle = \langle \underline{a}, \underline{b} \rangle$.*

Lemma 3.13 *Let $\mathcal{M} \subseteq \mathcal{L}_n(\mathcal{B})$ and $\mathcal{N} \subseteq \mathcal{L}_m(\mathcal{B})$ be two subspaces. Let $T : \mathcal{M} \longrightarrow \mathcal{N}$ be a linear map. The following are equivalent:*

(1) T is an isometry,
(2) There exists an orthonormal basis $\mathcal{A} = \left\{ \underline{e_1}, \ldots, \underline{e_k} \right\}$ of $\mathcal{M}$ such that $\left\{ T\underline{e_i} : i = 1, \ldots, k \right\}$ is an orthonormal set of $\mathcal{N}$,

(3) *For every orthonormal set $\mathcal{A} = \left\{\underline{e_1}, \ldots, \underline{e_k}\right\}$ of $\mathcal{M}$, the set $\left\{T\underline{e_1}, \ldots T\underline{e_k}\right\}$ is an orthonormal set of $\mathcal{N}$.*

Moreover, if T is an isometry, then it is injective.

PROOF. We start by proving that (2) implies (1). Let $\mathcal{A} = \left\{\underline{e_1}, \ldots, \underline{e_k}\right\}$ be an orthonormal subset of $\mathcal{M}$ such that $\left\{T\underline{e_1}, \ldots, T\underline{e_k}\right\}$ is orthonormal. Let $\underline{a}, \underline{b} \in \mathcal{M}$. We can write $\underline{a} = \sum_{i=1}^{k} a_i \underline{e_i}$ and $\underline{b} = \sum_{i=1}^{k} b_i \underline{e_i}$ with $a_i, b_i \in \mathcal{B}$ $(i = 1, \ldots, n)$. Then

$$\langle T\underline{a}, T\underline{b} \rangle = \bigvee_{i,j=1}^{n} \left\langle a_i T\underline{e_i}, b_j T\underline{e_j} \right\rangle = \bigvee_{i,j=1}^{n} a_i b_j \left\langle T\underline{e_i}, T\underline{e_j} \right\rangle$$

$$= \bigvee_{i=1}^{n} a_i b_i = \langle \underline{a}, \underline{b} \rangle .$$

Hence T is an isometry.

Now, (1) implies (3) and (3) implies (2) are both trivial.

Assume now that T is an isometry. Assume $T\underline{a} = T\underline{b}$. Then, using the same notations as above, we have

$$a_i = \left\langle a, \underline{e_i} \right\rangle = \left\langle T\underline{a}, T\underline{e_i} \right\rangle = \left\langle T\underline{b}, T\underline{e_i} \right\rangle = \left\langle \underline{b}, \underline{e_i} \right\rangle = b_i$$

for all $i = 1, \ldots, k$. Hence $\underline{a} = \underline{b}$. $\quad\square$

Definition 3.14 *Let $\mathcal{M}$ and $\mathcal{N}$ be two subspaces of respectively $\mathcal{L}_m(\mathcal{B})$ and $\mathcal{L}_n(\mathcal{B})$. A surjective isometry $T : \mathcal{M} \longrightarrow \mathcal{N}$ is called an isomorphism, and then $\mathcal{M}$ and $\mathcal{N}$ are called isomorphic subspaces.*

It is clear that the inverse of an isomorphism is an isomorphism, and that the composition of two isomorphisms is again an isomorphism. It follows that isomorphic is an equivalence relation. It is also an important observation that isomorphisms map orthonormal bases to orthonormal bases: if $\left\{\underline{a_1}, \ldots, \underline{a_n}\right\}$ is an orthonormal basis for a subspace $\mathcal{M}$ and $T : \mathcal{M} \longrightarrow \mathcal{N}$ is an isomorphism then $\left\{T\underline{a_1}, \ldots, T\underline{a_n}\right\}$ is an orthonormal set since T is an isometry (Theorem (3.13)). Moreover, if $\underline{b} \in \mathcal{N}$ then there exists $\underline{c} \in \mathcal{M}$ such that $T(\underline{c}) = \underline{b}$. Since $\underline{c} = \sum_{i=1}^{n} c_i \underline{a_i}$ for some $c_1, \ldots, c_n \in \mathcal{B}$ we conclude that $\underline{b} = \sum_{i=1}^{n} c_i T(\underline{a_i})$. Hence $\left\{T\underline{a_1}, \ldots, T\underline{a_n}\right\}$ is an orthonormal generating subset of $\mathcal{N}$, hence a basis of $\mathcal{N}$.

Theorem 3.15 *If $\mathcal{M}$ is a subspace then there exists an $m \in \mathbb{N}$ and an isomorphism $T : \mathcal{M} \longrightarrow \mathcal{L}_m(\mathcal{B})$. Moreover T can be chosen to take stochastic*

vectors to stochastic vectors, and if $\mathcal{M}$ is a stochastic subspace then T can be chosen so that T and T^{-1} map stochastic vectors to stochastic vectors.

PROOF. Let $\left\{\underline{e_1}, \ldots, \underline{e_m}\right\}$ be an orthonormal basis for $\mathcal{M}$ and let us denote the canonical basis of $\mathcal{L}_m\left(\mathcal{B}\right)$ by $\left\{\underline{\delta_1}, \ldots, \underline{\delta_m}\right\}$. We define $T : \mathcal{M} \longrightarrow \mathcal{L}_m\left(\mathcal{B}\right)$ by setting for all $\underline{a} \in \mathcal{M}$:

$$T\underline{a} = \left(\langle \underline{a}, e_1 \rangle, \ldots, \langle \underline{a}, e_m \rangle\right).$$

Then T is linear and $T\underline{e_i} = \underline{\delta_i}$ for $i \in \{1, \ldots, m\}$. By Lemma (3.13), T is an isometry and T is surjective by construction (if $\underline{b} \in \mathcal{L}_m\left(\mathcal{B}\right)$ then $\underline{b} = (b_1, \ldots, b_m)$ then $T\left(\sum_{i=1}^{m} b_i \underline{e_i}\right) = \underline{b}$). So T is an isomorphism.

Moreover, T preserves stochastic vectors. Indeed, let $\underline{a}$ be a stochastic vector in $\mathcal{M}$. Let $n \in \mathbb{N}$ such that $\mathcal{M}$ is a subspace of $\mathcal{L}_n\left(\mathcal{B}\right)$. Denote by $\left\{\underline{\delta'_1}, \ldots, \underline{\delta'_n}\right\}$ the canonical orthonormal basis of $\mathcal{L}_n\left(\mathcal{B}\right)$. For $i = 1, \ldots, m$ then

$$\langle \underline{a}, \underline{e_i} \rangle = \left\langle \underline{a}, \sum_{r=1}^{n} \langle \underline{e_i}, \underline{\delta'_r} \rangle \underline{\delta'_r} \right\rangle = \bigvee_{r=1}^{n} \langle \underline{e_i}, \underline{\delta'_r} \rangle \langle \underline{a}, \underline{\delta'_r} \rangle.$$

Hence, for $i \neq j$ and $i, j = 1, \ldots, m$ we have

$$\langle \underline{a}, \underline{e_i} \rangle \langle \underline{a}, \underline{e_j} \rangle = \bigvee_{r,s=1}^{n} \langle \underline{e_i}, \underline{\delta'_r} \rangle \langle \underline{a}, \underline{\delta'_r} \rangle \langle \underline{e_j}, \underline{\delta'_s} \rangle \langle \underline{a}, \underline{\delta'_s} \rangle$$

$$= \bigvee_{r=1}^{n} \langle \underline{e_i}, \underline{\delta'_r} \rangle \langle \underline{a}, \underline{\delta'_r} \rangle \langle \underline{e_j}, \underline{\delta'_r} \rangle \quad \text{since } \underline{a} \text{ is stochastic}$$

$$\leq \bigvee_{r=1}^{n} \langle \underline{e_i}, \underline{\delta'_r} \rangle \langle \underline{e_j}, \underline{\delta'_r} \rangle = \langle \underline{e_i}, \underline{e_j} \rangle = 0.$$

Hence, by definition, $T\underline{a}$ is stochastic.

Now, it is easy to check that $T^{-1}\left(a_1, \ldots, a_m\right) = \sum_{k=1}^{m} a_i \underline{e_i}$. Assume that $\mathcal{M}$ is stochastic and that the basis $\left\{\underline{e_1}, \ldots, \underline{e_n}\right\}$ is stochastic. If $\left(a_1, \ldots, a_m\right) \in \mathcal{L}_m\left(\mathcal{B}\right)$ is stochastic, then for $r, s = 1, \ldots, m$:

$$\left\langle \sum_{k=1}^{m} a_k \underline{e_k}, \underline{\delta'_r} \right\rangle \left\langle \sum_{k=1}^{m} a_k \underline{e_k}, \underline{\delta'_s} \right\rangle = \bigvee_{k,l=1}^{m} a_k a_l \langle \underline{e_k}, \underline{\delta'_r} \rangle \langle \underline{e_l}, \underline{\delta'_s} \rangle$$

$$= \bigvee_{k=1}^{m} a_k \langle \underline{e_k}, \underline{\delta'_r} \rangle \langle \underline{e_k}, \underline{\delta'_s} \rangle \quad \text{as } \underline{a} \text{ is stochastic}$$

$$= a_k \delta_r^s$$

with δ^s_r is the Kronecker symbol. Note that we used that by definition, an orthonormal basis of a subspace is stochastic. Hence $T^{-1}(a_1, \ldots, a_m)$ is a stochastic vector as well. Hence T^{-1} maps stochastic vectors to stochastic vectors. $\square$

Corollary 3.16 *Any two orthonormal bases of a subspace $\mathcal{M}$ have the same cardinality.*

PROOF. Let $\mathcal{A} = \left\{\underline{a_1}, \ldots, \underline{a_m}\right\}$ and $\mathcal{B} = \left\{\underline{b_1}, \ldots, \underline{b_n}\right\}$ be two orthonormal bases of $\mathcal{M}$. By Theorem (3.15), there exists isomorphisms $T : \mathcal{M} \longrightarrow \mathcal{L}_m(\mathcal{B})$ and $S : \mathcal{M} \longrightarrow \mathcal{L}_n(\mathcal{B})$. Hence $T \circ S^{-1} : \mathcal{L}_n(\mathcal{B}) \longrightarrow \mathcal{L}_m(\mathcal{B})$ is an isomorphism. In particular, it maps orthonormal basis to orthonormal basis. Hence $n = m$ by Theorem (3.6). $\square$

We call the common cardinality of all orthonormal bases for a subspace $\mathcal{M}$ the *dimension of* $\mathcal{M}$. It follows from Theorem (3.15) that if $\mathcal{M}$ has dimension m, then $\mathcal{M}$ is isomorphic to $\mathcal{L}_m(\mathcal{B})$. A source of examples of subspaces is given by:

Proposition 3.17 *For any $\underline{a_1} \in \mathcal{L}_n(\mathcal{B})$ we denote by $\underline{a_1}^\perp$ the set*

$$\left\{\underline{b} \in \mathcal{L}_n(\mathcal{B}) : \left\langle \underline{a_1}, \underline{b} \right\rangle = 0\right\}.$$

If $\underline{a_1}$ is stochastic then $\underline{a_1}^\perp$ is a stochastic subspace of $\mathcal{L}_n(\mathcal{B})$ of dimension $n-1$.

PROOF. Using Example (2.13), we extend the stochastic vector $\underline{a_1}$ to an orthonormal basis $\left\{\underline{a_1}, \ldots, \underline{a_n}\right\}$ of $\mathcal{L}_n(\mathcal{B})$. If $\underline{b} \perp \underline{a_1}$ then, writing $\underline{b} = \sum_{i=1}^n b_i \underline{a_i}$ we see that $\left\langle \underline{b}, \underline{a_1} \right\rangle = 0$ if and only if $b_1 = 0$. Hence

$$\underline{a_1}^\perp = \left\{\sum_{i=2}^n b_i \underline{a_i} : b_2, \ldots, b_n \in \mathcal{B}\right\}$$

is the subspace generated by the stochastic orthonormal set $\left\{\underline{a_2} \ldots, \underline{a_n}\right\}$ of cardinality $n-1$. $\square$

We are now ready to show:

Theorem 3.18 *If $\mathcal{A} = \left\{\underline{a_1}, \ldots, \underline{a_m}\right\}$ is a stochastic orthonormal set in $\mathcal{L}_n(\mathcal{B})$ with $m < n$ then $\mathcal{A}$ can be extended to an orthonormal basis for $\mathcal{L}_n(\mathcal{B})$.*

PROOF. We proceed by induction on n. The result is trivial for $n = 1$. Assume that for some $n \in \mathbb{N}$, any stochastic orthonormal set of cardinality $m < n$ in $\mathcal{L}_n(\mathcal{B})$ can be extended to a basis for $\mathcal{L}_n(\mathcal{B})$. Let $\mathcal{A} = \left\{\underline{a_1}, \ldots, \underline{a_m}\right\}$ be a stochastic orthonormal subset of $\mathcal{L}_{n+1}(\mathcal{B})$ with $m < n+1$. By Proposition (3.17) and Theorem (3.15) there exists an isomorphism $T : \underline{a_1}^{\perp} \longrightarrow \mathcal{L}_n(\mathcal{B})$ such that T and T^{-1} preserve stochastic vectors. Moreover, $\left\{\underline{a_2}, \ldots, \underline{a_m}\right\} \subseteq \underline{a_1}^{\perp}$. Let $\underline{b_i} \in \mathcal{L}_n(\mathcal{B})$ be given by $T\underline{a_i} = \underline{b_i}$ for $i = 2, \ldots, m$. It follows from Theorem (3.13) that $\left\{\underline{b_2}, \ldots, \underline{b_m}\right\}$ is an orthonormal set in $\mathcal{L}_n(\mathcal{B})$ of cardinal $m - 1 < n$. By our induction hypothesis, there exist stochastic vectors $\underline{b_{m+1}}, \ldots, \underline{b_{n+1}}$ such that $\left\{\underline{b_2}, \ldots, \underline{b_{n+1}}\right\}$ is an orthonormal basis for $\mathcal{L}_n(\mathcal{B})$. By Theorem (3.13), $\left\{T^{-1}\underline{b_2}, \ldots, T^{-1}\underline{b_{n+1}}\right\}$ is a stochastic orthonormal set in $\mathcal{L}_{n+1}(\mathcal{B})$ which is a basis for $\underline{a_1}^{\perp}$. Since $\underline{a_i} = T^{-1}\underline{b_i}$ for $i = 2, \ldots, m$, we conclude by Corollary (3.3) that $\left\{a, T^{-1}\underline{b_2}, \ldots, T^{-1}\underline{b_{n+1}}\right\}$ is an orthonormal basis of $\mathcal{L}_{n+1}(\mathcal{B})$ which extends $\mathcal{A}$. $\square$

It follows from Theorem (3.18) that if $\mathcal{M}$ is a stochastic subspace of $\mathcal{L}_n(\mathcal{B})$ then

$$\mathcal{M}^{\perp} = \{\underline{b} \in \mathcal{L}_n(\mathcal{B}) : \forall \underline{a} \in \mathcal{M} \quad \underline{b} \perp \underline{a}\}$$

is also a stochastic subspace and $\mathcal{L}_n(\mathcal{B}) = \mathcal{M} + \mathcal{M}^{\perp}$. One can now study projection operators and the order structure of subspaces but we leave this for later work.

4   Stochastic and Unitary Matrices

In the sequel, a matrix on $\mathcal{L}_n(\mathcal{B})$ will mean an $n \times n$ Boolean matrix, and a vector in $\mathcal{L}_n(\mathcal{B})$ will mean a Boolean vector and will be identified with a $n \times 1$ column vector. Moreover, if A is a matrix then we denote the $(i, j)^{\text{th}}$ entry by $(A)_{ij}$, or simply $(A)_i$ if A is a column vector.

Let A be a matrix on $\mathcal{L}_n(\mathcal{B})$. Then the map $\underline{x} \in \mathcal{L}_n(\mathcal{B}) \mapsto A\underline{x}$ is linear and will be identified with A. Indeed, for all $\underline{b}, \underline{c} \in \mathcal{L}_n(\mathcal{B})$, $c \in \mathcal{B}$, and $i = 1, \ldots, n$ we have

$$\left(A\left(c\underline{b}\right)\right)_i = \bigvee_{j=1}^{n} a_{ij}\left(c\underline{b}\right)_j = \bigvee_{j=1}^{n} a_{ij}cb_j = c\bigvee_{j=1}^{n} a_{ij}b_j = c\left(A\underline{b}\right)_i$$

and

$$\left(A\left(\underline{b}+\underline{c}\right)\right)_i = \bigvee_{j=1}^{n} a_{ij}\left(\underline{b}+\underline{c}\right)_j = \bigvee_{j=1}^{n} a_{ij}\left(b_j \vee c_j\right)$$

$$= \left(\bigvee_{j=1}^{n} a_{ij}b_j\right) \vee \left(\bigvee_{j=1}^{n} a_{ij}c_j\right)$$

$$= (A\underline{b})_i \vee (A\underline{c})_i = (A\underline{b} + A\underline{c})_i\,.$$

Conversely, any operator T on $\mathcal{L}_n\left(\mathcal{B}\right)$ can be represented by a matrix on $\mathcal{L}_n\left(\mathcal{B}\right)$ with respect to the canonical basis. Indeed, define $a_{ij} = \left\langle T\underline{\delta_j}, \underline{\delta_i}\right\rangle$ for all $i, j = 1, \ldots, n$. Then $T\underline{\delta_j} = \sum_{i=1}^{n} a_{ij}\underline{\delta_i}$. Defining the matrix $A_T = [a_{ij}]_{n\times n}$ we have

$$\left(A_T\underline{\delta_i}\right)_k = \bigvee_{j=1}^{n} a_{kj}\left(\underline{\delta_i}\right)_j = \bigvee_{j=1}^{n} a_{kj}\delta_{ji} = a_{ki} = \left(T\underline{\delta_i}\right)_k$$

for all $i, k = 1, \ldots, n$ and it follows that the action of A_T is given by T. The matrix A_T is called the *matrix corresponding to T* in the canonical basis of $\mathcal{L}_n\left(\mathcal{B}\right)$. If $A = [a_{ij}]_{n\times n}$ is a matrix on $\mathcal{L}_n\left(\mathcal{B}\right)$ then its transpose $[a_{ji}]_{n\times n}$ is denoted by A^*.

It is straightforward to check that if $T : \mathcal{L}_n\left(\mathcal{B}\right) \longrightarrow \mathcal{L}_m\left(\mathcal{B}\right)$ and $S : \mathcal{L}_m\left(\mathcal{B}\right) \longrightarrow \mathcal{L}_k\left(\mathcal{B}\right)$ then the matrix of $S \circ T$ is given by the product $A_S A_T$, the matrix of λT for $\lambda \in \mathcal{B}$ is given by λA_T and if $S : \mathcal{L}_n\left(\mathcal{B}\right) \longrightarrow \mathcal{L}_m\left(\mathcal{B}\right)$ then the matrix of $S + T$ is $A_S + A_T$. Moreover, for all $\underline{a} \in \mathcal{L}_n\left(\mathcal{B}\right)$ and $\underline{b} \in \mathcal{L}_m\left(\mathcal{B}\right)$ we check that $\langle T\underline{a}, \underline{b}\rangle = \langle \underline{a}, T^*\underline{b}\rangle$ where $T^* : \mathcal{L}_m\left(\mathcal{B}\right) \longrightarrow \mathcal{L}_n\left(\mathcal{B}\right)$ is the linear map of matrix A_T^* (and where we use the same notation for the inner products on $\mathcal{L}_n\left(\mathcal{B}\right)$ and $\mathcal{L}_m\left(\mathcal{B}\right)$). Thus, linear maps always have an adjoint. It is routine to check that the adjoint is unique. We thus have, as with standard linear algebra, a natural isomorphism between the *-algebra of linear maps and the *-algebra of Boolean matrices.

Invertibility of Boolean matrices was studied in [7,8,12,20] and the following result is well-known. We present here a short proof which relies upon our previous work with orthonormal bases and generalize the invertibility result to show that invertible rectangular matrices have to be square. Note that if a matrix A is invertible, then its columns and its rows both form generating families. We now show that these families are actually orthonormal bases of $\mathcal{L}_n\left(\mathcal{B}\right)$ and therefore are stochastic.

Theorem 4.1 *Let A be an $n \times m$ Boolean matrix. The following are equivalent:*

(1) *A is invertible, i.e. there exists a (necessarily unique) $m \times n$ Boolean matrix A^{-1} such that $A^{-1}A = I_n$ and $AA^{-1} = I_m$,*
(2) *A is unitary, i.e. $n = m$ and $AA^* = A^*A = I_n$,*
(3) *The columns of A form an orthonormal basis for $\mathcal{L}_n\left(\mathcal{B}\right)$,*

(4) The rows of A form an orthonormal basis of $\mathcal{L}_m(\mathcal{B})$.

In particular, if any of 1-4 holds, then $n = m$.

PROOF. Assume (3) holds. Then by Theorem (3.6), there are n columns of A and thus $n = m$. By Theorem (3.1), the rows of A are a basis for $\mathcal{L}_n(\mathcal{B})$ as well, so (4) holds. The same reasoning shows that (4) implies (3) and in particular $n = m$ again.

Moreover, let us denote the columns of A by $\underline{a_1}, \ldots, \underline{a_m}$ and the rows of A by $\underline{r_1}, \ldots, \underline{r_n}$. By construction $A^*A = \left[\langle \underline{a_i}, \underline{a_j} \rangle\right]_{m \times m}$ and $AA^* = \left[\langle \underline{r_i}, \underline{r_j} \rangle\right]_{n \times n}$ so A is unitary if and only if both (3) and (4) holds. Since (3) and (4) are equivalent and imply $n = m$, either imply (2).

Assume now that A is invertible and write $A = [a_{ij}]_{n \times m}$ and $A^{-1} = [b_{ij}]_{m \times n}$. Then $A^{-1}A = I_m$ and $AA^{-1} = I_n$ implies that $\bigvee_{j=1}^{m} a_{ij} = \bigvee_{j=1}^{n} b_{ij} = 1$ and $b_{ki}a_{ij} = 0$ ($k \neq j \in \{1, \ldots, m\}$ and $i \in \{1, \ldots n\}$) and $a_{ik}b_{kj} = 0$ ($i \neq j \in \{1, \ldots n\}$ and $k \in \{1, \ldots m\}$). Moreover if $i \in \{1, \ldots, n\}$ and $j \neq k \in \{1, \ldots, m\}$ then:

$$a_{ij}a_{ik} = \left(\bigvee_{s=1}^{n} b_{ks}a_{ij}\right)a_{ik} = \left(\bigvee_{\substack{s=1 \\ s \neq i}}^{n} a_{ij}b_{ks}\right)a_{ik} \leq \left(\bigvee_{s=1, s \neq i}^{n} a_{ik}b_{ks}\right) = 0.$$

Hence, the columns of A^* form an orthonormal subset of $\mathcal{L}_m(\mathcal{B})$ and thus by Theorem (3.1), the columns of A form an orthonormal basis of $\mathcal{L}_n(\mathcal{B})$. So (1) implies (3) and the proof is complete. $\square$

As a consequence of Theorem (4.1), we see that invertible operators are always isomorphisms by Lemma (3.13), since they map the canonical basis to the orthonormal basis of their column vectors.

Theorem (4.1) allows us to establish the following remarkable fact: bases, as per Definition (2.9), are necessarily orthonormal, hence of cardinality the dimension of the Boolean vector space. Thus, for Boolean vector spaces, being a basis in a traditional sense is the same as being an orthonormal basis.

Theorem 4.2 *If $\mathcal{A} = \{\underline{a_1}, \ldots, \underline{a_m}\}$ is a basis for $\mathcal{L}_n(\mathcal{B})$ then $n = m$ and $\mathcal{A}$ is an orthonormal basis.*

PROOF. Define

$$T : \begin{vmatrix} \mathcal{L}_n\,(\mathcal{B}) & \longrightarrow & \mathcal{L}_m\,(\mathcal{B}) \\ \underline{b} & \longmapsto & (b_1, \ldots, b_m) \end{vmatrix}$$

where $\underline{b} = \sum b_i \underline{a_i}$. Now T is a linear bijection. Denote the inverse of T by S. It is easily checked that S is a linear bijection and $ST = I_n$ and $TS = I_m$. We conclude from Theorem (4.1) that $n = m$ and that matrix A_S of S is unitary. It is easily checked that $S\underline{\delta_i} = \underline{a_i}$ for $i = 1, \ldots, n$, i.e. the columns of A_S are the vectors $\underline{a_1}, \ldots, \underline{a_n}$ which by Theorem (4.1) form an orthonormal basis. □

We record the following observation as well:

Corollary 4.3 *Let $T : \mathcal{L}_n\,(\mathcal{B}) \longrightarrow \mathcal{L}_m\,(\mathcal{B})$ be a linear bijection. Then $n = m$ and T is an isomorphism.*

In view of Theorem (4.1), we introduce a type of matrix which will be of great interest to us in the next section. First, given A, B two $n \times n$ matrices, we shall say that $A \leq B$ when $\langle A\underline{a}, \underline{b} \rangle \leq \langle B\underline{a}, \underline{b} \rangle$ for all $\underline{a}, \underline{b} \in \mathcal{L}_n\,(\mathcal{B})$. The relation $\leq$ is easily seen to be an order on the set of $n \times n$ matrices. It is shown in [7] that $[a_{ij}]_{n \times n} \leq [b_{ij}]_{n \times n}$ if and only if $a_{ij} \leq b_{ij}$ for all $i, j \in \{1, \ldots, n\}$. Now we set:

Definition 4.4 *A matrix A is stochastic when $A^*A \geq I$ and $AA^* \leq I$.*

It is shown in [7] that products of stochastic matrices are stochastic matrices, and that a matrix is stochastic if and only if it maps stochastic vectors to stochastic vectors, or equivalently when its columns are stochastic vectors.

Note that A is unitary, or equivalently invertible, if and only if A and A^* are both stochastic. So unitarity is the same as bi-stochasticity. As an interesting observation, if we call a matrix A symmetric when $A^* = A$, then a symmetric stochastic matrix is always a unitary of order 2, namely $A^2 = I$. Conversely, if $A^2 = I$ then A is invertible with $A^{-1} = A^*$, so symmetric stochastic matrices are exactly given by unitaries of order 2, i.e. a reflection.

We have encountered such matrices before. Example (2.13) shows how to obtain such reflections. Let $\underline{a} = (a_1, \ldots . a_n)$ be a stochastic vector. Then the matrix

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \\ a_2 & a_3 & \cdots & a_1 \\ \vdots & \vdots & & \vdots \\ a_n & a_1 & \cdots & a_{n-1} \end{bmatrix}$$

22

is symmetric and stochastic.

Note however that the product of reflections need not be a reflection, as the product of the reflections $\begin{bmatrix} 0\ 1\ 0 \\ 1\ 0\ 0 \\ 0\ 0\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1\ 0\ 0 \\ 0\ 0\ 1 \\ 0\ 1\ 0 \end{bmatrix}$ is given by $\begin{bmatrix} 0\ 0\ 1 \\ 1\ 0\ 0 \\ 0\ 1\ 0 \end{bmatrix}$ which is not a reflection.

5 Invariant Vectors

Eigenvalues and eigenvectors of Boolean matrices have been previously studied [1,9,11,18]. Though invariant vectors are special case of eigenvectors, as far as we know the results in this section are new.

The following consequence of Lemma (3.8) will be used.

Lemma 5.1 *If $\underline{a} \in \mathcal{L}_n(\mathcal{B})$ then there exists an orthovector $\underline{b} \in \mathcal{L}_n(\mathcal{B})$ such that $\|\underline{b}\| = \|\underline{a}\|$ and $\underline{b} \leq \underline{a}$.*

PROOF. Apply Lemma (3.8) with the interval $[0, \|a\|]$ of $\mathcal{B}$ in lieu of $\mathcal{B}$. □

Let $A_1, \ldots, A_m$ be matrices on $\mathcal{L}_n(\mathcal{B})$ with $A_k = \left[a_{ij}^k\right]_{n \times n}$ $(k = 1, \ldots, m)$. The *joint trace* of $A_1, \ldots, A_m$ is

$$\mathrm{tr}\,(A_1, \ldots, A_m) = \bigvee_{i=1}^n a_{ii}^1 a_{ii}^2 \ldots a_{ii}^m.$$

In particular, the *trace* of $[a_{ij}]_{n \times n}$ is given by $\mathrm{tr}\,(A) = \bigvee_{i=1}^n a_{ii}$. A vector $\underline{b}$ is an *invariant vector* for A if $A\underline{b} = \underline{b}$, and more generally a *common invariant vector* of $A_1, \ldots, A_m$ if $A_i\underline{b} = \underline{b}$ for $i = 1, \ldots, m$.

Lemma 5.2 *Let A, B be two matrices on $\mathcal{L}_n(\mathcal{B})$. Then*

(1) $\mathrm{tr}\,(AB) = \mathrm{tr}(BA)$,
(2) If B is invertible then $\mathrm{tr}\,(BAB^*) = \mathrm{tr}(A)$.

23

PROOF. We compute

$$\mathrm{tr}(AB) = \bigvee_{i=1}^{n} (AB)_{ii} = \bigvee_{i=1}^{n} \bigvee_{k=1}^{n} a_{ik}b_{ki} = \bigvee_{k=1}^{n} \bigvee_{i=1}^{n} b_{ki}a_{ik} = \bigvee_{k=1}^{n} (BA)_{kk} = \mathrm{tr}\,(BA)\,.$$

If B is invertible then $B^{-1} = B^*$ by Theorem (4.1) and thus (1) implies (2). $\square$

Theorem 5.3 *Stochastic matrices $A_1, \ldots, A_m$ on $\mathcal{L}_n(\mathcal{B})$ have a common invariant stochastic vector if and only if $\mathrm{tr}\,(A_1, \ldots, A_m) = 1$.*

PROOF. Suppose $\underline{b}$ is a stochastic vector and $A_i\underline{b} = \underline{b}$ for $i = 1, \ldots, m$. Then $\bigvee_{j=1}^{n} a_{ij}^k b_j = b_i$ for $k = 1, \ldots, m$ and $i, j = 1, \ldots, n$. Multiplying both sides by b_i and since $\underline{b}$ is stochastic, we obtain $a_{ii}^k b_i = b_i$. Hence, $b_i \le a_{ii}^k$, $k = 1, \ldots, m$, so $b_i \le a_{ii}^1 a_{ii}^2 \ldots a_{ii}^m$. Therefore

$$\mathrm{tr}\,(A_1, \ldots, A_m) = \bigvee_{i=1}^{n} a_{ii}^1 a_{ii}^2 \ldots a_{ii}^m \ge \bigvee_{i=1}^{n} b_i = 1.$$

Conversely, suppose $\mathrm{tr}\,(A_1, \ldots, A_m) = \bigvee_{i=1}^{n} a_{ii}^1 a_{ii}^2 \ldots a_{ii}^m = 1$. By Lemma (3.8), there exists a stochastic vector $\underline{b} = (b_1, \ldots, b_n)$ such that $b_j \le a_{jj}^1 a_{jj}^2 \ldots a_{jj}^m$. Since $b_j \le a_{jj}^k$ ($k = 1, \ldots, m$) and A_k is stochastic, we have that $a_{ij}^k b_j = 0$ for $i \ne j$, $i, j = 1, \ldots, n$ and $k = 1, \ldots, m$. Hence

$$(A_k\underline{b})_i = \bigvee_{j=1}^{n} a_{ij}^k b_j = a_{ii}^k b_i = b_i.$$

Therefore, $A_k\underline{b} = \underline{b}$ ($k = 1, \ldots, m$) so $\underline{b}$ is a common invariant stochastic vector for $A_1, \ldots, A_m$. $\square$

Corollary 5.4 *A stochastic matrix A has an invariant stochastic vector if and only if $\mathrm{tr}(A) = 1$.*

Corollary 5.5 *If A is a stochastic matrix and B is invertible on $\mathcal{L}_n(\mathcal{B})$ then A has an invariant stochastic vector if and only if BAB^* does.*

Corollary 5.6 *A stochastic vector $\underline{b} = (b_1, \ldots, b_n)$ is a common invariant vector for stochastic matrices $A_1, \ldots, A_m$ if and only if $b_i \le a_{ii}^1 a_{ii}^2 \ldots a_{ii}^m$ for all $i = 1, \ldots, n$.*

Stochastic matrices $A_1, \ldots, A_m$ on $\mathcal{L}_n(\mathcal{B})$ are *simultaneously reducible* if there exists an invertible matrix B on $\mathcal{L}_n(\mathcal{B})$ and matrices $C_1, \ldots, C_m$ on

$\mathcal{L}_{n-1}(\mathcal{B})$ such that for $i = 1, \ldots, m$ we have

$$A_i = B \begin{bmatrix} 1 & 0 \\ 0 & C_i \end{bmatrix} B^*.$$

Notice that the matrices $C_1, \ldots, C_m$ are stochastic since $B^* A_i B = \begin{bmatrix} 1 & 0 \\ 0 & C_i \end{bmatrix}$. In particular, if there is only one matrix A in the above definition, we say that A is *reducible*.

Theorem 5.7 *Unitary matrices $A_1, \ldots, A_m$ on $\mathcal{L}_n(\mathcal{B})$ are simultaneously reducible if and only if $\mathrm{tr}(A_1, \ldots, A_m) = 1$.*

PROOF. If $A_1, \ldots, A_m$ are simultaneously reducible then $A_i = B \begin{bmatrix} 1 & 0 \\ 0 & C_i \end{bmatrix} B^*$ for some invertible matrix B and some matrix C_i, $i = 1, \ldots, m$. Since B is unitary, $B\underline{\delta_1}$ is stochastic and

$$A_i \left(B\underline{\delta_1} \right) = B \begin{bmatrix} 1 & 0 \\ 0 & C_i \end{bmatrix} \underline{\delta_1} = B\underline{\delta_1}$$

for $i = 1, \ldots, m$. Hence, $A_1, \ldots, A_m$ have a common invariant vector, and thus by Theorem (5.3) we have $\mathrm{tr}(A_1, \ldots, A_m) = 1$.

Conversely, assume that $\mathrm{tr}(A_1, \ldots, A_m) = 1$. Then $A_1, \ldots, A_m$ have a common stochastic invariant vector $\underline{b} = (b_1, \ldots, b_n)$ by Theorem (5.3). We define the symmetric stochastic matrix B by

$$B = \begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_n \\ b_2 & b_2^c & 0 & \cdots & 0 \\ b_3 & 0 & b_3^c & \cdots & 0 \\ \vdots & & & & \\ b_n & 0 & 0 & \cdots & b_n^c \end{bmatrix}.$$

Let $D_i = BA_iB$ for $i = 1, \ldots, m$. With the notation $A_k = \left[a_{ij}^k \right]_{n \times n}$, we compute the $(1,1)$ entry of D_i as

$$\bigvee_{j=1}^{n} b_{1j} \left(\bigvee_{r=1}^{n} a_{jr}^i b_{r1} \right) = \bigvee_{j=1}^{n} b_j \left(\bigvee_{r=1}^{n} a_{jr}^i b_r \right) = \bigvee_{j=1}^{n} b_j b_j = 1.$$

Since a product of unitary matrices is unitary, D_i is a unitary matrix and thus must have the form

$$D_i = \begin{bmatrix} 1 & 0 \\ 0 & C_i \end{bmatrix}$$

for some matrix C_i $(i = 1, \ldots, m)$. Since $A_i = BD_iB$ for $i = 1, \ldots, m$, we are finished. $\square$

Corollary 5.8 *A unitary matrix A is reducible if and only if* $\mathrm{tr}(A) = 1$.

We now give an example to show that Theorem (5.7) does not hold for stochastic matrices. Consider the stochastic matrix $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$. It is of trace 1, yet if it were reducible then there exists a unitary B such that $A = B \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} B^* = I$ which is a contradiction.

Notice if A is unitary and $\underline{b}$ is an invariant vector for A, then $\underline{b}$ is also an invariant vector for A^*. Indeed, $A\underline{b} = \underline{b}$ implies that $A^*\underline{b} = A^*A\underline{b} = \underline{b}$.

We now give an example that motivates the next result. Let $A = [a_{ij}]_{3\times 3}$ be a 3×3 symmetric stochastic matrix. We shall show that A has an invariant stochastic vector and hence A is reducible. Indeed, we have that

$$
\begin{aligned}
a_{11}^c a_{22}^c a_{33}^c &= (a_{12} \vee a_{13})(a_{12} \vee a_{32})(a_{13} \vee a_{23}) \\
&= (a_{12} \vee a_{13})(a_{12} \vee a_{23})(a_{13} \vee a_{23}) \\
&= (a_{12}a_{12} \vee a_{12}a_{23} \vee a_{13}a_{12} \vee a_{13}a_{23})(a_{13} \vee a_{23}) \\
&= a_{12}(a_{13} \vee a_{23}) = 0.
\end{aligned}
$$

Thus $\mathrm{tr}\,(A) = (a_{11}^c a_{22}^c a_{33}^c)^c = 0^c = 1$ so the result follows from Corollaries (5.4) and (5.8). The next theorem generalizes this calculation.

Theorem 5.9 *If A is an $n \times n$ symmetric stochastic matrix with n odd, then A has an invariant stochastic vector.*

PROOF. Since $A = [a_{ij}]_{n\times n}$ is symmetric, we have that

$$
\begin{aligned}
a_{11}^c a_{22}^c \ldots a_{nn}^c &= (a_{12} \vee a_{13} \vee \ldots \vee a_{1n})(a_{12} \vee a_{23} \vee \ldots \vee a_{2n}) \\
&\quad \ldots (a_{1n} \vee a_{2n} \vee \ldots \vee a_{n-1,n}).
\end{aligned}
$$

26

Since A is stochastic, we conclude that if we expand the right hand-side, the only nonzero terms are of the form $a_{ij}a_{ij}a_{rs}a_{rs}\ldots a_{uv}a_{uv}$ with $i \neq r$, $r \neq u$ and so on. By construction, there are n factors in this product. This would imply that n must be even. This is a contradiction, so all terms in the expansion are zero and thus

$$\operatorname{tr}(A) = (a_{11}^c a_{22}^c \ldots a_{nn}^c)^c = 1.$$

The result follows from Corollary (5.4). $\quad\square$

We now show that Theorem (5.9) does not hold if n is even. Consider the stochastic symmetric matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Then $\operatorname{tr}(A) = 0$ so A has no stochastic invariant vector. Now, generalizing, we see that if B is a $k \times k$ stochastic symmetric matrix, then $\begin{bmatrix} 0 & B \\ B & 0 \end{bmatrix}$ has trace 0 and thus has no invariant stochastic vector. Thus, for all even n there exists a stochastic symmetric $n \times n$ matrix with no invariant stochastic vector.

We can find more invariant stochastic vectors in the natural way. An *invariant orthogonal set* for matrices $A_1, \ldots, A_m$ on $\mathcal{L}_n(\mathcal{B})$ is a set of mutually orthogonal invariant vectors for $A_1, \ldots, A_m$. For example, if $\underline{b}, \underline{c}$ are stochastic vectors, then $\{\underline{b}, \underline{c}\}$ is an invariant orthogonal set for the unitary matrix A if and only if $c_i \leq a_{ii}b_i^c$ for $i = 1, \ldots, n$ or equivalently $b_i \leq a_{ii}c_i^c$ for $i = 1, \ldots, n$.

Theorem 5.10 *A unitary matrix A possesses an invariant orthogonal set of m stochastic vectors if and only if there exists an invertible matrix B such that*

$$A = B \begin{bmatrix} I_m & 0 \\ 0 & C \end{bmatrix} B^*$$

where I_m is the identity operator on $\mathcal{L}_m(\mathcal{B})$.

PROOF. Suppose A is an $n \times n$ matrix with the given form. Then $m \leq n$ and we can define $\underline{b}_j = B\underline{\delta}_j$, $j = 1, \ldots, m$. We conclude from Theorem (4.1) that $\underline{b}_1, \ldots, \underline{b}_m$ are stochastic vectors and we have $A\underline{b}_j = \underline{b}_j$ for $j = 1, \ldots, m$ by construction. Moreover, for $i \neq j$ we have

$$\left\langle \underline{b}_j, \underline{b}_i \right\rangle = \left\langle B\underline{\delta}_i, B\underline{\delta}_j \right\rangle = \left\langle B^*B\underline{\delta}_i, \underline{\delta}_j \right\rangle = \left\langle \underline{\delta}_i, \underline{\delta}_j \right\rangle = 0.$$

Hence $\{\underline{b_1}, \ldots, \underline{b_m}\}$ is an invariant orthogonal set of stochastic vectors.

Conversely, suppose that A possesses an invariant orthogonal set of stochastic vectors $\{\underline{b_1}, \ldots, \underline{b_m}\}$ and write $\underline{b_j} = (b_{1j}, \ldots, b_{nj})$ for $j = 1, \ldots, n$. Letting

$$
B_1 = \begin{bmatrix} b_{11} & b_{21} & b_{31} & \cdots & b_{n1} \\ b_{21} & b_{21}^c & 0 & \cdots & 0 \\ b_{31} & 0 & b_{31}^c & \cdots & 0 \\ \vdots & & & & \vdots \\ b_{n1} & 0 & \cdots & 0 & b_{n1}^c \end{bmatrix}
$$

and $D_1 = B_1 A B_1$ as in the proof of Theorem (5.7), we have that

$$
D_1 = \begin{bmatrix} 1 & 0 \\ 0 & C_1 \end{bmatrix}
$$

where C_1 is a stochastic matrix and $A = B_1 D_1 B_1$. Letting $C_1 = [c_{ij}]_{n \times n}$ and $D_1 = [d_{ij}]_{n \times n}$ we have

$$
\begin{aligned}
c_{11} = d_{22} &= \bigvee_{j=1}^{2} b_{2j} \left(\bigvee_{k=1}^{2} a_{jk} b_{k2} \right) \\
&= b_{21} (a_{11} b_{21} \vee a_{12} b_{21}^c) \vee b_{21}^c (a_{21} b_{21} \vee a_{22} b_{21}^c) \\
&= a_{11} b_{21} \vee a_{22} b_{21}^c.
\end{aligned}
$$

More generally

$$
c_{ii} = d_{i+1,i+1} = a_{ii} b_{i+1,1} \vee a_{i+1,i+1} b_{i+1,1}^c
$$

for $i = 1, \ldots, n - 1$. Hence

$$
\operatorname{tr}(C_1) = \bigvee_{i=1}^{n-1} \left(a_{ii} b_{i+1,1} \vee a_{i+1,i+1} b_{i+1,1}^c \right) = \bigvee_{i=1}^{n-1} a_{ii} b_{i,1}^c.
$$

Since $b_{i2} \le a_{ii} b_{i1}^c$ $(i = 1, \ldots, n)$, we conclude that $\underline{b_2}$ is an invariant stochastic vector of C_1 by Corollary (5.6). Hence, there exists a symmetric stochastic matrix B_2 such that

$$
C_1 = B_2 \begin{bmatrix} 1 & 0 \\ 0 & C_2 \end{bmatrix} B_2.
$$

It follows that

$$A = B_1 \begin{bmatrix} 1 & 0 \\ 0 & B_2 \begin{bmatrix} 1 & 0 \\ 0 & C_2 \end{bmatrix} B_2 \end{bmatrix} B_1$$

$$= B_1 \begin{bmatrix} 1 & 0 \\ 0 & B_2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & C_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & B_2 \end{bmatrix} B_1$$

$$= B_3 \begin{bmatrix} I_2 & 0 \\ 0 & C_2 \end{bmatrix} B_3^*$$

with $B_3 = B_1 \begin{bmatrix} 1 & 0 \\ 0 & B_2 \end{bmatrix}$. The proof is then completed by a simple induction.   □

Theorem (5.10) can be easily generalized to the following:

Corollary 5.11 *Unitary matrices $A_1, \ldots, A_m$ possess an invariant orthogonal set of stochastic vectors if and only if there exists an invertible matrix B and matrices $C_1, \ldots, C_n$ such that*

$$A_i = B \begin{bmatrix} I_m & 0 \\ 0 & C_i \end{bmatrix} B^*$$

for $i = 1, \ldots, m$ and I_m the identity operator on $\mathcal{L}_m(\mathcal{B})$.

We now illustrate Theorem (5.10) with an example. Let $\mathcal{B}$ be the power set of $\{1, 2, 3, 4, 5\} = \Omega$ endowed with its natural Boolean algebra structure. Consider the stochastic symmetric matrix A over $\mathcal{L}_5(\mathcal{B})$ defined by

$$A = \begin{bmatrix} \{1\} & \{2\} & \{3\} & \{4\} & \{5\} \\ \{2\} & \{4,5\} & \emptyset & \emptyset & \{1,3\} \\ \{3\} & \emptyset & \{4,5\} & \{1\} & \{2\} \\ \{4\} & \emptyset & \{1\} & \{2,3,5\} & \emptyset \\ \{5\} & \{1,3\} & \{2\} & \emptyset & \{4\} \end{bmatrix}.$$

There are many stochastic invariant vectors for A and we choose

$$\underline{b} = (\{1\}, \emptyset, \emptyset, \{2, 3, 5\}, \{4\}).$$

29

We now form the stochastic symmetric matrix

$$
B = \begin{bmatrix}
\{1\} & \emptyset & \emptyset & \{2,3,5\} & \{4\} \\
\emptyset & \Omega & \emptyset & \emptyset & \emptyset \\
\emptyset & \emptyset & \Omega & \emptyset & \emptyset \\
\{2,3,5\} & \emptyset & \emptyset & \{1,4\} & \emptyset \\
\{4\} & \emptyset & \emptyset & \emptyset & \{1,2,3,5\}
\end{bmatrix}
$$

We can then reduce A by

$$
BAB = \begin{bmatrix}
\Omega & \emptyset & \emptyset & \emptyset & \emptyset \\
\emptyset & \{4,5\} & \emptyset & \{2\} & \{1,3\} \\
\emptyset & \emptyset & \{4,5\} & \{1,3\} & \{2\} \\
\emptyset & \{2\} & \{1,3\} & \emptyset & \{4,5\} \\
\emptyset & \{1,3\} & \{2\} & \{4,5\} & \emptyset
\end{bmatrix}.
$$

Thus

$$
A = B \begin{bmatrix} 1 & 0 \\ 0 & C \end{bmatrix} B
$$

yet $\operatorname{tr}(C) = \{4,5\} \neq \Omega$ so no further reduction is possible.

6   Powers of Stochastic Matrices

As mentioned in section 2, powers of stochastic matrices may be important for the study of Boolean Markov chains. Various applications of powers of lattice matrices are discussed in [2,19]. If A is a Boolean matrix, the smallest natural number p such that there exists a natural number e with $A^{e+p} = A^e$ is called the *period* of A and is denoted by $p(A)$. The smallest natural number e such that $A^{e+p(A)} = A^e$ is called the *exponent* or *index* of A and is denoted by $e(A)$. It is known that for any $n \times n$ Boolean matrix A, both $p(A)$ and $e(A)$ exist and $e(A) \leq (n-1)^2 + 1$ [2,19]. We shall use:

Definition 6.1 *Let $n \in \mathbb{N}$. The least common multiple of $\{1,2,\ldots,n\}$ is denoted by $[n]$.*

It is also known that $p(A)$ divides $[n]$.

In this section, we show that for a stochastic matrix, we can improve the upper bound for $e(A)$ to $e(A) \leq n - 1$. Although we do not improve on

$p(A)|[n]$, we give an alternative proof of this result for stochastic matrices because it is embedded in our proof that $e(A) \le n - 1$.

If A is a 2×2 matrix, then it follows from the previous known results that $A^4 = A^2$. Moreover, it is easy to check that if A is a 2×2 stochastic matrix then $A^3 = A$. In the same way, for 3×3 matrix A we have $A^{11} = A^5$. However, one can check that if A is a 3×3 stochastic matrix then $A^8 = A^2$. Displaying the first eight powers of A would be cumbersome, so we refrain from doing so. However, we can easily prove the special case that $A^6 = I$ for any unitary 3×3 matrix A. In this case, we have

$$A = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$$

where each row and column is a stochastic vector. We then have

$$A^2 = \begin{bmatrix} a_1 \vee a_2 b_1 \vee a_3 c_1 & b_3 c_1 & b_1 c_2 \\ a_3 c_2 & a_2 b_1 \vee b_2 \vee b_3 c_2 & a_2 c_1 \\ b_3 a_2 & a_3 b_1 & a_3 c_1 \vee b_3 c_2 \vee c_3 \end{bmatrix},$$

$$A^3 = \begin{bmatrix} a_1 \vee a_3 b_1 c_2 \vee a_2 b_3 c_1 & a_2 b_1 & a_3 c_1 \\ a_2 b_1 & a_2 b_3 c_1 \vee b_2 \vee a_3 b_1 c_2 & b_3 c_2 \\ a_3 c_1 & b_3 c_2 & a_3 b_1 c_2 \vee a_2 b_3 c_1 \vee c_3 \end{bmatrix}.$$

Since A^3 is symmetric and unitary (as a product of unitary, or by inspection), we conclude that $A^6 = A^3 A^3 = I$.

From these observations and our work in Section 5, we can already draw some interesting conclusions. For example, let A be a 3×3 unitary matrix with $\mathrm{tr}(A) = 1$. Applying Corollary (5.8), there exists an invertible matrix B and a 2×2 unitary matrix C such that

$$A = B \begin{bmatrix} 1 & 0 \\ 0 & C \end{bmatrix} B^*. \tag{6.1}$$

Since C is symmetric (all 2×2 unitaries are), we have $C^2 = I$ and thus

$$A^2 = B \begin{bmatrix} 1 & 0 \\ 0 & C^2 \end{bmatrix} B^* = I.$$

We conclude that any 3×3 unitary matrix A with $\operatorname{tr}(A) = 1$ is symmetric.

As another example, let A be a 4×4 unitary matrix with $\operatorname{tr}(A) = 1$. As before, there exists an invertible matrix B such that (6.1) holds where C is now a 3×3 unitary matrix. Since $C^6 = I$, we conclude that $A^6 = I$ and thus A^3 is symmetric.

We now begin the proof of the main result of this section. Let $A = [a_{ij}]_{n \times n}$ be a stochastic matrix on $\mathcal{L}_n(\mathcal{B})$. We shall use:

Definition 6.2 *A nonzero element of $\mathcal{B}$ of the form*

$$a_{i_1 1} a_{i_2 2} \ldots a_{i_n n}$$

for $i_1, \ldots, i_n \in \{1, \ldots, n\}$ is called an atom of A.

Of course there are a finite numbers of atoms of A.

Lemma 6.3 *Let $A = [a_{ij}]_{n \times n}$ be a stochastic matrix on $\mathcal{L}_n(\mathcal{B})$. Let $\omega_1, \ldots, \omega_m$ be the distinct atoms of A.*

(1) If $i, j \in \{1, \ldots, m\}$ and $i \neq j$ then $\omega_i \omega_j = 0$,

(2) $\bigvee_{i=1}^{m} \omega_i = 1$,

(3) For all $i, j \in \{1, \ldots, n\}$ we have $a_{ij} = \bigvee \{\omega_k : \omega_k \leq a_{ij}\}$,

(4) If $\omega_i \leq a_{kj}$ then $A\omega_i \underline{\delta_j} = \omega_i \underline{\delta_k}$.

PROOF. For (1), letting $\omega_i = a_{i_1 1} a_{i_2 2} \ldots a_{i_n n}$ and $\omega_j = a_{j_1 1} a_{j_2 2} \ldots a_{j_n n}$, if $i \neq j$ then $i_k \neq j_k$ for some $k \in \{1, \ldots, n\}$ and thus $\omega_j \omega_i = 0$ since $a_{i_k k} a_{j_k k} = 0$.

(2) will follow from (3). For (3), since

$$a_{11} = \bigvee \{a_{11} (a_{i_2 2} \ldots a_{i_n n}) : i_2, \ldots, i_n = 1, \ldots, n\}$$

as A is stochastic, the results holds for a_{11}. It holds similarly for a_{ij} with $i, j \in \{1, \ldots, n\}$. Last, for (4), if $\omega_i \leq a_{kj}$ then

$$A\omega_i \underline{\delta_j} = \omega_i A \underline{\delta_j} = \omega_i (a_{1j}, a_{2j}, \ldots, a_{nj}) = (\omega_i a_{1j}, \ldots, \omega_i a_{nj})$$
$$= \omega_i a_{jk} \underline{\delta_k} = \omega_i \underline{\delta_k}.$$

This concludes our proof. $\square$

The main result for this section is:

Theorem 6.4 *If A is a stochastic $n \times n$ matrix then $A^{[n]+n-1} = A^{n-1}$.*

PROOF. Let $\omega_1, \ldots, \omega_m$ be the distinct atoms of A. By Lemma (6.3,2), we have $\underline{\delta_i} = \sum_{j=1}^{m} \omega_j \underline{\delta_i}$ for all $i \in \{1, \ldots n\}$. Since $\{\underline{\delta_1}, \ldots, \underline{\delta_n}\}$ is a basis for $\mathcal{L}_n(\mathcal{B})$, the set $\{\omega_j \underline{\delta_i} : i = 1, \ldots n; j = 1, \ldots m\}$ is a generating set of $\mathcal{L}_n(\mathcal{B})$. Set $r = [n] + n - 1$. If we can show that $A^r \omega_j \underline{\delta_i} = A^{n-1} \omega_j \underline{\delta_i}$ for $i = 1, \ldots n$ and $j = 1, \ldots m$ then we are done.

Consider first $\omega_1 \underline{\delta_1}$ and call the vectors $A^0 \omega_1 \underline{\delta_1}$, $A \omega_1 \underline{\delta_1}$, $A^2 \omega_1 \underline{\delta_1}, \ldots, A^{n-1} \omega_1 \underline{\delta_1}$ the *iterates* of A at $\omega_1 \underline{\delta_1}$. By Lemma (6.3,4), the iterates of $\omega_1 \underline{\delta_1}$ have the form: $\omega_1 \underline{\delta_1}, \omega_1 \underline{\delta_{i_1}}, \omega_1 \underline{\delta_{i_2}}, \ldots, \omega_1 \underline{\delta_{i_{n-1}}}$ for $i_1, \ldots, i_{n-1} \in \{1, \ldots, n\}$.

Suppose there is only one distinct iterate of A at $\omega_1 \underline{\delta_1}$. Then

$$A \omega_1 \underline{\delta_1} = \omega_1 \underline{\delta_{i_1}} = \omega_1 \underline{\delta_1}.$$

Then we have

$$A^{n-1} \omega_1 \underline{\delta_1} = A^n \omega_1 \underline{\delta_1} = \ldots = A^r \omega_1 \underline{\delta_1}. \tag{6.2}$$

Suppose now there are two distinct iterates of A at $\omega_1 \underline{\delta_1}$. Then $\omega_1 \underline{\delta_1} \neq \omega_1 \underline{\delta_{i_1}}$. If $A \omega_1 \underline{\delta_{i_2}} = \omega_1 \underline{\delta_{i_1}}$ then

$$\omega_1 \underline{\delta_{i_3}} = A \omega_1 \underline{\delta_{i_2}} = A \omega_1 \underline{\delta_{i_1}} = \omega_1 \underline{\delta_{i_2}} = \omega_1 \underline{\delta_{i_1}}$$

and we can conclude again that (6.2) holds. Otherwise, $A^2 \omega_1 \underline{\delta_1} = \omega_1 \underline{\delta_1}$ and thus $A^{n-1} \omega_1 \underline{\delta_1} = \omega_1 \underline{\delta_1}$ or $A^{n-1} \omega_1 \underline{\delta_1} = \omega_1 \underline{\delta_{i_1}}$. Either way, we have

$$A^{2+(n-1)} \omega_1 \underline{\delta_1} = A^{n-1} \omega_1 \underline{\delta_1}. \tag{6.3}$$

Suppose instead that there are three distinct iterates of A at $\omega_1 \underline{\delta_1}$. Thus $\omega_1 \underline{\delta_1}$, $\omega_1 \underline{\delta_{i_1}}$ and $\omega_1 \underline{\delta_{i_2}}$ are distinct. If $\omega_1 \underline{\delta_{i_3}} = \omega_1 \underline{\delta_{i_2}}$ then $A^r \omega_1 \underline{\delta_1} = A^{n-1} \omega_1 \underline{\delta_1} = \omega_1 \underline{\delta_{i_3}}$ so (6.2) holds again. If $\omega_1 \underline{\delta_{i_1}} = \omega_1 \underline{\delta_{i_3}}$ then $A \omega_1 \underline{\delta_1} \in \{\omega_1 \underline{\delta_1}, \omega_1 \underline{\delta_{i_2}}\}$ and (6.3) holds. If $\omega_1 \underline{\delta_1} = \omega_1 \underline{\delta_{i_3}}$ then $A^{n-1} \omega_1 \underline{\delta_1} \in \{\omega_1 \underline{\delta_1}, \omega_1 \underline{\delta_{i_1}}, \omega_1 \underline{\delta_{i_2}}\}$ and we have

$$A^{3+n-1} \omega_1 \underline{\delta_1} = A^{n-1} \omega_1 \underline{\delta_1}. \tag{6.4}$$

Generalizing this observation, suppose that all the iterates $\omega_1 \underline{\delta_1}, \omega_1 \underline{\delta_{i_1}}, \ldots, \omega_1 \underline{\delta_{i_{n-1}}}$ are distinct. Since there are only n possibilities for $A^n \omega_1 \underline{\delta_1}$, we conclude that $A^n \omega_1 \underline{\delta_1} = \omega_1 \underline{\delta_1}$ or $\omega_1 \underline{\delta_{i_j}}$ for some $j \in \{1, \ldots, n-1\}$. But then

$$A^{t+(n-1)} \omega_1 \underline{\delta_1} = A^{n-1} \omega_1 \underline{\delta_1} \tag{6.5}$$

33

for some $t \in \{1, 2, \ldots, n\}$. Notice (6.3) and (6.4) are special cases of (6.5).

Let us now suppose (6.5) holds for some $t \in \{1, \ldots, n\}$. Since $r = kt + (n-1)$ for some $k \in \mathbb{N}$ we have

$$
\begin{aligned}
A^r \omega_1 \underline{\delta_1} &= A^{kt+n-1} \omega_1 \underline{\delta_1} = \left(A^t\right)^k A^{n-1} \omega_1 \underline{\delta_1} \\
&= \left(A^t\right)^{k-1} A^t A^{n-1} \omega_1 \underline{\delta_1} = \left(A^t\right)^{k-1} A^{n-1} \omega_1 \underline{\delta_1} \\
&= \left(A^t\right)^{k-2} A^t A^{n-1} \omega_1 \underline{\delta_1} = \left(A^t\right)^{k-2} A^{n-1} \omega_1 \underline{\delta_1} \\
&= \ldots = A^{n-1} \omega_1 \underline{\delta_1}.
\end{aligned}
$$

In a similar way, we can prove that $A^r \omega_j \underline{\delta_i} = A^{n-1} \omega_j \underline{\delta_i}$ for $j = 1, \ldots, m$ and $i = 1, \ldots, n$, so the proof is complete. $\square$

Corollary 6.5 *If A is an $n \times n$ unitary matrix then $A^{[n]} = I$.*

As examples, $A^{15} = A^3$ for any 4×4 stochastic matrix and $A^{64} = A^4$ for any 5×5 stochastic matrix. We now give a final example. Let (a, b, c) be a stochastic vector and form the stochastic matrix

$$
A = \begin{bmatrix} b \vee c & a & 0 \\ a & b & a \\ 0 & c & b \vee c \end{bmatrix}.
$$

We then have

$$
A^2 = \begin{bmatrix} 1 & 0 & a \\ 0 & a \vee b & 0 \\ 0 & c & c \vee b \end{bmatrix}
$$

and $A^{2n+1} = A$, $A^{2n} = A^2$ for $n \in \mathbb{N}$. This example illustrates an important difference between Boolean Markov chains and traditional Markov chains given by real stochastic matrices. An important property of traditional Markov chains is that the sites (called states in the traditional case) can be decomposed into equivalence classes. This is important because sites in the same equivalence class share a similar behavior [3].

To be precise, let $M = [p_{ij}]_{n \times n}$ be a real stochastic matrix, i.e. $p_{ij} \geq 0$ and $\sum_{i=1}^{n} p_{ij} = 1$ for every $j = 1, \ldots, n$. The real p_{ij} represents the transition probability from site j to site i. A site i is *accessible* from a site j if there exists $n \in \mathbb{N}$ such that $(M^n)_{ij} > 0$, and we then denote $j \to i$. It is easy to check that $\to$ is transitive and that the relation $\longleftrightarrow$ defined by $i \longleftrightarrow j \iff (i \to j \wedge j \to i)$ is an equivalence relation on the sites of the Markov chain.

Let us now extend this concept to Boolean Markov chains whose transition matrix is a Boolean stochastic matrix A. Thus, $j \to i$ whenever $(A^n)_{ij} > 0$ for some $n \in \mathbb{N}$. For the example above, we note that $1 \to 2$ and $2 \to 3$ yet $1 \not\to 3$. Thus $\to$ is not transitive. If we define $\longleftrightarrow$ by $i \longleftrightarrow j \iff (i \to j \wedge j \to i)$ then we have, in the above example, that in fact $1 \longleftrightarrow 2$ and $2 \longleftrightarrow 3$ yet $1 \not\leftrightarrow 3$. Hence $\longleftrightarrow$ is no longer an equivalence relation.

References

[1] T. S. Blyth, *On eigenvectors of Boolean Matrices*, Proc. Roy. Soc. Edinburgh **Sect. A67** (1967), 196–204.

[2] K. Cechlárová, *Powers of matrices over distributive lattices — a review*, Fuzzy Sets Sys. **138** (2003), 627–641.

[3] Y. Give'on, *Lattice matrices*, Information and Control **7** (1964), 477–484.

[4] S. Gudder, *Quantum Markov chains*, Submitted.

[5] ———, *Sequential products of quantum measurments*, Rep. Math. Phys. (To appear).

[6] P. V. Jagannadham, *Linear transformations on Boolean vector spaces*, Math. Ann. **16** (1966), 240–247.

[7] R. D. Luce, *A note on Boolean matrix theory*, Proc. Amer. Math. Soc. **3** (1952), 382–388.

[8] D. E. Rutherford, *Inverses of Boolean matrices*, Proc. Glasgow Math. Assoc. **6** (1963), 49–53.

[9] ———, *The eigenvalue problem for Boolean matrices*, Proc. Roy. Soc. Edinburgh **Sect. A67** (1963/1965), 25–38.

[10] ———, *Orthogonal boolean matrices*, Proc. Roy. Soc. Edinburgh **Sect A67** (1964/1965), 126–135.

[11] R. L. Sindak, *Eigenvectors and maximal vectors in Boolean vector spaces*, Proc. Amer. Math. Soc. **47** (1975), 323–328.

[12] L. A. Skornyakov, *Invertible matrices over distributive structures*, (Russian) Sibirsk. Mat. Zh. **27** (1986), 182–185, English translation: Siberian Math. J. **27** (1986), 289–292.

[13] D. Stirzaker, *Stochastic Processes and Models*, Oxford Univ. Press, 2005.

[14] N. V. Subrahmanyam, *Boolean vector spaces I*, Math. Z. **83** (1964), 422–433.

[15] ———, *Boolean vector spaces II*, Math. Z. **87** (1965), 401–419.

[16] ———, *Boolean vector spaces III*, Math. Z. **100** (1967), 295–313.

[17] J. H. M. Wedderburn, *Boolean linear associative algebra*, Ann. Math. **35** (1934), 185–194.

[18] Yi-Jia Tan, *Eigenvalues and eigenvectors for matrices over distributive lattices*, Linear Algebra Appl. **283** (1998), 257–272.

[19] Yi-Jia Tan, *On the powers of matrices over a distributive lattices*, Linear Algebra Appl. **336** (2001), 1–14.

[20] M. Yoeli, *A note on a generalization of Boolean matrix theory*, Amer. Math. Monthly **68** (1961), 552–557.