Linear Algebra and its Applications 578 (2019) 27–52

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

The Birkhoff theorem for unitary matrices of prime-power dimension

Alexis De Vos *, Stijn De Baerdemacker

ARTICLE INFO

Article history:
Received 20 December 2018
Accepted 6 May 2019
Available online 10 May 2019
Submitted by R. Brualdi

MSC:
15A21
15A51

Keywords:
Unitary matrix
Permutation matrix
Birkhoff theorem

ABSTRACT

The unitary Birkhoff theorem states that any unitary matrix with all row sums and all column sums equal unity can be decomposed as a weighted sum of permutation matrices, such that both the sum of the weights and the sum of the squared moduli of the weights are equal to unity. If the dimension n of the unitary matrix equals a power of a prime p, i.e. if $n = p^w$, then the Birkhoff decomposition does not need all $n!$ possible permutation matrices, as the epicirculant permutation matrices suffice. This group of permutation matrices is isomorphic to the general affine group $\mathrm{GA}(w, p)$ of order only $p^w(p^w - 1)(p^w - p)...(p^w - p^{w-1}) \ll (p^w)!$.

1. Introduction

Let $\mathrm{D}(n)$ be the semigroup of $n \times n$ doubly stochastic matrices; let $\mathrm{P}(n)$ be the group of $n \times n$ permutation matrices. Birkhoff [1] has demonstrated

* Corresponding author.
E-mail address: alexis.devos@ugent.be (A. De Vos).

https://doi.org/10.1016/j.laa.2019.05.007

Theorem 1. *Every D(n) matrix D can be written*

$$D = \sum_{\sigma} c_\sigma P_\sigma$$

with all $P_\sigma \in P(n)$ *and the weights* c_σ *real, satisfying both* $0 \leq c_\sigma \leq 1$ *and* $\sum_\sigma c_\sigma = 1$.

The question arises whether a similar theorem holds for matrices from the unitary group U(n). This question is discussed by De Baerdemacker et al. [2] [3]. For this purpose, the subgroup XU(n) of U(n) is introduced [4] [5]. It consists of all U(n) matrices with all line sums (i.e. all row sums and all column sums) equal to 1. Whereas U(n) is an n^2-dimensional Lie group, the group XU(n) is only $(n-1)^2$-dimensional. A unitary Birkhoff theorem has been proved for XU(n) matrices [2] [3]. Remarkable is the fact that the case $n = p$ with p an arbitrary prime [3] has been treated in a very different way from the case where n is an arbitrary integer [2]. As a result, the decomposition, tailored to prime numbers [3], can be restricted to n^2 terms, whereas the general case [2] leads to a summation over all $n!$ (or at least over $n!/2$) permutation matrices, albeit with a large number of degrees of freedom. In the present paper, we will treat the two cases in a unified way. Moreover, the unified approach will be applied to the case $n = p^w$, i.e. n equal to an arbitrary power w of an arbitrary prime p.

In general, the Birkhoff theorem for unitary matrices is based on the two following lemmas. Let G(n) be a finite subgroup of XU(n).

Lemma 1. *If an XU(n) matrix X can be written*

$$X = \sum_{\sigma} c_\sigma G_\sigma$$

with all $G_\sigma \in G(n)$*, then the weights* c_σ *satisfy* $\sum_\sigma c_\sigma = 1$.

The proof is trivial: all line sums of G_σ equal unity; therefore, all line sums of the matrix $c_\sigma G_\sigma$ equal c_σ and thus all line sums of the matrix $\sum_\sigma c_\sigma G_\sigma$ are equal to $\sum_\sigma c_\sigma$. As all line sums of X are equal to 1, we thus need $\sum_\sigma c_\sigma = 1$.

Lemma 2. *If every XU(n) matrix X can be written*

$$X = \sum_{\sigma} a_\sigma G_\sigma$$

with all $G_\sigma \in G(n)$*, then there exists a decomposition*

$$X = \sum_{\sigma} b_\sigma G_\sigma \ ,$$

such that not only $\sum_\sigma b_\sigma = 1$*, but also* $\sum_\sigma |b_\sigma|^2 = 1$.

This fact follows from the Klappenecker–Rötteler theorem [6].

Lemmas 1 and 2 have the following consequence: it suffices to find an appropriate finite subgroup $\mathrm{G}(n)$ of $\mathrm{XU}(n)$ such that every member X of $\mathrm{XU}(n)$ can be written as $X = \sum_\sigma a_\sigma G_\sigma$, in order to guarantee the existence of a decomposition $X = \sum_\sigma b_\sigma G_\sigma$ with both $\sum_\sigma b_\sigma = 1$ and $\sum_\sigma |b_\sigma|^2 = 1$. In case there exist more than one candidate subgroup $\mathrm{G}(n)$, it is profitable to choose the smallest one, such that the sum $\sum_\sigma b_\sigma G_\sigma$ is as short as possible.

In the following sections, i.e. in Sections 3 to 6, we will present such subgroup $\mathrm{G}(n)$ in three different cases:

- arbitrary dimension n,
- prime dimension n, and
- prime-power dimension n.

However, for finding $\mathrm{G}(n)$, we will need some properties of $\mathrm{XU}(n)$ matrices, which will be presented in next section, i.e. Section 2.

2. The group XU(n)

Remark 1. For sake of convenience, in the present paper, the rows and colums of a matrix are not numbered starting from 1, but instead starting from 0. Thus the upper-left entry of any $m \times m$ square matrix A is $A_{0,0}$ and its lower-right entry is $A_{m-1,m-1}$.

We recall that the group $\mathrm{XU}(n)$ is an $(n-1)^2$-dimensional subgroup of the n^2-dimensional unitary group $\mathrm{U}(n)$. Any member X of $\mathrm{XU}(n)$ can be written

$$X = T \begin{pmatrix} 1 & \\ & U \end{pmatrix} T^{-1} , \tag{1}$$

where U is a member of $\mathrm{U}(n-1)$ and where the constant matrix T is an $n \times n$ transformation matrix with following properties:

- it is unitary;
- all its upper-row entries $T_{0,k}$ and left-column entries $T_{j,0}$ are equal.

As a consequence, we have

$$|T_{0,k}|^2 = |T_{j,0}|^2 = 1/n \ .$$

A possible choice for T is $1/\sqrt{n}$ times a dephased complex Hadamard matrix [7]. Thus (1) constitutes a 1-to-1 mapping between X and U. Because of

$$T_{j,0} = T_{0,k} = 1/\sqrt{n} \ , \tag{2}$$

eqn (1) leads to

$$X_{k,l} = \frac{1}{n} + \sum_{r=1}^{n-1}\sum_{s=1}^{n-1} T_{k,r}\, U_{r-1,s-1}\, (T^{-1})_{s,l} \ .$$

With T being unitary, i.e. with $T^{-1} = T^\dagger$, this becomes

$$X_{k,l} = \frac{1}{n} + \sum_{r=1}^{n-1}\sum_{s=1}^{n-1} U_{r-1,s-1}\, T_{k,r}\overline{T_{l,s}} \ .$$

We thus can write the matrix X as a sum of $1 + (n-1)^2$ matrices:

$$X = W + \frac{1}{n}\sum_{r=1}^{n-1}\sum_{s=1}^{n-1} U_{r-1,s-1} M_{r,s} \ , \tag{3}$$

where W is the $n \times n$ van der Waerden matrix, i.e. the doubly stochastic matrix with all entries equal to $\frac{1}{n}$, and where $M_{r,s}$ is an $n \times n$ matrix defined by

$$(M_{r,s})_{k,l} = n\, T_{k,r}\overline{T_{l,s}} \ . \tag{4}$$

In Sections 4, 5, and 6, the matrix W and the $(n-1)^2$ matrices $M_{r,s}$ will be decomposed as a weighted sum of matrices G_σ, such that we will obtain a decomposition of X of the desired form $\sum_\sigma a_\sigma G_\sigma$ and will be able to apply Lemmas 1 and 2.

The labels r and s of the matrix $M_{r,s}$ run from 1 to $n-1$, in contrast to the indices k and l of its entries, which run from 0 to $n-1$. We thus have $(n-1)^2$ such matrices, each having n^2 entries. Each entry of the matrix $M_{r,s}$ equals the leftmost entry of its row times the uppermost entry of its column. Taking into account (2), one indeed easily checks

$$(M_{r,s})_{0,l}\ (M_{r,s})_{k,0} = (M_{r,s})_{k,l} \ . \tag{5}$$

Both the zeroth row and the zeroth column of $M_{r,s}$ equal a line of the Hadamard matrix T (up to complex conjugation and up to the factor $\sqrt{n}$):

$$\begin{aligned} (M_{r,s})_{0,l} &= \sqrt{n}\ \overline{T_{l,s}} \\ (M_{r,s})_{k,0} &= \sqrt{n}\ T_{k,r} \ . \end{aligned} \tag{6}$$

Because T is $1/\sqrt{n}$ times a Hadamard matrix, we have $|T_{l,s}| = 1/\sqrt{n}$ and $|T_{k,r}| = 1/\sqrt{n}$, such that $|(M_{r,s})_{0,l}| = 1$ and $|(M_{r,s})_{k,0}| = 1$, and thus, because of (5), we conclude that all entries $(M_{r,s})_{k,l}$ have unit modulus.

3. Underlying framework

In the present section, we consider an arbitrary doubly transitive group $\mathrm{G}(n)$ of $n \times n$ permutation matrices. We denote by N the order of the group. We generalize the ideas and computations in Reference [2], where $\mathrm{G}(n)$ is equal to the group $\mathrm{P}(n)$ of all $n \times n$ permutation matrices, thus $\mathrm{G}(n)$ being isomorphic to the symmetric group $\mathbf{S}_n$ and N being equal to $n!$.

In the next three sections, we will apply the Lemmas 1 and 2 to three different choices of $\mathrm{G}(n)$:

- In case of arbitrary n, we choose the group of all $n \times n$ permutation matrices (i.e. a group isomorphic to the symmetric group $\mathbf{S}_n$). See Section 4.
- In case of n equal to some prime p, we choose the group of all $n \times n$ supercirculant permutation matrices (i.e. a group isomorphic to a semidirect-product group $\mathbf{C}_n : \mathbf{C}_{n-1}$). See Section 5.
- In case of n equal to some power w of some prime p (i.e. equal to p^w), we choose the group of all $n \times n$ epicirculant permutation matrices (i.e. a group isomorphic to the general affine group $\mathrm{GA}(w, p)$). See Section 6.

The meaning of the words 'supercirculant' and 'epicirculant' will be made clear below. The mentioned groups are doubly transitive, as it is known that the symmetric group $\mathbf{S}_n$ is n-transitive, the alternating group $\mathbf{A}_n$ is $(n-2)$-transitive, and the affine groups are 2-transitive [8], in contrast to e.g. the cyclic group $\mathbf{C}_n$, which is only 1-transitive.

In each of the three cases, we will prove below (i.e. in Sections 4, 5, and 6, respectively) that every $\mathrm{XU}(n)$ matrix X can be written as

$$X = \sum_{\sigma} c_\sigma G_\sigma \tag{7}$$

with all G_σ member of the appropriate group $\mathrm{G}(n)$. Because of Lemmas 1 and 2, we are then allowed to put the case that both $\sum_\sigma c_\sigma = 1$ and $\sum_\sigma |c_\sigma|^2 = 1$. For the explicit computation of the weights c_σ, we note that the $\mathrm{G}(n)$ matrices form an n-dimensional reducible representation of some abstract group $\mathbf{G}$. We assume that $\mathbf{G}$ has μ different irreducible representations. According to Lemma (29.1) of [9], because $\mathbf{G}$ is 2-transitive, the n-dimensional natural representation is the sum of the 1-dimensional trivial representation and an $(n-1)$-dimensional irreducible representation, which we will call the standard representation.

As soon as the group $\mathbf{G}$ is determined, the coefficients c_σ of the decomposition (7) can be found using the following procedure. We start by writing down a set of linear matrix equations in c_σ, involving the μ irreducible representations $D^{(\nu)}$ (with $0 \leq \nu \leq \mu - 1$) of $\mathbf{G}$ and a set of arbitrary $n_\nu \times n_\nu$ unitary matrices $U^{(\nu)}$, with the exception of $\nu = 0$ and $\nu = 1$ (see further):

$$U^{(\nu)} = \sum_{\sigma} c_\sigma D^{(\nu)}_\sigma \ . \tag{8}$$

Here, n_ν is the dimension of the νth representation $D^{(\nu)}_\sigma$ of G_σ. Note that we have μ such matrix equations (8). Each matrix eqn constitutes n_ν^2 scalar equations. We thus have a total of $\sum_{\nu=0}^{\mu-1} n_\nu^2 = N$ scalar equations with N unknowns c_σ:

$$\sum_{\sigma} c_\sigma (D^{(\nu)}(\sigma))_{k,l} = (U^{(\nu)})_{k,l} \ .$$

Thank to Schur's orthogonality theorem, this set of N linear equations with N unknowns is an invertible matrix [2]. The solution of this set of equations is:

$$\begin{aligned} c_\sigma &= \frac{1}{N} \sum_{\nu} n_\nu \sum_{i=0}^{n_\nu - 1} \sum_{j=0}^{n_\nu - 1} (D^{(\nu)}(\sigma))_{i,j} \ (U^{(\nu)}(\sigma))_{i,j} \\ &= \frac{1}{N} \sum_{\nu} n_\nu \ \mathrm{Tr} \left(D^{(\nu)}(\sigma)^\dagger \, U^{(\nu)}(\sigma) \right) \ . \end{aligned} \tag{9}$$

We choose for $\nu = 0$ the trivial representation, i.e. the 1-dimensional irreducible representation with all characters equal to 1. We choose for $\nu = 1$ the standard representation, i.e. the $(n-1)$-dimensional irreducible representation obtained by applying (1) to the permutation matrix G_σ:

$$G_\sigma = T \begin{pmatrix} 1 & \\ & D^{(1)}(\sigma) \end{pmatrix} T^{-1}$$

and thus

$$\begin{pmatrix} 1 & \\ & D^{(1)}(\sigma) \end{pmatrix} = T^{-1} G_\sigma T \ .$$

In (9), the matrix $U^{(0)}(\sigma)$ equals the 1×1 unit matrix and the matrix $U^{(1)}(\sigma)$ equals the $(n-1) \times (n-1)$ lower-right block of

$$\begin{pmatrix} 1 & \\ & U \end{pmatrix} = T^{-1} X T \ .$$

A key observation is that (among the μ matrix equations (8)), the two matrix equalities

$$\begin{aligned} U^{(0)} &= \sum c_\sigma D^{(0)}(\sigma) \\ \text{and} \quad U^{(1)} &= \sum c_\sigma D^{(1)}(\sigma) \ , \end{aligned}$$

together with the choices $U^{(0)} = 1$ and $U^{(1)} = U$, in fact mean

$$1 = \sum c_\sigma \; 1$$
$$\text{and} \;\; T^{-1}XT = \sum c_\sigma T^{-1} G_\sigma T$$

and thus suffice to guarantee $\sum c_\sigma = 1$ and $\sum c_\sigma G_\sigma = X$, respectively. As a result, for the remaining matrices $U^{(\nu)}(\sigma)$ with $2 \leq \nu \leq \mu - 1$, we are allowed to choose any unitary matrix of the right dimension n_ν. This usually allows a large number of degrees of freedom. Here, we propose two different strategies to take advantage of this freedom.

3.1. First strategy

For each matrix $U^{(\nu)}(\sigma)$ with $2 \leq \nu \leq \mu - 1$, we choose the $n_\nu \times n_\nu$ unit matrix. Then (9) becomes

$$c_\sigma = \frac{1}{N} \left[\; n_0 \operatorname{Tr}\left(D^{(0)\,\dagger}(\sigma)\right) + n_1 \operatorname{Tr}\left(D^{(1)\,\dagger}(\sigma) U\right) + \sum_{\nu=2}^{\mu-1} n_\nu \operatorname{Tr}\left(D^{(\nu)\,\dagger}(\sigma)\right) \; \right] . \tag{10}$$

We take advantage of Schur's orthogonality relation:

$$\sum_\nu n_\nu \operatorname{Tr}\left(D^{(\nu)\,\dagger}(\sigma)\right) = \sum_\nu n_\nu \operatorname{Tr}\left(D^{(\nu)\,\dagger}(\sigma)\, D^{(\nu)}(\epsilon)\right) = \delta_\sigma \, N \; ,$$

where ϵ is the identity permutation and where $\delta_\epsilon = 1$ while $\delta_\sigma = 0$ if $\sigma \neq \epsilon$. Because moreover $D^{(1)\,\dagger}(\sigma) = D^{(1)}(\sigma^{-1})$ and $n_1 = n - 1$, we obtain the explicit expression for the weight:

$$c_\sigma = \delta_\sigma + \frac{n-1}{N} \operatorname{Tr}\left(D^{(1)}(\sigma^{-1}) U\right) - \frac{n-1}{N} \chi^{(1)}(\sigma^{-1}) \; . \tag{11}$$

The number $\chi^{(\nu)}(G)$ denotes the character of the element G of the group $\mathbf{G}$ according to the νth representation. It is equal to $\operatorname{Tr}(D^{(\nu)}(G))$. In particular, we have $\operatorname{Tr}(D^{(1)}(G)) = \operatorname{Tr}(G) - 1$.

3.2. Second strategy

The second strategy is only applicable if the group $\mathbf{G}$ has an anti-standard irreducible representation, non-equivalent to the standard representation. The anti-standard representation, which we will assign the label $\nu = 2$ (if it exists), has the same characters as the standard representation (with label $\nu = 1$), except for a factor -1 if the corresponding permutation is an odd permutation. We note that, if $\mathbf{G}$ has an anti-standard representation, it also has an anti-trivial representation (a.k.a. sign representation), i.e. the 1-dimensional irrep with all characters equal to 1 (even permutations) or to -1 (odd permutations). As we thus have $n_0 = 1$ for the trivial representation, $n_1 = n - 1$ for the standard representation, $n_2 = n - 1$ for the anti-standard representation, and $n_3 = 1$ for

the anti-trivial representation, and as $N = \sum_\nu n_\nu^2$, a necessary condition for the second strategy is

$$N \geq 2 + 2(n-1)^2 \ . \tag{12}$$

As in the first strategy, we again choose the 1×1 unit matrix for $U^{(0)}(\sigma)$ and the $(n-1) \times (n-1)$ matrix U for $U^{(1)}(\sigma)$. However, in this second strategy, we also choose the matrix U for each matrix $U^{(2)}(\sigma)$. For each matrix $U^{(\nu)}(\sigma)$ with $3 \leq \nu \leq \mu - 1$, we choose the $n_\nu \times n_\nu$ unit matrix. Then (9) becomes

$$c_\sigma = \frac{1}{N} \left[\ n_0 \, \mathrm{Tr} \left(D^{(0)\,\dagger}(\sigma) \right) + n_1 \, \mathrm{Tr} \left(D^{(1)\,\dagger}(\sigma) \right) + n_2 \, \mathrm{Tr} \left(D^{(2)\,\dagger}(\sigma) \right) + \sum_{\nu=3}^{\mu-1} n_\nu \, \mathrm{Tr} \left(D^{(\nu)\,\dagger}(\sigma) \right) \ \right] . \tag{13}$$

Again taking advantage of Schur's orthogonality relation and $n_1 = n_2 = n-1$, we obtain

$$\begin{aligned} c_\sigma &= \delta_\sigma + \frac{2(n-1)}{N} \, \mathrm{Tr} \left(D^{(1)}(\sigma^{-1}) U \right) - \frac{2(n-1)}{N} \, \chi^{(1)}(\sigma^{-1}) \quad \text{if } \sigma \text{ even} \\ &= 0 \qquad\qquad \text{if } \sigma \text{ odd} \ . \end{aligned} \tag{14}$$

In the second strategy, the group $\mathbf{G} \cap \mathbf{A}_n$ thus takes over the role of $\mathbf{G}$ and $N/2$ takes over the role of N.

4. The case of arbitrary dimension n

If n is an arbitrary positive integer, then we choose for the $n \times n$ Hadamard matrix T of Section 2 the $n \times n$ discrete Fourier transform F, with entries

$$F_{k,l} = \frac{1}{\sqrt{n}} \, \omega^{kl} \ ,$$

where ω is equal to the nth root of unity. One finds [2]:

Lemma 3. *Every XU(n) matrix X can be written*

$$X = \sum_\sigma c_\sigma P_\sigma$$

with all $P_\sigma \in P(n)$.

The proof is provided by [3], by means of induction on n. Combining Lemmas 1, 2, and 3 leads to the unitary Birkhoff theorem:

Theorem 2. *Every XU(n) matrix X can be written*

$$X = \sum_{\sigma} c_\sigma P_\sigma$$

with all $P_\sigma \in P(n)$, such that both $\sum_\sigma c_\sigma = 1$ and $\sum_\sigma |c_\sigma|^2 = 1$.

4.1. First strategy

We can apply result (11) with $N = n!$. The only possible values of $\chi^{(1)}$ are $\mathrm{Tr}(\mathrm{P}_\sigma)-1$ and thus $-1, 0, 1, 2, ..., n-1$, with exception of $n-2$.

Appendix A gives an algorithm for, given an XU matrix, computing the $n!$ weights c_σ, using the algebraic language `GAP`.

4.2. Second strategy

The character tables of the groups $\mathbf{S}_2$ and $\mathbf{S}_3$ show no anti-standard representation. For $n > 3$, the group $\mathbf{S}_n$ has an anti-standard representation. In this case, we can apply result (14) with $N = n!$. The restriction $n > 3$ is not surprising, as (12) with $N = n!$ is fulfilled neither if $n = 2$ nor if $n = 3$.

5. The case of prime dimension $n = p$

We call an $n \times n$ matrix A supercirculant iff each row k equals row $k-1$ shifted x positions to the right. Thus $A_{k,l} = A_{k-1,l-x}$, where addition and subtraction are modulo n. We equivalently may write

$$A_{0,a} = A_{k,a+kx} \ .$$

We call x the pitch of the matrix. If $x = 1$, then the supercirculant matrix is called circulant; if $x = n-1$, then the supercirculant matrix is called anticirculant.

If p denotes a prime, then the $p \times p$ supercirculant permutation matrices are denoted $S_{a,x}$, where x is the pitch and a (called the shift) is the column with the unit entry in the upper row (i.e. row 0). The unit entries of such $p \times p$ permutation matrix thus are located at the p positions $(0, a)$, $(1, a+x)$, $(2, a+2x)$, ..., and $(p-1, a+(p-1)x)$, where sums are to be taken modulo p. Because x and p are co-prime, the consecutive columns with a 1, i.e. the columns a, $a+x$, $a+2x$, ..., and $a+(p-1)x$, are all different.

If n equals some prime p, then, just like in Section 4, we choose for the Hadamard matrix T the $p \times p$ discrete Fourier transform F:

$$F_{k,l} = \frac{1}{\sqrt{p}}\, \omega^{kl} \ ,$$

where ω is equal to the pth root of unity. Thus (4) becomes

$$(M_{r,s})_{k,l} = \omega^{kr-ls} \ .$$

From [3], we know that M can be written as a weighted sum of p supercirculant permutation matrices:

$$M_{r,s} = \sum_{a=0}^{p-1} (M_{r,s})_{0,a} \ S_{a,x(r,s)} \ , \tag{15}$$

where the pitch x of the matrix $S_{a,x}$ is a function of r and s. Indeed, the condition

$$(M_{r,s})_{k,a+kx} = (M_{r,s})_{0,a}$$

yields

$$kr - (a + kx)s = -as$$

and thus $r - xs = 0$. Thus x has to satisfy the eqn

$$sx = r \bmod p \ .$$

This eqn has one solution:

$$x = rs^{-1} \bmod p \ ,$$

where s^{-1} is the inverse of s modulo p. As p is prime, each non-zero integer has exactly one inverse. With $(M_{r,s})_{0,a} = \omega^{-as}$, we finally obtain

$$M_{r,s} = \sum_{a=0}^{p-1} \omega^{-as} S_{a,rs^{-1}} \ .$$

The supercirculant $p \times p$ permutation matrices form a group $\mathrm{S}(p)$, subgroup of $\mathrm{P}(p)$ (proof in Appendix B), isomorphic to the semidirect product of the cyclic group of order p and the multiplicative group of integers modulo p. The group thus is isomorphic to the semidirect product of two cyclic groups:

$$\mathbf{C}_p : \mathbf{C}_{p-1} \ ,$$

a non-Abelian group of order $p(p-1)$.

Lemma 4. *If n is prime, then every XU(n) matrix X can be written*

$$X = \sum_{\sigma} c_\sigma S_\sigma$$

with all $S_\sigma \in S(n)$.

The proof is as follows. If n is a prime p, then all matrices $M_{r,s}$ are supercirculant with a pitch $x = rs^{-1}$ modulo p. Also the van der Waerden matrix W is supercirculant, as it is circulant:

$$W = \sum_{a=0}^{n-1} \frac{1}{n} S_{a,1} \ .$$

Hence, according to (3), X is a weighted sum of supercirculant permutation matrices.

Combining Lemmas 1, 2, and 4 leads to

Theorem 3. *If n is prime, then every XU(n) matrix X can be written*

$$X = \sum_{\sigma} c_\sigma S_\sigma$$

with all $S_\sigma \in S(n)$, such that both $\sum_\sigma c_\sigma = 1$ and $\sum_\sigma |c_\sigma|^2 = 1$.

5.1. First strategy

We can apply result (11) with $N = p(p-1)$. The only possible values of $\chi^{(1)}$ are -1, 0, and $p-1$, as demonstrated in Appendix C. Thus we find a unitary Birkhoff decomposition with only $p(p-1)$ terms. For a prime exceeding 3, this number is substantially smaller than the number $p!/2$ of Subsection 4.2. The resulting unitary Birkhoff theorem is also slightly stronger than the theorem in [3], where the Birkhoff decomposition consists of p^2 terms.

5.2. Second strategy

The group S(2), isomorphic to the cyclic group $\mathbf{C}_2$, has only two irreducible representations: the trivial one and the standard one. Also the group S(n) with n equal to an odd prime p, has no inequivalent anti-standard representation. Indeed, because all odd supercirculant permutations have non-unit pitch (see Appendix D) and thus have unit trace (see Appendix C) and hence have zero character $\chi^{(1)}$, all characters of the anti-standard representation equal the corresponding characters of the standard representation. Therefore, the standard and anti-standard representations are equivalent. We conclude that we cannot apply the second strategy of Subsection 3.2. The absence of any inequivalent anti-standard representation is no surprise, as $N = n(n-1)$ does not satisfy (12).

6. The case of prime-power dimension $n = p^w$

For $n = p^w$ with arbitrary positive w, we can choose for T of Section 2 the Kronecker product of w small (i.e. $p \times p$) Fourier matrices F:

$$T = F \otimes F \otimes ... \otimes F = F^{\otimes w} \ .$$

The $n \times n$ matrix T has following entries:

$$T_{a,b} = \frac{1}{\sqrt{n}} \ \omega^{f(a,b)} \ ,$$

where $f(x, y)$ is the sum of the ditwise product of the p-ary numbers x and y:

$$f(x, y) = \sum_j x_j y_j \bmod p \ .$$

As a consequence, we have

$$(M_{r,s})_{k,l} = \omega^{f(k,r)-f(s,l)} \ . \tag{16}$$

Among the n^2 entries of this matrix, n^2/p are equal to 1, n^2/p are equal to ω, ..., and n^2/p are equal to ω^{p-1}.

Remark 2. For sake of convenience, below, the rows and the colums of a matrix will sometimes be pointed at, not by a number, but instead by a vector. This will allow matrix computations for the row and column numbers. For this purpose, any number $z = z_0 + z_1 p + z_2 p^2 ... + z_{w-1} p^{w-1}$ has an associated boldfaced $w \times 1$ vector $\mathbf{z} = (z_0, z_1, z_2, ..., z_{w-1})^T$, consisting of the w dits of the number z.

We call a matrix A epicirculant if row k equals row 0, 'shifted to the right' according to

$$A_{\mathbf{0},\mathbf{a}} = A_{\mathbf{k},\,\mathbf{a}+\mathbf{x}\mathbf{k}} \ ,$$

where $\mathbf{a}$ is the $w \times 1$ vector associated with the column number a and where $\mathbf{x}$ is a $w \times w$ matrix called the pitch matrix, consisting of w^2 entries, all $\in \{0, 1, ..., p-1\}$. A matrix of the form (16) is automatically epicirculant. It is a weighted sum of epicirculant permutation matrices E: we have

$$M_{r,s} = \sum_{a=0}^{p-1} (M_{r,s})_{0,a} \ E_{\mathbf{a},\mathbf{x}(r,s)} \ . \tag{17}$$

Here, $\mathbf{x}$ is an appropriate $w \times w$ pitch matrix, depending on r and s. Proof is in Appendix E. We note that vector $\mathbf{a}$ and matrix $\mathbf{x}$ constitute a pair, fully specifying an affine transformation [10].

If n is a prime power, say $n = p^w$, then the epicirculant $p^w \times p^w$ permutation matrices form a group E(n), subgroup of P(n) (proof in Appendix F), isomorphic to the general

affine group $\mathrm{GA}(w,p)$, a semidirect product of the direct product of cyclic groups of order p and the general linear group $\mathrm{GL}(w,p)$:

$$\mathrm{GA}(w,p) = \mathbf{C}_p^w : \mathrm{GL}(w,p)$$

of order

$$p^w(p^w-1)(p^w-p)(p^w-p^2)...(p^w-p^{w-1}) \ . \tag{18}$$

We note that $\mathrm{GA}(w,p)$ is a maximal subgroup of the symmetric group $\mathbf{S}_{p^w}$ (O'Nan–Scott theorem) [11].

Each of the w subgroups $\mathbf{C}_p$ consists of p matrices, each a Kronecker product with a total of w factors:

$$I \otimes I \otimes ... \otimes I \otimes M \otimes I \otimes ... \otimes I \ ,$$

where I denotes the $p \times p$ unit matrix and M a $p \times p$ circulant permutation matrix $S_{a,1}$.

Lemma 5. *If n is a prime power, then every XU(n) matrix X can be written*

$$X = \sum_\sigma c_\sigma E_\sigma$$

with all $E_\sigma \in E(n)$.

The proof is as follows. If n is a prime power p^w, then all matrices $M_{r,s}$ are epicirculant with an invertible pitch matrix $\mathbf{x}$. Also the van der Waerden matrix W is epicirculant, as it is circulant:

$$W = \sum_{a=0}^{n-1} \frac{1}{n} \ E_{\mathbf{a},\mathbf{1}} \ ,$$

where the pitch matrix $\mathbf{1}$ denotes the $w \times w$ unit matrix. Hence, according to (3), X is a weighted sum of epicirculant permutation matrices.

Combining Lemmas 1, 2, and 5 leads to

Theorem 4. *If n is a prime power, then every XU(n) matrix X can be written*

$$X = \sum_\sigma c_\sigma E_\sigma$$

with all $E_\sigma \in E(n)$, such that both $\sum_\sigma c_\sigma = 1$ and $\sum_\sigma |c_\sigma|^2 = 1$.

6.1. First strategy

We can apply result (11) with N given by (18). The only possible values of $\chi^{(1)}$ are -1, 0, $p-1$, p^2-1, p^3-1, ..., and p^w-1, as demonstrated in Appendix G.

6.2. Second strategy

For $w > 1$ and $p > 2$, the general affine groups have, besides the standard representation, also an inequivalent anti-standard representation. For a proof, it suffices to point to a single example of an odd epicirculant permutation matrix with trace different from unity. We choose the $p^w \times p^w$ matrix

$$E = I \otimes I \otimes ... \otimes I \otimes M \ ,$$

i.e. the Kronecker product of $w-1$ matrices I (i.e. the $p \times p$ unit matrix) and the $p \times p$ supercirculant matrix $M = S_{0,q}$. The $w \times w$ pitch matrix associated with E is the diagonal matrix $\mathrm{diag}(q, 1, 1, ..., 1)$.

On the one hand, we have the following property of the Kronecker product of two square matrices:

$$\mathrm{Det}(A \otimes B) = [\,\mathrm{Det}(A)\,]^{\dim(B)}\ [\,\mathrm{Det}(B)\,]^{\dim(A)} \ . \tag{19}$$

Therefore, we have $\mathrm{Det}(E) = \mathrm{Det}(M)^{(p^{w-1})}$. We choose the number q such that $\mathrm{Det}(M) = -1$ and thus $\mathrm{Det}(E) = -1$. This is always possible. Suffice it to choose q equal to $g(p)$, where g is a generator of the modulo p multiplication group [12]. Unfortunately, there is no algorithm known for finding such generator except brute force [13]. Nevertheless, we can prove that $\mathrm{Det}(S_{0,g(p)}) = -1$, without a priori knowing the value of $g(p)$: see Appendix D.

On the other hand, we have $\mathrm{Tr}(E) = p^{w-1}\,\mathrm{Tr}(M) = p^{w-1}\,1 = p^{w-1}$. Because $w > 1$, we have $\mathrm{Tr}(E) > 1$ and thus $\chi^{(1)} > 0$. We thus conclude that we can apply result (14) with N according to (18).

The above reasoning is not valid for $p = 2$, because, in that case, $\mathrm{Det}(M) = -1$ does not imply $\mathrm{Det}(E) = -1$. For the case $p = 2$, we will prove that all $2^w \times 2^w$ epicirculant matrices are even permutations. For this purpose, it is sufficient to demonstrate that all group generators are even. From reversible computing [14] [15] [16], it is known that the group $\mathrm{GA}(w, 2)$ is generated by following matrices:

$$A = I \otimes I \otimes ... \otimes I \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes I \otimes I \otimes ... \otimes I$$

Table 1
Applicability of the second strategy for the Birkhoff decomposition of an XU(n) matrix with $n = p^w$.

	$p = 2$	$p \geq 3$
$w = 1$	no	no
$w = 2$	yes	yes
$w \geq 3$	no	yes

$$B = I \otimes I \otimes ... \otimes I \otimes \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \otimes I \otimes I \otimes ... \otimes I$$

$$C = I \otimes I \otimes ... \otimes I \otimes \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \otimes I \otimes I \otimes ... \otimes I \ ,$$

with a total of $w-1$ (for A) or $w-2$ (for B and C) factors I. In the context of computing, these matrices represent `NOT` gates, respectively controlled `NOT` gates. Applying (19), we have:

$$\mathrm{Det}(A) = [\ \mathrm{Det} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}]^{(p^{w-1})} = (-1)^{2^{w-1}} = 1$$

$$\mathrm{Det}(B) = [\ \mathrm{Det} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}]^{(p^{w-2})} = (-1)^{2^{w-2}} = 1$$

$$\mathrm{Det}(C) = [\ \mathrm{Det} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}]^{(p^{w-2})} = (-1)^{2^{w-2}} = 1 \ ,$$

except if $w = 2$. Thus, for $w > 2$, all members of GA$(w, 2)$ represent even permutations and the second strategy (Subsection 3.2) is not applicable.

This leaves us with the case $p = 2$ and $w = 2$. The epicirculant matrices form a group E(4) isomorphic to the symmetric group $\mathbf{S}_4$. As stated in Section 4.2, the second strategy is applicable. The results on the applicability of the second strategy are summarized in Table 1.

Table 2
Number of Birkhoff terms in the decomposition of an arbitrary $n \times n$ unit-linesum unitary matrix.

n	1	2	3	4	5	6	7	8	9	10	11
	1	2	6	12	20	360	42	1,344	216	1,814,400	110

n	12	13	14	15	16	17
	239,500,800	156	43,589,145,600	653,837,184,000	322,560	272

7. Conclusion

According to [2], every unit-linesum $n \times n$ unitary matrix can be decomposed as a weighted sum of the $n \times n$ permutation matrices, such that both the sum of the weights and the sum of the squared moduli of the weights equal unity. Such Birkhoff sum contains $n!$ terms. In the present paper, we demonstrate the following:

- If $n \geq 4$, then $n!/2$ terms suffice.
- If $n = p^w$ with p an arbitrary prime and w an arbitrary integer, then $p^w(p^w - p^{w-1})(p^w - p^{w-2})...(p^w - p)(p^w - 1)$ suffice.
- If $n = p^w$ with p an arbitrary odd prime and w an integer ≥ 2, then $p^w(p^w - p^{w-1})(p^w - p^{w-2})...(p^w - p)(p^w - 1)/2$ suffice.

For numerical examples, see Table 2.

We see that in some cases the upper bounds on the number of Birkhoff terms are quite small. This triggers the question whether anything can be said about lower bounds. Because the $n!$ permutation matrices of dimension n are member of XU(n) and because there exist $(n-1)^2+1$ linear independent $n \times n$ permutation matrices [17], any Birkhoff decomposition of an arbitrary XU(n) matrix needs at least $(n-1)^2+1$ terms. Hence, $(n-1)^2+1$ is a lower bound on the number of Birkhoff terms. This amount also appears in discussions on the classical Birkhoff theorem, e.g. in the Marcus–Ree theorem [18] [19] [20]. As some numbers in Table 2 are far in excess of the Marcus–Ree number, there is still room for better upper bounds. In particular, the case of n equal to the product of two different primes is left for further investigation. Another subject of deeper investigation is the geometric interpretation of the results. The classical Birkhoff theorem on doubly stochastic matrices has a clear interpretation, i.e. a polytope within an $(n-1)^2$-dimensional Euclidean space, with the $n!$ permutation matrices as corners [21]. The line segments between two corners form straight edges. In contrast, the Birkhoff theorems on unitary matrices lead to a closed curve between each two permutation matrices [2] [22], within a compact $(n-1)^2$-dimensional curved space.

Declaration of Competing Interest

There is no competing interest.

Appendix A. Algorithm

We present here a GAP program that, given an $n \times n$ XU matrix, computes the N weights c_σ of the Birkhoff decomposition, according to Subsection 4.1 and applying eqn (11):

```
# Introducing the matrix XU

XU := (1/4) * [[ 3-1*E(4) , -1+1*E(4) ,  1-1*E(4) ,  1+1*E(4) ],
               [ 1-1*E(4) ,  3-1*E(4) ,  1+1*E(4) , -1+1*E(4) ],
               [ 0+0*E(4) ,  0+0*E(4) ,  2+2*E(4) ,  2-2*E(4) ],
               [ 0+2*E(4) ,  2+0*E(4) ,  0-2*E(4) ,  2+0*E(4) ]];
dim := DimensionsMat(XU); n := dim[1];

# Generating the group G

permutation1 := (1,2);
ll := List([1..n], x -> x+1); ll[n] := 1;
permutation2 := PermList(ll);
generator1 := PermutationMat( permutation1, n);
generator2 := PermutationMat( permutation2, n);
G := Group(generator1, generator2); N := Order(G);

# Introducing the constant matrix T

T := NullMat(n,n);
for row    in [1..n] do
for column in [1..n] do
T[row][column] := (1/Sqrt(n)) * E(n)^( (row-1)*(column-1) ); od; od;

# Defining the standard irreducible representation

StandardIrrep := function(matriks)
local A,a;
A := T * matriks * T^(-1);
a := NullMat(n-1,n-1);
for row    in [1..(n-1)] do
for column in [1..(n-1)] do
a[row][column] := A[row+1][column+1]; od; od;
return a;
end;
```

```
# Computing the matrix U and the weights c

U := StandardIrrep(XU);

PermuList := [];
IrrepList := [];
for j in [1..N]
do Append(PermuList, [              Elements(G)[j] ]);
   Append(IrrepList, [StandardIrrep(Elements(G)[j])]); od;

WeightList := [];
for j in [1..N]
do if PermuList[j]=IdentityMat(n) then C1 := 1; else C1 := 0; fi;
   C2 := ((n-1)/N) * Trace(IrrepList[j]^(-1)*U);
   C3 := ((n-1)/N) * Trace(IrrepList[j]^(-1)  );
   Append(WeightList, [C1 + C2 - C3]); od;

# Checking the properties of the weights c

check1 := 0; check2 := 0; check3 := NullMat(n,n);

for j in [1..N]
do check1 := check1 + WeightList[j];
   check2 := check2 + WeightList[j] * ComplexConjugate(WeightList[j]);
   check3 := check3 + WeightList[j] * PermuList[j]; od;

check1 = 1; check2 = 1; check3 = XU;

# Presenting the weights c

output := WeightList;
```

Above, the two permutation matrices `generator1` and `generator2` generate the group $\mathbf{G}$, equal to the symmetric group $\mathbf{S}_n$. For Subsections 4.2, 5.1, 6.1, and 6.2, similar algorithms can be constructed, applying appropriate generators, the appropriate matrix T, and the appropriate formula for the weights (i.e. either eqn (11) or eqn (14)).

Appendix B. The group of supercirculant permutation matrices

The supercirculant $n \times n$ permutation matrices form a group. Indeed, the product of two such matrices (say $S_{a,x}$ and $S_{b,y}$) yields a third such matrix. In order to prove this fact, we compute the matrix entry at position (u, v):

$$\begin{aligned}(S_{a,x}\ S_{b,y})_{u,v} &= \sum_f (S_{a,x})_{u,f}\ (S_{b,y})_{f,v} \\ &= \sum_f \delta_{f,\,a+ux}\ \delta_{v,\,b+fy} \\ &= \delta_{v,\,b+(a+ux)y} \\ &= \delta_{v,\,b+ay+uxy} = (S_{b+ay,\,xy})_{u,v}\end{aligned}$$

and hence

$$S_{a,x}\ S_{b,y} = S_{b+ay,\,xy}\ . \tag{20}$$

If n is a prime p, each non-zero number x has an inverse number x^{-1}. Applying (20), we find

$$S_{a,x}\ S_{-ax^{-1},\,x^{-1}} = S_{0,1}\ .$$

The right-hand side being the $p\times p$ unit matrix, the result proves that each supercirculant matrix has an inverse matrix that also is supercirculant:

$$(S_{a,x})^{-1} = S_{-ax^{-1},\,x^{-1}}\ .$$

We conclude by considering two applications of eqn (20):

- choosing $x = y = 1$ leads to

  $$S_{a,1}\ S_{b,1} = S_{a+b,\,1}$$

  illustrating that the p matrices $S_{a,1}$ are isomorphic to the addition modulo p;
- choosing $a = b = 0$ leads to

  $$S_{0,x}\ S_{0,y} = S_{0,\,xy}$$

  illustrating that the $p-1$ matrices $S_{0,x}$ are isomorphic to the multiplication modulo p.

Each supercirculant matrix can be decomposed as the product of a zero-shift matrix and a unit-pitch matrix:

$$\begin{aligned}S_{a,x} &= S_{0,x}\,S_{a,1} \\ &= S_{ax^{-1},1}\,S_{0,x}\ .\end{aligned}$$

Appendix C. The trace of a supercirculant permutation matrix

We compute the trace of the supercirculant permutation matrix $S_{a,x}$:

$$\mathrm{Tr}(S_{a,x}) = \sum_u (S_{a,x})_{u,u} = \sum_u \delta_{u,\, a+ux} \ .$$

If the eqn

$$u(1-x) = a$$

is fulfilled, then the corresponding number u points to a unit entry in position (u,u) of the matrix $S_{a,x}$. We notice:

- If $x \neq 1$, then $u = a(1-x)^{-1}$ is the one and only solution;
- if $x = 1$ and $a \neq 0$, then the eqn has no solution u;
- if $x = 1$ and $a = 0$, then u may have any value from $\{0,\ 1,\ 2,\ ...,\ p-1\}$.

Thus we conclude:

- $\mathrm{Tr}(S_{a,x}) = 1$, if $x \neq 1$,
- $\mathrm{Tr}(S_{a,1}) = 0$, if $a \neq 0$, and
- $\mathrm{Tr}(S_{0,1}) = p$.

Appendix D. The determinant of a supercirculant permutation matrix

As mentioned in Appendix B, each supercirculant matrix can be decomposed as follows:

$$S_{a,x} = S_{0,x}\, S_{a,1} \ .$$

Hence:

$$\mathrm{Det}(S_{a,x}) = \mathrm{Det}(S_{0,x})\, \mathrm{Det}(S_{a,1}) \ .$$

We have $S_{a,1} = (S_{1,1})^a$ and therefore $\mathrm{Det}(S_{a,1}) = (\mathrm{Det}(S_{1,1}))^a$. If p is odd, then $\mathrm{Det}(S_{1,1}) = 1$, such that $\mathrm{Det}(S_{a,1}) = 1$. In other words: for odd primes, all of the p circulant permutation matrices have unit determinant. The situation is different for the $p-1$ matrices $S_{0,x}$. Half of them have unit determinant and half of them have determinant equal to -1. In order to prove this fact, the key observation is the fact that the cyclic group is Abelian; so there exists a similarity transformation that diagonalizes *all* matrices $S_{0,x}$. We now prove that the following matrix F serves our purpose:

$$F_{u,v} = \begin{cases} 1 & \text{if } u = v = 0 \\ 0 & \text{if } u = 0 \text{ and } v \neq 0 \\ 0 & \text{if } u \neq 0 \text{ and } v = 0 \\ \frac{\omega^{v\varphi(u)}}{\sqrt{p-1}} & \text{if } u \neq 0 \text{ and } v \neq 0 \ , \end{cases}$$

where $\omega = \exp(\frac{2\pi i}{p-1})$ is the $(p-1)$th root of unity, and the function $\varphi(a)$ gives the 'position' of the number a in the cyclic group $\mathbf{C}_{p-1}$ (multiplicative group modulo p), as a power of the (a priori unknown) generator g, i.e.

$$a = g^{\varphi(a)} \ .$$

From this definition, the following interesting properties of φ can be deduced:

$$\begin{aligned} \varphi(1) &= 0 \\ \varphi(g) &= 1 \\ \varphi(ab) &= \varphi(a) + \varphi(b) \ . \end{aligned}$$

These properties are key in the following derivation. We compute the similarity transformation given by $F^\dagger S_{0,x} F$. Because both F and $S_{0,x}$ are block diagonal with a single 1 in the upper-left corner, we only need to compute the lower-right part:

$$\begin{aligned} (F^\dagger S_{0,x} F)_{u,v} &= \sum_{k=1}^{p-1} \sum_{l=1}^{p-1} \overline{F_{k,u}} \, (S_{0,x})_{k,l} \, F_{l,v} \\ &= \frac{1}{p-1} \sum_{k=1}^{p-1} \sum_{l=1}^{p-1} \omega^{-u\varphi(k)} \, \delta_{l,xk} \, \omega^{v\varphi(l)} \\ &= \frac{1}{p-1} \sum_{k=1}^{p-1} \omega^{-u\varphi(k)+v\varphi(xk)} \\ &= \frac{1}{p-1} \sum_{k=1}^{p-1} \omega^{-u\varphi(k)+v\varphi(x)+v\varphi(k)} \\ &= \omega^{v\varphi(x)} \, \delta_{u,v} \ . \end{aligned}$$

This result leads to two conclusions:

- By choosing $x = 1$, we find that $(F^\dagger F)_{u,v} = \delta_{u,v}$ and thus that F is unitary.
- By choosing x arbitrary, we find that the matrix $S_{0,x}$ has the eigenvalues $\omega^{v\varphi(x)}$ plus an additional 1 from the upper-left matrix block.

The determinant is just the product of all eigenvalues:

$$\begin{aligned}\mathrm{Det}(S_{0,x}) &= \prod_{v=1}^{p-1} \omega^{v\varphi(x)} = \omega^{\varphi(x)\sum_{v=1}^{p-1} v} = \omega^{\varphi(x)\,\frac{p(p-1)}{2}} \\ &= e^{\frac{2\pi i}{p-1}\,\varphi(x)\,\frac{p(p-1)}{2}} = e^{\pi i \varphi(x) p}\ .\end{aligned}$$

Now, if p is an odd prime, then $e^{\pi i p} = -1$, such that $\mathrm{Det}(S_{0,x}) = (-1)^{\varphi(x)}$, which proves that the sign of the determinant of $S_{0,x}$ alternates in the chain of successive elements of $\mathbf{C}_{p-1}$. More in particular, the position of $x = g$ always is $\varphi(g) = 1$, so we have $\mathrm{Det}(S_{0,g}) = -1$.

We note that the above results for both $S_{a,1}$ and $S_{0,x}$ are only valid for odd primes p. If p is even, i.e. if $p = 2$, then there exist only two supercirculant matrices $S_{0,1} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, with determinant equal to 1, and $S_{1,1} = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, with determinant equal to -1.

Appendix E. The pitch matrix

In (17), the epicirculant matrix $E_{\mathbf{a},\mathbf{x}}$ needs a unit entry in position $(\mathbf{k}, \mathbf{a} + \mathbf{x}\mathbf{k})$ if

$$(M_{r,s})_{\mathbf{k},\mathbf{a}+\mathbf{x}\mathbf{k}} = (M_{r,s})_{\mathbf{0},\mathbf{a}}$$

implying

$$f(k,r) - f(s, a + \sum_u \sum_v x_{u,v} k_v p^u) = f(0,r) - f(s,a)$$

or

$$\sum_j k_j r_j - \sum_j s_j \left(a_j + (\sum_u \sum_v x_{u,v} k_v p^u)_j \right) = -\sum s_j a_j$$

and thus

$$\sum_j s_j \sum_v x_{j,v} k_v = \sum_j k_j r_j$$

or

$$\sum_v k_v \sum_j x_{j,v} s_j = \sum_v k_v r_v$$

and thus

$$\sum_v k_v \, (\sum_j x_{j,v} s_j - r_v) = 0\ .$$

We fulfill this condition by the set of w non-coupled eqns

$$\sum_j s_j x_{j,v} = r_v \ . \tag{21}$$

For each eqn, we expect p^{w-1} solutions (as we can choose $w-1$ out of the w dits $x_{j,v}$ arbitrarily from $\{0,1,...,p-1\}$). However, many solutions have to be rejected. Indeed, each column of the matrix $E_{\mathbf{a},\mathbf{x}}$ in (17) should contain one and only one unit entry. For this purpose, it is necessary and sufficient that the matrix $\mathbf{x}$ is invertible. Proof is as follows. We require that for any two different row numbers ($k' \neq k$) the unit entry of the permutation matrix is in another column:

$$\mathbf{a} + \mathbf{x}\mathbf{k}' \neq \mathbf{a} + \mathbf{x}\mathbf{k}$$

and thus $\mathbf{x}(\mathbf{k}' - \mathbf{k}) \neq \mathbf{0}$. This requires that for any non-zero number K we have

$$\mathbf{x}\mathbf{K} \neq \mathbf{0} \ .$$

This, in turn, requires that the rows of $\mathbf{x}$ are linearly independent and thus that the matrix $\mathbf{x}$ is invertible.

We now prove that, for any pair (r,s), the set (21) has at least one acceptable solution, i.e. a solution such that the matrix $\mathbf{x}$ is invertible. Indeed:

- Because both r and s are non-zero, at least one dit r_u is non-zero and at least one dit s_j is non-zero. Let r_α be the least-significant non-zero dit of r; let s_β be the least-significant non-zero dit of s.
- We choose all dits $x_{j,v} = 0$, except the dits $x_{v,v}$, $x_{\beta,v}$, and $x_{\alpha,\beta}$. Thus eqns (21) become

$$\begin{aligned} s_v x_{v,v} + s_\beta x_{\beta,v} &= r_v \mod p \text{ if } v \neq \beta \\ s_\alpha x_{\alpha,\beta} + s_\beta x_{\beta,\beta} &= r_\beta \mod p \ . \end{aligned} \tag{22}$$

- For $v \neq \alpha$ and $v \neq \beta$, we choose $x_{v,v} = 1$. Further we choose $x_{\alpha,\alpha} = 0$ and $x_{\alpha,\beta} = 1$. Thus eqns (22) become

$$\begin{aligned} s_\beta x_{\beta,v} &= r_v - s_v \mod p \ \text{ if } v \neq \alpha \ \text{ and } v \neq \beta \\ s_\beta x_{\beta,\alpha} &= r_\alpha \mod p \\ s_\beta x_{\beta,\beta} &= r_\beta - s_\alpha \mod p \end{aligned} \tag{23}$$

  which lead to a single solution set $x_{\beta,v}$.

The resulting pitch matrix $\mathbf{x}$ consists of a non-zero diagonal, one non-zero row, and one extra unit entry. E.g. for $w = 7$, $\alpha = 2$, and $\beta = 4$, we have:

$$\begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & 0 & & 1 & & \\ & & & 1 & & & \\ x_{4,0} & x_{4,1} & x_{4,2} & x_{4,3} & x_{4,4} & x_{4,5} & x_{4,6} \\ & & & & & 1 & \\ & & & & & & 1 \end{pmatrix} .$$

We note that here $\mathrm{Det}(\mathbf{x})$ equals $x_{4,2}$. In general, we have

$$\mathrm{Det}(\mathbf{x}) = \pm\, x_{\beta,\alpha} = \pm\, r_\alpha\, s_\beta^{-1} \ .$$

Because $\mathrm{Det}(\mathbf{x}) \neq 0$, we have that $\mathbf{x}$ is invertible.

Appendix F. The group of epicirculant permutation matrices

The epicirculant permutation matrices form a group. An arbitrary entry (at location $(\mathbf{k},\mathbf{l})$) of such matrix $E_{\mathbf{a},\mathbf{x}}$ is $\delta_{\mathbf{l},\,\mathbf{a}+\mathbf{xk}}$. The product of two such matrices yields a third such matrix. Indeed:

$$\begin{aligned} (E_{\mathbf{a},\mathbf{x}}\ E_{\mathbf{b},\mathbf{y}})_{u,v} &= \sum_f (E_{\mathbf{a},\mathbf{x}})_{u,f}\ (E_{\mathbf{b},\mathbf{y}})_{f,v} \\ &= \sum_f \delta_{\mathbf{f},\,\mathbf{a}+\mathbf{xu}}\ \delta_{\mathbf{v},\,\mathbf{b}+\mathbf{yf}} \\ &= \delta_{\mathbf{v},\,\mathbf{b}+\mathbf{ya}+\mathbf{yxu}} \\ &= (E_{\mathbf{b}+\mathbf{ya},\,\mathbf{yx}})_{u,v} \end{aligned}$$

and hence

$$E_{\mathbf{a},\mathbf{x}}\ E_{\mathbf{b},\mathbf{y}} = E_{\mathbf{b}+\mathbf{ya},\,\mathbf{yx}} \ .$$

Straightforward application of this result leads to

$$E_{\mathbf{a},\mathbf{x}}\ E_{-\mathbf{x}^{-1}\mathbf{a},\,\mathbf{x}^{-1}} = E_{\mathbf{0},\mathbf{1}} \ .$$

The right-hand side being the $p^w \times p^w$ unit matrix, the result proves that each epicirculant matrix has an inverse matrix that also is epicirculant:

$$(E_{\mathbf{a},\mathbf{x}})^{-1} = E_{-\mathbf{x}^{-1}\mathbf{a},\,\mathbf{x}^{-1}} \ .$$

Each epicirculant matrix can be decomposed as the product of a matrix with zero shift vector $\mathbf{a}$ and a matrix with unit pitch matrix $\mathbf{x}$:

$$\begin{aligned} E_{\mathbf{a},\mathbf{x}} &= E_{\mathbf{0},\mathbf{x}}\, E_{\mathbf{a},\mathbf{1}} \\ &= E_{\mathbf{x}^{-1}\mathbf{a},\mathbf{1}}\, E_{\mathbf{0},\mathbf{x}} \ . \end{aligned}$$

Appendix G. The trace of an epicirculant permutation matrix

We compute the trace of the epicirculant permutation matrix $E_{\mathbf{a},\mathbf{x}}$:

$$\mathrm{Tr}(E_{\mathbf{a},\mathbf{x}}) = \sum_u (E_{\mathbf{a},\mathbf{x}})_{u,u} = \sum_u \delta_{\mathbf{u},\,\mathbf{a}+\mathbf{x}\mathbf{u}} \ .$$

If the eqn

$$(\mathbf{1}-\mathbf{x})\mathbf{u} = \mathbf{a}$$

is fulfilled, then the corresponding number u points to a unit entry in position (u,u) of the matrix $E_{\mathbf{a},\mathbf{x}}$. Here, $\mathbf{1}$ denotes the $w \times w$ unit matrix. We notice:

- If $(\mathbf{1}-\mathbf{x})$ is invertible, then $\mathbf{u} = (\mathbf{1}-\mathbf{x})^{-1}\mathbf{a}$ is the one and only solution;
- if $(\mathbf{1}-\mathbf{x}) = \mathbf{0}$ and $\mathbf{a} \neq \mathbf{0}$, then the eqn has no solutions $\mathbf{u}$;
- if $(\mathbf{1}-\mathbf{x}) = \mathbf{0}$ and $\mathbf{a} = \mathbf{0}$, then u may have any value from $\{0, 1, 2, ..., p^w - 1\}$;
- if $(\mathbf{1}-\mathbf{x})$ is neither invertible nor zero, then $(\mathbf{1}-\mathbf{x})$ has rank λ with $1 \leq \lambda \leq w-1$ and $\mathbf{u}$ can have as many values as there are solutions of the eqn $(\mathbf{1}-\mathbf{x})\mathbf{u} = \mathbf{0}$, i.e. as the size of the kernel of $(\mathbf{1}-\mathbf{x})$, i.e. $p^{w-\lambda}$.

Thus we conclude:

- $\mathrm{Tr}(E_{\mathbf{a},\mathbf{1}}) = 0$, if $\mathbf{a} \neq \mathbf{0}$,
- $\mathrm{Tr}(E_{\mathbf{0},\mathbf{1}}) = p^w$, and
- $\mathrm{Tr}(E_{\mathbf{a},\mathbf{x}}) = p^{w-\lambda}$, if $(\mathbf{1}-\mathbf{x})$ has rank $\lambda \neq 0$.

References

[1] G. Birkhoff, Tres observaciones sobre el algebra lineal, Universidad Nacional de Tucumán: Revista Matemáticas y Física Teórica 5 (1946) 147–151.
[2] S. De Baerdemacker, A. De Vos, L. Chen, L. Yu, The Birkhoff theorem for unitary matrices of arbitrary dimension, Linear Algebra Appl. 514 (2017) 151–164.
[3] A. De Vos, S. De Baerdemacker, The Birkhoff theorem for unitary matrices of prime dimension, Linear Algebra Appl. 493 (2016) 455–468.
[4] A. De Vos, S. De Baerdemacker, The NEGATOR as a basic building block for quantum circuits, Open Syst. Inf. Dyn. 20 (2013) 1350004.
[5] A. De Vos, S. De Baerdemacker, On two subgroups of U(n), useful for quantum computing, in: Proceedings of the 30th International Colloquium on Group-Theoretical Methods in Physics, Gent, July 2014, J. Phys. 597 (2015) 012030.
[6] A. Klappenecker, M. Rötteler, Quantum software reusability, Internat. J. Found. Comput. Sci. 14 (2003) 777–796.
[7] W. Tadej, K. Życzkowski, A concise guide to complex Hadamard matrices, Open Syst. Inf. Dyn. 13 (2006) 133–177.

[8] mathworld.wolfram.com/TransitiveGroup.html, 2018.
[9] M. Burrow, Representation Theory of Finite Groups, Dover, New York, 1965.
[10] Wikipedia, Affine group, https://wikipedia.org/wiki/Affine_group, 2018.
[11] M. Liebeck, C. Praeger, J. Saxl, A classification of the maximal subgroups of the finite alternating and symmetric groups, J. Algebra 111 (1987) 365–383.
[12] mathworld.wolfram.com/ModuloMultiplicationGroup.html, 2018.
[13] K. Conrad, Cyclicity of $(\mathrm{Z}/(p))^{\times}$, http://www.math.uconn.edu/~kconrad/blurbs/grouptheory/cyclicmodp.pdf, 2018.
[14] T. Beth, M. Rötteler, Quantum algorithms: applicable algebra and quantum physics, in: G. Alber, T. Beth, M. Horodecki, P. Horodecki, R. Horodecki, M. Rötteler, H. Weinfurter, R. Werner, A. Zeilinger (Eds.), Quantum Information, Springer Verlag, Berlin, 2001, pp. 96–150.
[15] K. Patel, I. Markov, J. Hayes, Optimal synthesis of linear reversible circuits, Quantum Inf. Comput. 8 (2008) 282–294.
[16] A. De Vos, Reversible Computing, Wiley–VCH, Weinheim, 2010.
[17] H. Farahat, Sets of linearly independent permutation matrices, J. Lond. Math. Soc. 2 (1970) 696–698.
[18] M. Marcus, R. Ree, Diagonals of doubly stochastic matrices, Q. J. Math. 10 (1959) 296–302.
[19] R. Bruali, Notes on the Birkhoff algorithm, Canad. Math. Bull. 25 (1982) 191–199.
[20] F. Dufossé, B. Uçar, Notes on the Birkhoff-von Neumann Decomposition of Doubly Stochastic Matrices, Research Report, 8852 Institut National de Recherche en Informatique et en Automatique, 2016.
[21] I. Bengtsson, A. Ericsson, M. Kuś, W. Tadej, K. Życzkowski, Birkhoff's polytope and unistochastic matrices, $N = 3$ and $N = 4$, Comm. Math. Phys. 259 (2005) 307–324.
[22] A. De Vos, S. De Baerdemacker, From reversible computation to quantum computation by Lagrange interpolation, unpublished, https://arxiv.org/abs/1502.00819, 2015.